The Abolition of the Button

A Treatise on the Conservation of Interface Work

Flyxion
Independent Researcher

October 2026

Preface: On the Status of the Main Hypothesis

This book defends a hypothesis that it does not prove. The hypothesis, stated with some care in Chapter 1 and referred to throughout as the Main Hypothesis, is that the Interface Elimination Thesis is logically untenable. The Interface Elimination Thesis is the view, implicit in a great deal of public and commercial discussion of direct neural coupling, that the discretization, interpretation, authorization and error control performed by ordinary interfaces are overhead imposed by an inconvenient physical intermediary, and that this overhead will vanish once the intermediary is removed. The reader should know at the outset that the following pages establish the functional half of this claim, which is very nearly a tautology, and nothing that establishes the comparative half, which is the half that matters to the proponents of the thesis. What the pages establish, under assumptions that are stated and in several cases not defended, is a family of weaker propositions about information, cost, drift, regress and authorization. Each of these propositions is compatible with the Interface Elimination Thesis being true, false, or a matter of degree. The distance between those propositions and the Main Hypothesis is not crossed in any chapter, and each chapter closes with a short statement recording that it has not been crossed there either.

The arrangement is deliberate. A treatise that proved its thesis would invite the reader to accept it, whereas a treatise that announces its failure to prove its thesis invites the reader to check the mathematics, which is the more useful invitation. The mathematics is elementary throughout and rests on information theory, elementary decision theory and a modest amount of measure theory. Where a quantity would ordinarily be given a numerical value it is given a symbol, because any number supplied would be invented, and because a symbol has the merit of making no claim about the world that the surrounding argument cannot support. Assumptions are numbered and, where they are adopted without defense, are labeled as such. The proofs are by pencil. No neuron was consulted.

A word on scope is necessary. Neural interfaces that restore communication and control to people with paralysis or loss of speech are among the more admirable accomplishments of contemporary engineering, and nothing below bears on them except to explain why they work as well as they do: for those users the conventional channel has been lost, so that the cost of the conventional alternative is very large and the inequality that governs adoption is satisfied with room to spare. The target of the book is a thesis about interfaces in general and about universal adoption in particular, not any instrument, laboratory result or clinical program.

Appendix C states this again in plainer terms.

Contents

I The Metaphysics of the Missing Button **1**

1 The Elimination of the Intermediary **2**
1.1 A Short History of Becoming Unnecessary 2
1.2 Definitions 2
1.3 The Thesis and the Hypothesis 3
1.4 What Is Not Claimed 4
1.5 Plan of the Work 4

2 Anatomy of a Button **6**
2.1 The Button as an Achievement of Irrelevance 6
2.2 Quantization at the Source 6
2.3 The Keyboard as Deliberate Impoverishment 7
2.4 The Cost of Saying It Twice 7

3 Bandwidth Is Not Throughput **9**
3.1 Two Quantities with One Name 9
3.2 What Information Is Required 10
3.3 Which Bound Is Active 10

4 The Conservation of Interface Work **12**
4.1 A Cost Functional 12
4.2 Assumptions Adopted Without Defense 12
4.3 Where the Work Goes 13

5 The Relocation Principle **15**
5.1 Where the Keyboard Went 15
5.2 The End-to-End Argument 15
5.3 The Organizing Theorem 16

II The Mathematics of Reading What Nobody Has Said **17**

6 The Discretization Paradox **18**
6.1 The Map from Tissue to Symbol 18

6.2 The Boundary Is Not Optional 18
6.3 Debouncing, Hysteresis and the First Neural Interface Decoder 19
6.4 Zero Error Is Not Available 20

7 Semantic Indeterminacy **21**
7.1 No Intrinsic Command Meaning 21
7.2 The Quine Property 21
7.3 Co-Adaptation 22

8 Thinking Without Doing **23**
8.1 Four Predicates 23
8.2 Authorization Is Not Identifiable from Insufficient Observations 24
8.3 Authorization as a Property of a Protocol 24
8.4 Even a Perfect Decoder Does Not Authorize 25

9 The Regress of the Confirmation Command **27**
9.1 Asking Again 27
9.2 Perfectly Correlated Errors 27
9.3 A Floor Under Latent States 28

10 Drift and the Maintenance Burden **29**
10.1 The Stale Decoder 29
10.2 An Irreducible Maintenance Rate 29

11 Echo, Undo and the Ledger of Commands **31**
11.1 The Other Half of the Loop 31
11.2 Undo Requires a Record 31

III The Industrial Reinvention of the Human Hand **33**

12 Standards, Lifecycles and Orphaned Users **34**
12.1 The Interchangeable Peripheral 34
12.2 The Orphan Interval 34

13 The Right Not to Execute **36**
13.1 Friction as a Constitutive Feature 36
13.2 Neural Data and the Boundary of the Mind 37

14 The Economics of Universal Adoption **38**
14.1 The Adoption Inequality 38
14.2 The Disappearance of the Interface as a Product Category 39

15 Exit, Dependence and the Displacement of Governance **40**
15.1 The Interface as a Small Institution 40
15.2 The Thesis in Its Own Terms 40
15.3 Displacement of Dependence 41
15.4 Conflicting Claims 42
15.5 Formal Exit and Effective Exit 42
15.6 The Historical Irony 43
15.7 The Sovereignty Displacement Thesis 43
15.8 The Interface Once More 44

16 A Modest Proposal for the Imagined Button **45**
16.1 The Device 45
16.2 Dominance of the Button 45
16.3 A Note on Imagery 46

17 What Has Not Been Shown **47**

A Standard Results Used **49**

B Notation and Glossary **50**

C A Statement Concerning Assistive Neural Interfaces **51**

Part I

The Metaphysics of the Missing Button

Chapter 1

The Elimination of the Intermediary

1.1 A Short History of Becoming Unnecessary

The history of human-computer interaction can be told as a long effort to make machines accommodate human beings. The punched card gave way to the teletype, the teletype to the screen and keyboard, the keyboard to the pointing device, and the pointing device to the touch surface, and at each stage the machine was made to tolerate a little more of the ordinary variability of human gesture [3, 29]. After roughly seventy years of this labor, a remarkable proposal has emerged: that the next stage consists in engineering the human being to accommodate the machine, by routing the machine's attention past the hands, the voice and the eyes and into the tissue that was formerly their employer. The proposal is generally presented as the culmination of the history just described, as though the progressive removal of intermediaries had a natural terminus in the removal of the last one.

The proposal deserves a serious answer, and the answer is not that it is impossible. Devices that record from cortex and translate the recordings into cursor movement, text and synthesized speech have been demonstrated repeatedly in people who have lost other means of communication [19, 5, 39, 26, 40, 25]. The developers of commercial implants have described ambitions of very large channel counts [27] and have reported, for cursor control in clinical participants, information transfer rates that they describe as comparable to those of able-bodied mouse use [28]. The question taken up here is not whether these devices work but what is being claimed when it is said that they dispense with the interface, and whether that claim follows from the fact that they work.

1.2 Definitions

The vocabulary is kept deliberately small. A *command alphabet* is a finite set $\mathcal{C}$ of commands containing a distinguished element $\varnothing$, the null command, which stands for taking no action. The user's *intended command* is a random variable Θ taking values in $\mathcal{C}$. An *observation* is a random variable Y that depends on the user's activity: it may be the contact state of a switch, or it may be a recording $N \in \mathbb{R}^n$ of neural activity. An *executed command* $\widehat{\Theta}$ is the

command that the machine in fact carries out. An *interface* is any mechanism that produces an observation from the user's activity and an executed command from the observation, possibly with the aid of a bounded history; the map from observation to executed command is its *decoder*. For a button the decoder is a threshold on a contact state. For a neural interface it is a learned function of a recording. The interface makes an *error* when $\widehat{\Theta} \neq \Theta$, and its error probability is $P_e = \Pr(\widehat{\Theta} \neq \Theta)$.

Definition 1.1 (The Four Functions)**.** The *function of distinction* partitions the space of observations into classes that the machine treats as equivalent, so that continuous or variable activity is reduced to a finite set of cases. The *function of interpretation* assigns a command to each class. The *function of authorization* separates commands the user has committed to executing from commands the user has merely considered, prepared or rehearsed. The *function of correction* detects mismatches between the intended and the executed command and repairs them.

1.3 The Thesis and the Hypothesis

The target of the book can now be stated. The *Interface Elimination Thesis* is the claim that the four functions of Definition 1.1, as performed by keyboards, buttons, pointing devices and the like, are dispensable overhead, and that a system coupling neural activity directly to the machine performs the same tasks with strictly less total work because it does not need to perform them. The *Main Hypothesis* is that this thesis is logically untenable: that any system converting human activity into executable machine commands must perform all four functions, that eliminating a physical intermediary does not eliminate any of them but only relocates the place where they are performed, and that the thesis therefore rests on a confusion between a mechanism and the function it implements.

The first clause, the functional component of the Main Hypothesis, is very nearly a tautology, which is the reason it can be asserted with confidence. The second clause, the comparative component, is the reason the hypothesis as a whole cannot be proved. The distance between a function that must be performed somewhere and a function whose performance costs more than it did before is the distance the following chapters repeatedly approach and do not cross. The tautological clause is, however, worth stating as a theorem, because it is the organizing result of the book and because its proof is short enough to be checked in the margin.

Theorem 1.2 (Interface Ineliminability)**.** *Let X be a set of observations and $\mathcal{C}$ a set of commands, and let a system select commands by a deterministic function $f : X \to \mathcal{C}$. If the system can issue at least two distinct commands, then there exist $x_1, x_2 \in X$ with $f(x_1) \neq f(x_2)$. Consequently the system implements a nontrivial distinction over its observations.*

Proof. Suppose the system can issue distinct commands $c_1 \neq c_2$. Then c_1 and c_2 both lie in the image of f, so there are observations x_1, x_2 with $f(x_1) = c_1$ and $f(x_2) = c_2$, and plainly $f(x_1) \neq f(x_2)$. Define $x \sim_f y$ if and only if $f(x) = f(y)$. The equivalence classes partition

X into sets of observations that produce the same command, and because the image of f contains at least two commands the partition contains at least two nonempty classes. Any implementation of f must therefore distinguish at least two classes of observation. □

The theorem does not assert that the distinction is expensive, that it must be performed by a separate module, or that it must be performed by a physical object. It asserts only that the distinction exists wherever two commands can be issued. A keyboard performs it with switches. A neural interface performs it with a decoder. The first achievement of the interface-free computer is accordingly the construction of another interface, and the theorem would be of no interest to anyone if the Interface Elimination Thesis had not been stated in a form that appears to deny it.

1.4 What Is Not Claimed

Several claims are not made, and the list is long enough to be worth stating in full. It is not claimed that neural decoding is impossible, or that it is inferior for every user. It is not claimed that continuous control is inferior to discrete control: pointing devices, steering wheels and musical instruments are continuous, and a prosthetic limb driven by neural signals may be best controlled in the same way. It is not claimed that the cost of a neural interface exceeds that of a conventional interface for any particular person. It is not claimed that physical intermediaries have no costs of their own, since keyboards are inaccessible to many people and tiresome to all. It is not claimed that anything in the book applies to the clinical programs that restore function to people whose conventional channels are lost. The claim, to the extent that one is made, is confined to the logical structure of the argument that goes from the removal of a physical intermediary to the removal of its function.

1.5 Plan of the Work

Part I, which this chapter opens, examines the metaphysics of the missing button: what an interface does that is not merely transmission, why recording bandwidth differs from command throughput, and why the work of an interface is relocated rather than removed. Part II develops the mathematics of reading what nobody has said, through the problems of discretization, semantic indeterminacy, authorization, confirmation, drift and error correction. Part III turns to the industrial and institutional consequences of replacing standardized, inspectable and repairable interfaces with individually calibrated biological dependencies, extends the argument from interfaces to the governance claims of the sovereign-individual tradition, and ends with a proposal for a device that permits its user to control a computer by imagining that a button is pressed.

Standing of the Main Hypothesis. Chapter 1 has stated the Main Hypothesis and proved one theorem, Theorem 1.2, which is true of every system that issues more than one command and therefore supports the Main Hypothesis only in the sense in which it supports

everything. Nothing in the chapter shows that the Interface Elimination Thesis is untenable, and the chapter shows no cost increase of any kind. The hypothesis stands exactly where it stood before the chapter began.

Chapter 2

Anatomy of a Button

2.1 The Button as an Achievement of Irrelevance

A button is an extraordinarily sophisticated logical invention precisely because it eliminates almost everything about the person who presses it. The pressing hand may be tired or fresh, the person may be calm or agitated, the metabolic state of the body, the oscillations of cortex and the autobiographical associations of the moment are all without effect on the machine, which requires only the single fact that a contact has closed. The button is a mechanism for discarding information that the machine does not need, and the discarding is its function rather than a defect of its construction. A design that preserved more of the user would transmit more and would be correspondingly harder to interpret, since the machine would be obliged to decide which of the preserved features were relevant.

This observation has a respectable ancestry in the theory of regulation. Ashby's law of requisite variety holds that a regulator can absorb variety in a disturbance only by possessing variety of its own [1]. An interface that offers the machine a very large variety of signals does not thereby offer a very large variety of commands, since the machine must still map each signal to one of a small number of actions, and the surplus variety constitutes disturbance that the decoder must absorb. The button avoids the difficulty by never admitting the surplus in the first place.

2.2 Quantization at the Source

Definition 2.1 (Source Quantizer)**.** An interface is a *source quantizer* if the observation Y takes values in a finite alphabet in one-to-one correspondence with $\mathcal{C}$, so that the decoder is the identity on that alphabet.

A keyboard is a source quantizer, and so is any panel of labeled switches. In such an interface the user performs the discretization: the choice among alternatives is made by the user's motor act, and the machine reads the result. The consequence for error analysis is immediate. If $Y = \widehat{\Theta}$ deterministically, all error is attributable to the user's act, whether a slip, a mispress or a change of mind, and such error is visible to the user at once, through

key travel, sound and echo on the screen. In the idealized case the decoder contributes no error of its own and requires no calibration, properties that no learned decoder can match and that are obtained by the simple expedient of moving the quantizer from the machine into the finger. Actual keyboards involve scanning, debouncing, firmware, key mapping and the possibility of hardware failure, but these are small, inspectable and fixed, and they are not the learned and drifting decoders of the neural case.

Lemma 2.2. *For a source quantizer in the idealized case, the mutual information between intended and executed command satisfies* $\mathsf{I}(\Theta;\widehat{\Theta}) = \mathsf{H}(\Theta) - \mathsf{H}(\Theta \mid \widehat{\Theta})$*, and the residual conditional entropy* $\mathsf{H}(\Theta \mid \widehat{\Theta})$ *is generated entirely by the user's act and not by the decoder.*

Proof. The first equality is the definition of mutual information. Because $\widehat{\Theta}$ is a deterministic function of Y, and Y is a deterministic function of the user's act, any uncertainty about Θ remaining after $\widehat{\Theta}$ is known is uncertainty about the relation between intention and act. □

2.3 The Keyboard as Deliberate Impoverishment

The keyboard implements a finite symbolic vocabulary, an event-generation mechanism, a repeatable spatial organization, a learnable motor convention and a transparent process of error correction, and it does so with a physical object that costs little and can be replaced in a minute. The machine does not need to know the neural history of the letter A. It needs to know that an A-event has been registered in a particular context. The usefulness of the device arises from its impoverishment of information, and a history of the device that treats the impoverishment as a limitation to be overcome has misunderstood what the device is for.

Simon's account of artifacts as the interface between an inner environment and an outer one is relevant here [36]. On his account, an artifact works because the inner environment, the substance and organization of the artifact itself, is adapted to the outer environment in which it operates, and the interface between the two is the place where the adaptation is made. A button is an interface in this exact sense: its inner environment is a mechanism of contact and its outer environment is a human hand. To remove it is not to remove the boundary between the two environments but to relocate the boundary to a place where it cannot be seen.

2.4 The Cost of Saying It Twice

The human operator is a channel with its own limitations. Hick and Hyman showed that choice reaction time grows with the logarithm of the number of equally likely alternatives, and Fitts showed that movement time grows with the logarithm of the ratio of distance to target width [18, 17, 24]. These laws describe a channel whose capacity is bounded both by the rate at which decisions can be made and by the precision with which movements can be executed. The two bounds are different, and a device that removes the second does nothing to the first. The question of which bound is active in a given task is empirical, and Chapter 3 returns to it.

Standing of the Main Hypothesis. Chapter 2 has shown that a button discards information on purpose and that a keyboard moves quantization to the user. It has not shown that a neural interface must be worse at anything, and it has been explicit that the choice of alphabet, the cost of the physical object and the pace of the human operator all vary from task to task. The Main Hypothesis is no better supported than it was, and is perhaps worse supported, since the chapter has praised a device without comparing it to anything.

Chapter 3

Bandwidth Is Not Throughput

3.1 Two Quantities with One Name

A neural recording system is characterized by the number of channels it samples, the rate at which it samples them and the resolution of each sample, and the product of these gives a raw data rate that can be made impressive by increasing any factor. A command interface is characterized by the amount of information per unit time that it conveys about what the user intends to do. The two quantities are routinely discussed under the common name of bandwidth, with the effect that growth in the first is mistaken for growth in the second. The following results separate them.

Theorem 3.1 (Post-Processing Cannot Create Missing Information)**.** *Let Θ be an intended command, N a neural observation and $\widehat{\Theta}$ the output of a decoder that receives only N, so that $\Theta \to N \to \widehat{\Theta}$ is a Markov chain. Then*

$$\mathsf{I}(\Theta; \widehat{\Theta}) \leq \mathsf{I}(\Theta; N) \leq \mathsf{H}(\Theta).$$

Proof. By the chain rule, $\mathsf{I}(\Theta; N, \widehat{\Theta}) = \mathsf{I}(\Theta; N) + \mathsf{I}(\Theta; \widehat{\Theta} \mid N)$, and the Markov property gives $\mathsf{I}(\Theta; \widehat{\Theta} \mid N) = 0$, so $\mathsf{I}(\Theta; N, \widehat{\Theta}) = \mathsf{I}(\Theta; N)$. Applying the chain rule in the other order, $\mathsf{I}(\Theta; N, \widehat{\Theta}) = \mathsf{I}(\Theta; \widehat{\Theta}) + \mathsf{I}(\Theta; N \mid \widehat{\Theta})$, and since conditional mutual information is nonnegative, $\mathsf{I}(\Theta; \widehat{\Theta}) \leq \mathsf{I}(\Theta; N)$. For discrete Θ, $\mathsf{I}(\Theta; N) = \mathsf{H}(\Theta) - \mathsf{H}(\Theta \mid N) \leq \mathsf{H}(\Theta)$. □

The consequence is that a decoder cannot recover information about an intended command that is absent from its observations, and that no increase in the computational complexity of the decoder, including the use of a larger network, overturns the inequality. The qualification is that a decoder may receive additional inputs in the form of context, feedback or interaction history, in which case those inputs belong in the observation variable and the theorem is applied to the enlarged observation. The second inequality has a more humbling reading. The ceiling on command information is the entropy of the user's own command process, which is a property of the person and the task and not of the hardware. A million electrodes cannot cause a user to decide more things than the user decides.

Theorem 3.2 (Raw Signal Entropy Does Not Imply Command Information)**.** *For every $M > 0$ there exist a neural observation variable N and an intended command variable Θ such that $\mathsf{H}(N) > M$ and $\mathsf{I}(\Theta; N) = 0$.*

Proof. Let Θ be a fair binary command variable and let N be uniformly distributed on a set of 2^k possible observations, independently of Θ. Then $\mathsf{H}(N) = k$ bits, while independence gives $\mathsf{I}(\Theta; N) = 0$. Choosing any integer $k > M$ gives the result. □

A system can therefore record arbitrarily large amounts of data while obtaining no information about the command of interest. Conversely, a conventional binary button conveys one reliable bit per use without measuring millions of biological variables. Theorem 3.2 does not show that buttons have greater throughput than implants, only that measurement bandwidth alone is not evidence of command throughput.

3.2 What Information Is Required

A more operational statement is available through Fano's inequality [16, 6]. Let $m = |\mathcal{C}|$, let Θ be uniform on $\mathcal{C}$, and let the executed command have error probability P_e. Then

$$\mathsf{H}(\Theta \mid \widehat{\Theta}) \leq h_2(P_e) + P_e \log_2(m-1),$$

where h_2 is the binary entropy function, and therefore

$$\mathsf{I}(\Theta; \widehat{\Theta}) \geq \log_2 m - h_2(P_e) - P_e \log_2(m-1).$$

This gives the amount of command-relevant information that any interface, of any construction, must deliver in order to achieve a specified error rate, and it combines with Theorem 3.1 to show that the information must be present in the observation before it can be present in the command. The engineering question is therefore not how much neural data an implant records but how much information the recording supplies about the particular distinctions the user needs to communicate.

3.3 Which Bound Is Active

If the rate of decision-making is the binding constraint for a task, then replacing the motor act with a neural readout leaves the ceiling of Theorem 3.1 in place, and the interface may match but not exceed the conventional one. If motor execution is the binding constraint, a neural interface can in principle raise throughput. Reports from clinical programs describe cursor-control rates of the order of ten bits per second, which the developers compare to able-bodied mouse use [28]. Taken at face value these reports establish that the implant has reached parity with a device that costs a few currency units, and they say nothing about which constraint is binding. Parity with the incumbent is a considerable achievement for a person who cannot use the incumbent and a modest one for a person who can, and the difference is the subject of Chapter 14.

Standing of the Main Hypothesis. Chapter 3 has proved that a decoder cannot create information, that recording volume and command information are distinct, and that the ceiling on command throughput is set by the person. None of these results favors a conventional interface over a neural one for any task, and the chapter has explicitly declined to say which of the two bounds is active. The Main Hypothesis remains unsupported, and the chapter's conclusion that the question is empirical is a conclusion that the book will reach again.

Chapter 4

The Conservation of Interface Work

4.1 A Cost Functional

Let d denote an interface design, u a user and τ a task. The total cost of interaction is written

$$J(d, u, \tau) = \alpha E + \beta L + \gamma U + \delta K + \epsilon R,$$

where E is command error, L is latency, U is user effort, K is the burden of calibration and maintenance, R is the risk to safety and reliability, and $\alpha, \ldots, \epsilon$ are nonnegative weights. The weights are not specified, because any values supplied would be arbitrary and the argument does not need them. A technology is not superior merely because it reduces the number of visible physical intermediaries. It is superior if it improves the relevant components of J for the user and the task, and the number of visible intermediaries does not appear in J.

The work of the interface is divided among the parties who perform it. Write

$$W_{\text{interface}} = W_{\text{user}} + W_{\text{decoder}} + W_{\text{calibration}} + W_{\text{verification}},$$

for the work done respectively by the user in producing the signal, by the decoder in interpreting it, by whoever maintains the correspondence between the two, and by whoever checks that the executed command matches the intended one. This is an accounting identity under an appropriately specified cost model and not a physical conservation law. The accounting nevertheless has a use, which is to make visible the places where work can hide. A button places nearly all of the work in the user's finger and nearly none in the decoder. A neural interface may place less work in the user and more in the other three terms, and a statement that the total has fallen requires an argument about the sum and not an observation about a single term.

4.2 Assumptions Adopted Without Defense

Assumption 4.1 (Additivity)**.** The total work of an interface for a given task is the sum of the four terms above, each nonnegative.

Assumption 4.2 (Information Floor)**.** To achieve error probability P_e on alphabet $\mathcal{C}$ the interface must supply at least $I_{\min} = \mathsf{H}(\Theta) - 1 - P_e \log_2 |\mathcal{C}|$ bits of information about Θ per command. This follows from Fano's inequality in the weakened form obtained from Chapter 3 by using $h_2 \leq 1$ and $\log_2(m-1) \leq \log_2 m$.

Assumption 4.3 (Pricing of Bits)**.** Let b denote the number of bits per command that the user deliberately encodes through an act, and b' the number that the decoder extracts from aspects of the signal that the user did not deliberately encode. The user's work is a nondecreasing function $c_u(b)$ and the decoder, calibration and verification work together a nondecreasing function $c_d(b')$, and the requirement is $b + b' \geq I_{\min}$. The two contributions are taken to be nonoverlapping, so that the information delivered about Θ is their sum. This is an idealization, since information from two sources may be redundant, and the constraint in general involves the joint mutual information of the two sources with Θ.

Proposition 4.4 (Non-Entailment)**.** *Under Assumptions 4.1 to 4.3, the removal of a physical interface neither entails nor excludes a reduction of total interface work. Specifically, there exist cost functions c_u, c_d for which the neural design has strictly lower minimal work than the button, and cost functions for which it has strictly higher minimal work.*

Proof. For each design the minimal total work is $W^* = \min_{b+b' \geq I_{\min}} [c_u(b) + c_d(b')]$. Take the button to have $c_u^B(b) = \kappa_B b$ for $b \leq \mathsf{H}(\Theta)$ and no available extraction, so that $W_B^* = \kappa_B I_{\min}$. Take the neural design to have user cost $c_u^N(b) = \kappa_N b$ and decoder cost $c_d^N(b') = \lambda b'$. If $\kappa_N < \kappa_B$ and $\lambda < \kappa_B$ then $W_N^* \leq \kappa_N I_{\min} < W_B^*$. If $\kappa_N > \kappa_B$ and $\lambda > \kappa_B$ then every feasible allocation costs more than $\kappa_B I_{\min}$, so $W_N^* > W_B^*$. Both configurations satisfy the assumptions. □

The proposition is the honest content of what the book can prove about work. It states that the Interface Elimination Thesis is a contingent claim about relative prices and not a consequence of removing the intermediary. It also states, by symmetry, that the opposite claim is contingent, and that the Main Hypothesis, if read as saying that elimination necessarily raises total work, is not supported by the proposition. The distance between *not entailed* and *untenable* is the distance across which the book has agreed to look and not to walk.

4.3 Where the Work Goes

One observation deserves emphasis because it is easy to overlook. Neural motor interfaces in current use do not decode intention in the abstract. They decode the neural correlates of attempted or imagined movement, so that the user remains engaged in a deliberate encoding act and has exchanged a muscular act for an attempted one. In the terms of Assumption 4.3 the user still contributes $b > 0$ bits by deliberate act, and the interface has changed the medium of the act and not its existence. The keyboard disappears from the desk and reappears as a statistical classifier requiring calibration, and the finger disappears from the keyboard and reappears as an attempted finger.

Standing of the Main Hypothesis. Chapter 4 has proved Proposition 4.4, which says that the reduction of total work by neural coupling is neither implied nor ruled out by logic. This is a result against the Interface Elimination Thesis only in the sense that it removes a supposed entailment, and it is equally a result against the strong reading of the Main Hypothesis. The hypothesis has therefore been weakened by the chapter intended to support it, and the author considers this an instructive outcome.

Chapter 5

The Relocation Principle

5.1 Where the Keyboard Went

The four functions of Definition 1.1 can be traced through the transition from a conventional to a neural interface. The tracing is displayed in Table 5.1, in which the right-hand column is, in each case, a relocation and not an abolition.

Function	Conventional locus	Locus after elimination
Distinction	Contact closure at the switch	Threshold on a classifier output
Interpretation	Key-to-symbol map in firmware	Learned decoder parameters
Authorization	Deliberate press, with travel as friction	Confirmation event, or none
Correction	Backspace and undo	Backspace and undo, unchanged
Feedback	Key travel, sound, screen echo	Screen echo, unchanged

Table 5.1: The relocation of interface functions.

The last two rows are the instructive ones. Correction and feedback are not relocated at all, because they were never performed by the physical intermediary: they are performed by the display and the editing software, which the neural interface retains in full. The elimination thus removes the input half of a loop of which the output half is untouched, a point that Chapter 11 develops.

5.2 The End-to-End Argument

The principle that governs where such functions should live was articulated for network design by Saltzer, Reed and Clark, who argued that a function implemented at a low level of a system is often redundant or of little value compared with the cost of providing it, because the application at the endpoints must in any case verify the outcome [32]. Lampson's hints for

system design make a related point about the value of keeping interfaces simple and separable [22]. Applied to the present case, the argument has an ironic form: even a perfect decoder would not remove the need for end-to-end verification that the executed command is the intended one, since the only party who knows what was intended is the user, and the only evidence of what was executed is the display. The end-to-end check cannot be eliminated by improving the lower level, and it is the check, and not the lower level, that does the work on which reliability finally depends.

5.3 The Organizing Theorem

The preceding results can be gathered into a single statement. Its proof depends on Theorem 1.2 and on results proved in Part II, to which the reader is referred forward, and it is stated here because it is the form in which the Main Hypothesis is closest to being proved, and is not.

Theorem 5.1 (Unavoidable Interface Partition)**.** *Consider a system that receives observations and executes commands from a finite set $\mathcal{C}$ that includes the null command $\varnothing$. Suppose the system must distinguish multiple commands, must refrain from execution in some circumstances, and must satisfy a specified error criterion. Then any successful implementation provides, explicitly or implicitly, mechanisms that perform command discrimination, execution selection and error control.*

Proof. Because multiple commands must be distinguished, Theorem 1.2 requires a nontrivial partition, or a probabilistic discrimination, of the observations. Because $\varnothing$ is a possible output and execution is not always permitted, the system must implement a rule that separates execution from non-execution. Satisfying the error criterion requires the decision rule to meet the specified bound under the relevant observation distributions, and where observations are ambiguous this may require better measurements, context, feedback, abstention or other error-control mechanisms. □

Notice what the theorem does not establish. It does not prove that a neural interface must contain a separate module for each function, that it must use a conventional confirmation device, or that its implementation must be more complicated than an ordinary interface. It establishes that the underlying logical obligations persist whatever the physical source of the observations.

Standing of the Main Hypothesis. Chapter 5 has displayed the relocation of the functions and proved Theorem 5.1, which establishes that obligations persist and says nothing regarding their cost. The functional component of the Main Hypothesis is thereby established, while the comparative component, which requires that the persistence be a problem, has not been advanced. It is recorded that the chapter's principal result is a restatement of Theorem 1.2 with additional words.

Part II

The Mathematics of Reading What Nobody Has Said

Chapter 6

The Discretization Paradox

6.1 The Map from Tissue to Symbol

Human cognition and neural activity are not streams of ready-made digital commands. A person can entertain competing intentions, partially prepare actions, imagine alternatives, inhibit impulses and revise decisions, and the corresponding neural activity is high-dimensional, noisy and continuously varying. The machine, however, eventually requires a sufficiently definite command. Let the neural measurement be $N(t) \in \mathbb{R}^n$ and let the command alphabet be the finite set $\mathcal{C} = \{c_1, \ldots, c_m\}$. The interface must then implement a map

$$f : \mathbb{R}^n \longrightarrow \mathcal{C},$$

and this is precisely an act of discretization. The neural interface has not abolished the distinction between analog biological activity and digital computation. It has inherited responsibility for establishing that distinction, and because the relevant signals are noisy, context-sensitive and subject to change over time, the map requires estimation, thresholding, filtering and temporal integration. The supposed disappearance of the interface becomes an elaborate reconstruction of its functions.

6.2 The Boundary Is Not Optional

Every discretization has a boundary, and the boundary is where the trouble is. This can be made precise. Say that a point x has *margin* at most δ for f if some point within distance δ of x is mapped to a different command.

Theorem 6.1 (Unavoidable Margin)**.** *Let N have a density that is strictly positive on $\mathbb{R}^n$, and let $f : \mathbb{R}^n \to \mathcal{C}$ be measurable and take at least two values on sets of positive Lebesgue measure. Then for every $\delta > 0$ the set B_δ of points with margin at most δ has positive measure, and therefore* $\Pr(N \in B_\delta) > 0$.

Proof. Suppose B_δ has measure zero, and call a point *good* if it lies outside B_δ. A good point x has the property that f is constant on the ball $B(x, \delta)$. The good points have full measure

and are therefore dense. Take two good points x, y and a path between them; choose an integer k with $|y-x|/k < \delta/2$ and let $t_j = x + (j/k)(y-x)$ be equally spaced points on the straight segment between them. By density, choose good points x_j with $|x_j - t_j| < \delta/4$ for $0 < j < k$, and put $x_0 = x$, $x_k = y$. Then $|x_{j+1} - x_j| < \delta/2 + 2(\delta/4) = \delta$ for every j, so no compactness argument is needed. Consecutive good points in the chain lie in each other's balls, so f takes the same value at each, and hence $f(x) = f(y)$. Thus f is constant on the good points, hence almost everywhere, contradicting the assumption that f takes two values on sets of positive measure. So B_δ has positive measure, and the strictly positive density gives $\Pr(N \in B_\delta) > 0$. The step from positive Lebesgue measure to positive probability uses the positivity of the density and would fail without it. □

Theorem 6.1 says that no decoder with more than one effective output can avoid assigning positive probability to observations that a small perturbation would send to a different command. It applies equally to a mechanical switch, which is the reason that switches have a travel, a detent and a defined actuation force: the boundary is engineered to be crossed deliberately and to be crossed only by deliberate acts. In a neural decoder the boundary is wherever the learned parameters put it, and the perturbations that cross it include every source of variation in the signal.

6.3 Debouncing, Hysteresis and the First Neural Interface Decoder

Electrical engineers solved the boundary problem for switches decades ago, by debouncing contacts and by adding hysteresis to threshold circuits. A comparator with a single threshold θ will chatter when its input hovers near θ, flipping rapidly between outputs. The standard remedy is a pair of thresholds $\theta_{\mathrm{lo}} < \theta_{\mathrm{hi}}$, the output switching high only when the input exceeds θ_{hi} and low only when the input falls below θ_{lo}.

Proposition 6.2 (Hysteresis Suppresses Chatter)**.** *Let the underlying signal be a constant s and let observations be $s + \varepsilon_t$ with $|\varepsilon_t| \le a$ for all t. If $a < (\theta_{\mathrm{hi}} - \theta_{\mathrm{lo}})/2$ and $s \in [\theta_{\mathrm{lo}} + a, \theta_{\mathrm{hi}} - a]$, then the output never changes once an initial state is established.*

Proof. If the output is high it can fall only when an observation drops below θ_{lo}, but every observation is at least $s - a \ge \theta_{\mathrm{lo}}$. If the output is low it can rise only when an observation exceeds θ_{hi}, but every observation is at most $s + a \le \theta_{\mathrm{hi}}$. The interval is nonempty exactly when $a < (\theta_{\mathrm{hi}} - \theta_{\mathrm{lo}})/2$. □

The circuit with these properties is the Schmitt trigger, which Otto Schmitt devised in the 1930s while studying the propagation of impulses in squid nerve [33]. The neural interface decoder, in other words, was invented first as a model of the neuron and then as a component for tidying the output of a noisy switch, and a modern implant that employs dwell times and hysteresis to prevent unintended selections is rediscovering a component whose conceptual ancestry runs through the very tissue it now reads. The observation has no force as an argument, but it has a certain symmetry.

6.4 Zero Error Is Not Available

The boundary theorem has a classification-theoretic counterpart which shows that the difficulty is not merely one of engineering effort.

Theorem 6.3 (Overlapping Observations Preclude Perfect Classification)**.** *Let two commands c_0, c_1 have strictly positive prior probabilities and conditional observation distributions P_0, P_1 that are not mutually singular. Then no classifier using only the observation can achieve zero classification error.*

Proof. Suppose a classifier achieves zero error, and let A be the set of observations on which it outputs c_1. Zero error requires $P_0(A) = 0$ and $P_1(A) = 1$, that is, $P_0(A^c) = 1$ and $P_1(A^c) = 0$. Then A and A^c carry the full mass of P_1 and P_0 respectively and are disjoint, so P_0 and P_1 are mutually singular, contrary to hypothesis. □

The limitation applies equally to noisy conventional interfaces, and the book does not claim otherwise. The distinctive question is whether the neural interface offers a better or worse error distribution for a task, and that question cannot be answered by theorem.

Standing of the Main Hypothesis. Chapter 6 has proved that every discretization has a boundary, that boundaries invite chatter and that perfect classification is unavailable when observation distributions overlap. These results apply to all interfaces, including buttons, and the chapter has said so. The Main Hypothesis, which concerns a comparative claim, is untouched by results that hold on both sides of the comparison.

Chapter 7

Semantic Indeterminacy

7.1 No Intrinsic Command Meaning

A neural activation pattern has no intrinsic meaning as a command. It is not naturally equivalent to the instruction to save a file. Its meaning is established through a trained relationship among the user, the decoder, the task and the environment, so that neural decoding is not merely signal detection but the maintenance of an interpretive system whose operational semantics are acquired. A useful formalization introduces a contextual decoder

$$f_\theta(N, H, X) \longrightarrow \mathcal{C},$$

where H is interaction history, X is task context and θ denotes learned parameters. Identical or similar neural observations may then correspond to different appropriate actions under different contexts. The problem is not peculiar to neural interfaces, since ordinary language and conventional interfaces also depend on context, but conventional interfaces tend to provide explicit symbols and observable state transitions that reduce the ambiguity, and a neural interface must reconstruct them.

7.2 The Quine Property

The philosophical literature has a name for the situation of an interpreter confronted with behavior and no dictionary. Quine's thesis of the indeterminacy of translation holds that the totality of speech dispositions leaves open incompatible systems of analytical hypotheses [31], and Dretske's account of information distinguishes the signals a channel carries from the contents that may be assigned to them [7]. A decoder trained on finitely many observations is in Quine's position with respect to the brain, and the following proposition records the position formally.

Proposition 7.1 (Off-Support Underdetermination)**.** *Let f be any decoder from a space of observations to $\mathcal{C}$, trained on a finite set S of distinct observations with assigned commands. For any finite set S^* of observations disjoint from S and any assignment of commands to S^*,*

there is a decoder f' that agrees with f on S and with the assignment on S^*.

Proof. Define $f'(x) = f(x)$ for $x \notin S^*$ and let f' take the prescribed values on S^*. Because S and S^* are disjoint, f' agrees with f on S. □

The proposition is true only when decoders are drawn from the class of all functions, and the class actually used in practice is restricted by smoothness, architecture and regularization, a fact the proposition does not conceal. The restriction is an inductive bias, and the inductive bias is where the semantics of the interface now reside. The semantics of a button reside in a label and a convention that any person can read. The semantics of a decoder reside in parameters that no one can read directly and in a prior that is chosen by whoever trained it.

7.3 Co-Adaptation

The difficulty is aggravated in closed-loop use, where the user adapts to the decoder while the decoder is adapted to the user. Each party is a nonstationary environment for the other, and a fixed point of the joint process is not guaranteed to exist or to be unique. Systems have been built that exploit the adaptation deliberately, but the user's task then includes learning what the decoder presently means by the user's own neural activity. A conventional interface asks the user to learn a keyboard layout once. A co-adapting interface asks the user to remain in conversation with a model of the user, and to keep the conversation going indefinitely.

Standing of the Main Hypothesis. Chapter 7 has shown that the semantics of a neural command are conventions maintained by a trained decoder and are underdetermined by finite data in the absence of an inductive bias. The chapter has not shown that the conventions are less stable, less learnable or less useful than those of a keyboard. The Main Hypothesis, which requires a comparison, has been supported by none.

Chapter 8

Thinking Without Doing

8.1 Four Predicates

Ordinary interfaces permit a separation between mental events and publicly executable commands. A person can think about deleting a file without deleting it, imagine sending an insulting message without sending it, rehearse a musical phrase without performing it, or consider turning left and turn right. The separation is not an accidental inconvenience; it is part of the architecture of deliberate agency. A neural interface must therefore distinguish contemplating an action, desiring it, preparing it and authorizing it, and the four are not logically equivalent. Let $T(a)$ mean that the subject thinks about action a, $D(a)$ that the subject desires it, $P(a)$ that the subject prepares it, and $Z(a)$ that the subject authorizes it.

Proposition 8.1 (Non-Implication)**.** *The implications $T(a) \Rightarrow D(a)$, $D(a) \Rightarrow P(a)$ and $P(a) \Rightarrow Z(a)$ are not valid.*

Proof. Consider four worlds with valuations (T, D, P, Z) equal to $(1,0,0,0)$, $(1,1,0,0)$, $(1,1,1,0)$ and $(1,1,1,1)$. The first falsifies $T \Rightarrow D$, the second falsifies $D \Rightarrow P$, and the third falsifies $P \Rightarrow Z$. □

A reliable system therefore needs some means of determining authorization, and the empirical literature on the neural precursors of action gives little reason to think that authorization can be read from preparation. Libet and colleagues reported that a slow negative potential preceding self-initiated movement begins before the time at which subjects report awareness of the intention to move [23], and subsequent work found that patterns of activity could predict simple free choices several seconds in advance [37]. These findings are contested, and Schurger and colleagues have argued that the readiness potential may reflect the accumulation of spontaneous fluctuations to a threshold and not a decision that has already been made [34]. On either reading the lesson for interface design is the same: neural precursors of action exist, they precede commitment, and they are not commitment.

8.2 Authorization Is Not Identifiable from Insufficient Observations

The strongest formal statement about authorization concerns the case in which the distinction between authorized and unauthorized states is absent from the observation. Let $Z \in \{0, 1\}$ indicate whether an action is authorized, and let N be the neural observation available to the decoder.

Theorem 8.2 (Authorization Non-Identifiability)**.** *Suppose* $\Pr(N \in B \mid Z = 0) = \Pr(N \in B \mid Z = 1)$ *for every measurable set* B*, and that the two authorization states have equal prior probability. Then every decoder* g*, including randomized decoders, has classification error* $\Pr(\widehat{Z} \neq Z) = \frac{1}{2}$.

Proof. For a deterministic decoder let $A = \{n : g(n) = 1\}$. The error probability is $\Pr(Z = 0, N \in A) + \Pr(Z = 1, N \notin A)$, which under equal priors equals $\frac{1}{2}\Pr(N \in A \mid Z = 0) + \frac{1}{2}\Pr(N \notin A \mid Z = 1)$. By the assumed equality of conditional distributions this is $\frac{1}{2}\Pr(N \in A \mid Z = 1) + \frac{1}{2}\left[1 - \Pr(N \in A \mid Z = 1)\right] = \frac{1}{2}$. A randomized decoder is a mixture of deterministic rules, and the mixture of quantities each equal to $\frac{1}{2}$ is $\frac{1}{2}$. □

The general form of the result gives the minimum achievable error for equal priors as

$$P_e^* = \tfrac{1}{2}\big(1 - \mathrm{TV}(P_0, P_1)\big),$$

where P_0, P_1 are the conditional distributions of N under unauthorized and authorized states and TV is total variation distance. When the distributions overlap substantially authorization cannot be reliably inferred from that observation alone. This is not a claim that authorization is never neurally decodable. It is a conditional impossibility: if the relevant distinction is absent from the signal, no decoder, however sophisticated, can recover it. A larger recording system may improve distinguishability, but increasing the number of measurements does not guarantee it.

8.3 Authorization as a Property of a Protocol

The more interesting point is that authorization is not best understood as a neural property at all. Austin's analysis of performative utterances observed that to promise, to bet or to authorize is to perform an act whose validity depends on conventions and circumstances and not on an inner state alone [2]. On this understanding Z is a property of a history relative to a protocol π, which specifies which performances count as authorizing which actions.

Assumption 8.3 (Rehearsal Indistinguishability; adopted without defense)**.** There exist a rehearsed authorizing act and a performed authorizing act whose neural trajectories coincide to the resolution of the decoder.

Proposition 8.4 (Protocol Dependence)**.** *Let* ν *map histories to what the decoder observes, and let* h_1, h_2 *be histories with* $\nu(h_1) = \nu(h_2)$ *and* $Z_\pi(h_1) \neq Z_\pi(h_2)$*. Then no observation-only*

decoder $g \circ \nu$ is correct on both histories. Under Assumption 8.3 such a pair exists, namely a rehearsed and a performed authorizing act.

Proof. The function $g \circ \nu$ takes one value on h_1 and h_2 because their observations coincide, whereas Z_π takes different values on them. Under the assumption, the rehearsal does not count as authorization under π and the performance does. □

The first sentence of the proposition requires nothing about neural hardware or its costs. It is a statement about the relation between observations and protocols, and it holds for any observation map whatever, which is the reason it is the philosophical centerpiece of the argument.

The reader who finds Assumption 8.3 implausible is invited to find it so, in which case the Main Hypothesis is correspondingly less well supported. The practical remedy for the difficulty is a protocol that includes a distinguishing event, such as an explicit confirmation gesture, and the distinguishing event is, functionally, a button.

8.4 Even a Perfect Decoder Does Not Authorize

A decoder that was perfect with respect to the user's neural state would still not settle whether an action should be executed, because the consequences of error differ between actions. Let C_{FP} be the cost of executing an unauthorized action and C_{FN} the cost of failing to execute an authorized one, and let $\varpi(n) = \Pr(Z = 1 \mid N = n)$.

Theorem 8.5 (Optimal Execution Threshold)**.** *Under the two-action loss model, the Bayes-optimal decision executes if and only if*

$$\varpi(n) \geq \frac{C_{\mathrm{FP}}}{C_{\mathrm{FP}} + C_{\mathrm{FN}}},$$

either decision being permissible at equality.

Proof. The conditional expected loss of executing is $C_{\mathrm{FP}}\,[1 - \varpi(n)]$ and that of refusing is $C_{\mathrm{FN}}\varpi(n)$. Execution is optimal when $C_{\mathrm{FP}}\,[1 - \varpi(n)] \leq C_{\mathrm{FN}}\varpi(n)$, that is, when $C_{\mathrm{FP}} \leq (C_{\mathrm{FP}} + C_{\mathrm{FN}})\varpi(n)$, which rearranges to the stated threshold. □

Neural confidence is therefore insufficient to determine whether an action should be performed, since the consequences of error also matter. If unauthorized execution is very costly the optimal threshold approaches one, and a system adequate for moving a cursor may be unacceptable for deleting a database, transferring funds or operating dangerous machinery. Moreover, even a Bayes-optimal rule does not establish moral or legal authorization; it minimizes expected loss under an assumed model.

Standing of the Main Hypothesis. Chapter 8 has proved that authorization cannot be recovered from signals that do not contain it, that it is a property of protocols and not of tissue, and that no accuracy figure fixes an execution threshold. It has also noted that the

same chapter proves nothing about whether any actual neural signal contains authorization. The functional component of the Main Hypothesis has acquired a philosophical centerpiece. The comparative component has acquired no additional evidence, and the hypothesis as a whole therefore remains a hypothesis.

Chapter 9

The Regress of the Confirmation Command

9.1 Asking Again

Suppose an interface has some probability of issuing an erroneous command. To reduce errors, the designer introduces a confirmation step. The confirmation signal must itself be decoded, and it has an error probability of its own, so the designer may introduce a confirmation of the confirmation, and the resulting sequence $C_0 \to C_1 \to C_2 \to \cdots$ recalls the regress described by Carroll, in which a tortoise declines to draw a conclusion until the rule of inference has been added to the premises, and then declines again until the rule that licenses the addition has been added [4]. Practical systems terminate such regresses by acceptance thresholds, redundancy, independent channels and risk-sensitive procedures, and the regress is therefore not a proof that confirmation is impossible. It exposes the necessity of an external stopping rule: the decoder cannot establish the legitimacy of its own interpretation merely by interpreting additional signals, unless assumptions about the relation among the signals are introduced.

9.2 Perfectly Correlated Errors

The simplest formal statement concerns the extreme case in which the errors of repeated attempts are identical.

Proposition 9.1 (No Redundancy Gain Under Perfect Correlation)**.** *Let $E_1, \ldots, E_n$ be binary error indicators for repeated confirmation attempts, with $E_1 = E_2 = \cdots = E_n$ almost surely and $\Pr(E_1 = 1) = p$. For any odd n, a majority-vote confirmation rule has error probability exactly p.*

Proof. Because the indicators coincide almost surely, either every confirmation is correct or every confirmation is incorrect. The majority is incorrect exactly when $E_1 = 1$, which has probability p. □

By contrast, if the errors are conditionally independent with common probability $p < \frac{1}{2}$, the majority error is $\sum_{k=(n+1)/2}^{n} \binom{n}{k} p^k (1-p)^{n-k}$, which decreases toward zero as n grows. The crucial assumption is independence, or at least sufficiently weak dependence, and repeated decoding of the same biological state may share noise, model bias, contextual ambiguity and calibration error. Repeating a neural command is not automatically an independent verification. The system asks the brain whether it meant what the system believed it meant, using the same system that may have misunderstood it.

9.3 A Floor Under Latent States

The perfectly correlated case is an idealization, and a more realistic model supposes that errors are independent given a latent state of the system. Let S be a latent state with finitely many values and distribution ϖ_s, and let a_s be the probability, in state s, that a confirmation stage affirms a spurious command, the stages being independent given S. Let b_s be the corresponding probability that a stage affirms the intended command. With k confirmation stages in addition to the initial decode, the probability of a spurious execution is $\sum_s \varpi_s a_s^{k+1}$ and the probability of a successful execution of an intended command is $\sum_s \varpi_s b_s^{k+1}$.

Theorem 9.2 (Confirmation Floor)**.** *The spurious-execution probability is nonincreasing in k and converges to $\sum_{s:a_s=1} \varpi_s$. The success probability is also nonincreasing in k and converges to $\sum_{s:b_s=1} \varpi_s$. In particular, if some state with positive probability has $a_s = 1$, the spurious-execution probability is bounded below by that probability for every k.*

Proof. Each term a_s^{k+1} is nonincreasing in k because $0 \le a_s \le 1$, converges to 0 if $a_s < 1$ and equals 1 for all k if $a_s = 1$. A finite sum of such terms inherits both properties, and the same argument applies to b_s. □

The states with $a_s = 1$ correspond to saturating conditions: a decoder stuck on a class, a sustained involuntary pattern, a failed channel. In those states no number of confirmations helps, because every confirmation fails in the same way. Meanwhile the success probability falls with k in every state in which $b_s < 1$, so that each added confirmation taxes the intended commands as well as the spurious ones. The trade-off is familiar from any serial system, and it implies that the number of confirmations is itself a parameter to be chosen with regard to the costs of Theorem 8.5, and that no value of the parameter makes a correlated channel into an independent one.

Standing of the Main Hypothesis. Chapter 9 has proved that repeated confirmation fails under perfect correlation and has a floor under latent failure states, and it has noted that under independence repeated confirmation succeeds. The existence of both regimes means that the chapter supports neither the Interface Elimination Thesis nor the Main Hypothesis, and that which regime holds for a given decoder is a question about that decoder.

Chapter 10

Drift and the Maintenance Burden

10.1 The Stale Decoder

A button is not calibrated. A neural interface is, and the calibration decays. Recorded neural populations change over days and weeks in the signals they carry about the same behavior, a phenomenon described as representational drift in cortical areas [14], and instabilities within a day have been documented in intracortical interfaces [30]. The consequence is that a decoder trained at one time is a decoder for a population that no longer quite exists.

A minimal model suffices. Let the neural gain relating a scalar intention z to the observation be $g_t = g_0 + \xi_t$, where ξ_t is a random walk with $\mathrm{Var}(\xi_t) = t\sigma^2$, and let a fixed decoder estimate $\hat{z} = N_t/g_0$ from $N_t = g_t z$ in the absence of measurement noise.

Proposition 10.1 (Staleness Horizon)**.** *The mean squared error of the fixed decoder is* $\mathbb{E}\left[(\hat{z} - z)^2\right] = t\,v$*, with* $v = \sigma^2 \mathbb{E}[z^2]/g_0^2$ *when* z *is independent of the walk. Consequently the time at which the error reaches a tolerance* τ *is* $t^* = \tau/v$*.*

Proof. We have $\hat{z} - z = (g_t z)/g_0 - z = \xi_t z/g_0$, so $\mathbb{E}[(\hat{z} - z)^2] = \mathbb{E}[\xi_t^2]\mathbb{E}[z^2]/g_0^2 = t\sigma^2 \mathbb{E}[z^2]/g_0^2$ by independence, and setting this equal to τ gives $t^* = \tau/v$. □

10.2 An Irreducible Maintenance Rate

If recalibration restores the decoder at a fixed cost κ per event and the cost of error accrues at a rate c per unit of mean squared error, then recalibrating every T units of time gives a cost per unit time of

$$\rho(T) = \frac{\kappa}{T} + \frac{c\,v\,T}{2},$$

the first term being the amortized recalibration and the second the average error cost over a cycle.

Proposition 10.2 (Irreducible Maintenance Rate)**.** *The cost rate* ρ *is minimized at* $T^* = \sqrt{2\kappa/(cv)}$*, with minimum value* $\sqrt{2\kappa c v}$*, which is strictly positive whenever* $v > 0$ *and* $\kappa > 0$*.*

Proof. Setting $\rho'(T) = -\kappa/T^2 + cv/2 = 0$ gives $T^{*2} = 2\kappa/(cv)$, and $\rho'' > 0$ confirms a minimum. Substituting, $\rho(T^*) = \kappa\sqrt{cv/(2\kappa)} + (cv/2)\sqrt{2\kappa/(cv)} = 2\sqrt{\kappa cv/2} = \sqrt{2\kappa cv}$. $\square$

The proposition does not say that the rate is large, only that it is not zero, and the corresponding rate for an idealized button is zero. Self-calibrating decoders reduce κ by inferring the user's intended targets retrospectively from the structure of the task, as in typing interfaces in which the selected key is taken to have been the target [21]. The method is ingenious and it depends on a discrete set of targets whose identity can be inferred after the fact, which is to say on an alphabet. The keyboard rescues the decoder, and the rescue is reported without irony in the literature, which is a restraint worthy of the genre.

Standing of the Main Hypothesis. Chapter 10 has proved that a drifting channel requires recalibration at a rate bounded away from zero, under a model that the chapter has not validated against any recording. It has not shown that the rate is large, that it exceeds the maintenance cost of any particular conventional device, or that the idealized zero for the button survives contact with wear, dust and spilled coffee. The Main Hypothesis stands unsupported.

Chapter 11

Echo, Undo and the Ledger of Commands

11.1 The Other Half of the Loop

A neural cursor is a cursor. The decoded command must be displayed to the user, for otherwise the user cannot tell what has been decoded, and the display is an output interface of precisely the kind that conventional systems provide. A closed-loop neural interface therefore eliminates the input half of the human-computer loop and retains the output half in full, and the claim to have eliminated the interface should be read as a claim to have eliminated half of one.

Observation 11.1 (Echo Is Interface)*.* Any closed-loop use of a decoder requires a channel from the machine's state to the user, and the channel's design (what is shown, how soon and in what form) determines the effective cost L and error E of the interaction. The elimination of the input device does not reduce the design problem of the output channel.

11.2 Undo Requires a Record

Error correction in conventional systems depends on the possibility of undoing what was done. The standard implementation maintains a record of executed commands, each with a defined inverse, and the undo operation pops the record. The record is a ledger of discrete symbols. A continuous trajectory of neural signals does not directly admit such an inverse. Continuous-state systems can support restoration through checkpoints, inverse transformations or stored trajectories, so an alphabet is not strictly necessary, but each of these requires a finite stored record, and the records kept in conventional systems are in practice sequences of discrete commands. In such systems the capacity to take an action back is purchased with a finite vocabulary of actions, and the ledger is a button's memory.

Observation 11.2 (Undo Requires a Record)*.* An undo operation is a function from histories to prior states. Any finite store that supports it retains a summary of the history, and in conventional systems the summary is the sequence of commands executed.

The observation is stated informally, because the formalization depends on what is meant by a state, and the book's stock of assumptions is already large. The point is practical: the thing that makes a mistaken command recoverable is a log, and the log is usually a record of discretized commands, which are what the interface was supposed to make unnecessary.

Standing of the Main Hypothesis. Chapter 11 has observed that an elimination of input leaves output and correction in place, and that correction needs a ledger. The observations are informal and have not been compared with any measured error rate. The Main Hypothesis remains a hypothesis, and the chapter concludes with its customary confession that this is where it started.

Part III

The Industrial Reinvention of the Human Hand

Chapter 12

Standards, Lifecycles and Orphaned Users

12.1 The Interchangeable Peripheral

The conventional interface owes much of its economic character to the fact that it is a peripheral. A keyboard may be unplugged, replaced with another from a different manufacturer, repaired by a stranger, borrowed, left behind or thrown away, and none of these operations requires the consent of its original maker. The common conventions governing connectors, protocols and layouts mean that the user's skill is portable across devices: a person who has learned the layout on one keyboard can use any other. The standardization is an achievement of dull, unglamorous, cooperative engineering, and it has given the conventional interface properties that have no counterpart in an individually calibrated device, namely that the skill belongs to the user and not to the vendor.

An implanted decoder differs on each point. The hardware cannot be unplugged, the decoder is trained to a particular person and a particular device, and the knowledge of how to interpret the person's signals resides in parameters held by whoever maintains the software. To the extent that no open convention exists for such decoders, the user's relationship to the device is a dependency on a supplier. The book does not claim that no such convention could exist.

12.2 The Orphan Interval

Let T_b be the duration for which a user depends on a device and T_s the duration for which its supplier supports it. The user is orphaned during the interval (T_s, T_b) when it is nonempty.

Proposition 12.1 (Positive Expected Orphaning)**.** *The expected orphan duration* $\mathbb{E}\left[(T_b - T_s)^+\right]$ *is strictly positive if and only if* $\Pr(T_s < T_b) > 0$*.*

Proof. The random variable $(T_b - T_s)^+ = \max(T_b - T_s, 0)$ is nonnegative and positive exactly on the event $\{T_s < T_b\}$. Its expectation is therefore positive if and only if that event has positive probability. □

The proposition is trivial and the practical point is not. Companies are shorter-lived than bodies. Reporting in 2022 described users of a retinal prosthesis whose manufacturer had ceased to support it, who were left with implanted devices that could not be updated or repaired [38]. A discontinued keyboard is an annoyance resolved at a shop; a discontinued implant is a medical and a legal problem of a different order. This observation concerns the lifecycle of any implanted device and is not specific to neural interfaces that decode intent, but the asymmetry between implanted and detachable peripherals is part of what the Interface Elimination Thesis leaves out when it counts physical intermediaries as pure cost.

Standing of the Main Hypothesis. Chapter 12 has proved that orphaning has positive expected duration whenever it has positive probability, which is a proof of a tautology. It has not shown that the probability is large for any particular device, and it has recorded one case from another kind of implant. The Main Hypothesis, which is not about retinal prostheses, has not been advanced.

Chapter 13

The Right Not to Execute

13.1 Friction as a Constitutive Feature

An interface mediates between the possible, the considered and the committed. The delay, discretization and friction introduced by conventional interfaces are not necessarily defects awaiting technological elimination, because some of them are constitutive features of reliable and deliberate agency. The point is old and widely applied. Institutions that handle dangerous capabilities routinely require that two persons act in concert, that two keys be turned at separated locations, or that a second signature follow the first, and the designs are intentionally slow. Individuals commit themselves in advance to courses of action in order to remove the option of revising them under pressure, a practice that Elster analyzed through the figure of Ulysses bound to the mast [15]. In these cases friction is an instrument of governance.

The point can be given a simple formal form. Let a user form an intention at time s and let R be the random time, measured from s, at which the user would retract it on further deliberation, with a continuous distribution function F satisfying $F(0) = 0$. An interface that imposes a delay Δ before execution prevents the execution of exactly those intentions that are retracted within the delay.

Proposition 13.1 (Monotone Value of Delay)**.** *The probability that an intention which would be retracted is prevented from executing is $F(\Delta)$, which is nondecreasing in Δ and equals zero at $\Delta = 0$.*

Proof. Execution is prevented exactly when retraction occurs before $s + \Delta$, that is, when $R < \Delta$, which has probability $F(\Delta)$ by the continuity of F. This is nondecreasing as a distribution function, and $F(0) = 0$ by hypothesis. □

An interface with no delay thus cannot benefit from the user's second thoughts at all. If h is the harm of executing an intention that would have been retracted and β the cost of latency per unit time, the net benefit of a delay Δ is $G(\Delta) = hF(\Delta) - \beta\Delta$.

Proposition 13.2 (Latency Minimization Is Not the Objective)**.** *If F is differentiable at 0 with $hF'(0) > \beta$, the maximizer of G over $\Delta \geq 0$ is strictly positive.*

Proof. We have $G(0) = 0$ and $G'(0) = hF'(0) - \beta > 0$, so $G(\Delta) > 0$ for small $\Delta > 0$. Since $G(\Delta) \leq h - \beta\Delta \to -\infty$, the maximum over $\Delta \geq 0$ is attained, and it cannot be attained at 0. □

Zero latency is therefore not the optimum of an interface unless second thoughts are rare or the harm of acting on them is small. The claim that an interface should be as fast as thought is a claim about the parameters h and $F'(0)$, and the parameters depend on what the interface is used for. Combined with Theorem 8.5, which shows that the confidence required for execution depends on the cost of a mistaken execution, the result gives a precise sense in which an interface is a device for governing the passage from the possible to the committed, and not merely a device for reducing the time between a wish and its fulfillment.

13.2 Neural Data and the Boundary of the Mind

A neural record is a record of what a person has available, considers and prepares, at least as much as of what a person has committed to. Writers on the ethics of neurotechnology have argued that this creates distinctive interests in mental privacy, cognitive liberty and the integrity of the mind, and have proposed new protections [20, 42]. The argument of this chapter is complementary: an interface that reads preparation as commitment threatens not merely privacy but the structure of agency, since it removes the interval in which a person can decide not to do what the person has begun to prepare. The right not to execute is the right to have that interval, and a button provides it more cheaply than anything else yet devised.

Standing of the Main Hypothesis. Chapter 13 has shown that a positive delay is optimal when second thoughts are common and costly to ignore, and that the confidence required for execution depends on the stakes. Neither result compares a neural interface with a conventional one, since delays and thresholds can be added to either. The chapter has therefore established a principle of interface design that is neutral between the systems it was intended to discredit. The Main Hypothesis is not supported by it.

Chapter 14

The Economics of Universal Adoption

14.1 The Adoption Inequality

Let J_c denote the cost to a particular user of operating a conventional interface and J_n the cost of a neural alternative, both measured by the functional of Chapter 4. A user whose only concern is cost adopts the neural alternative if and only if $J_n < J_c$. The difference decomposes as

$$J_c - J_n = \alpha(E_c - E_n) + \beta(L_c - L_n) + \gamma(U_c - U_n) - \delta(K_n - K_c) - \epsilon(R_n - R_c).$$

Proposition 14.1 (Adoption Inequality)**.** *Suppose the conventional interface carries no calibration burden and no safety risk, $K_c = R_c = 0$, and the neural interface has $K_n \geq 0$, $R_n \geq 0$ and nonnegative operating costs E_n, L_n, U_n. Then $J_n < J_c$ implies*

$$\alpha E_c + \beta L_c + \gamma U_c > \delta K_n + \epsilon R_n.$$

Proof. Since $E_n, L_n, U_n \geq 0$ we have $J_c - J_n \leq \alpha E_c + \beta L_c + \gamma U_c - \delta K_n - \epsilon R_n$, and adoption requires $J_c - J_n > 0$. □

The proposition says that the operating cost of the conventional interface must exceed the maintenance and risk cost of the neural one. For a person whose conventional channel is unusable, the operating cost is very large, since the task cannot be done at all by ordinary means, and the inequality is satisfied. This is the situation of the clinical programs, and it is the reason for their success as a matter of cost. For a person with a working keyboard and a working hand, the operating cost of the conventional interface is bounded by what the user already pays for it, and the question is whether that amount exceeds $\delta K_n + \epsilon R_n$ for a device that requires surgery. The book does not claim that it does not. The inequality is an empirical condition for each user, and the assertion that universal adoption is inevitable is an assertion that the condition holds for everyone.

14.2 The Disappearance of the Interface as a Product Category

A further difficulty is rhetorical. The phrase *no interface* has the attractions of a slogan: it is easy to say, it names a state of affairs that sounds like pure gain and it invites no examination of what the interface did. A technology that reduces the number of visible parts is easily described as simpler, and simplicity is readily mistaken for a reduction in total work. The accounting of Chapter 4 is a mild corrective, and the inequality of this chapter supplies a more specific one: the claim that the interface has disappeared is a claim about where it is visible, and visibility does not occur in J.

Standing of the Main Hypothesis. Chapter 14 has derived a necessary condition for adoption and has observed that it is easily met for users who have lost the conventional channel and is a matter of fact for everyone else. The chapter has not measured any term of the inequality, and its conclusion that adoption depends on circumstances is a conclusion that nobody has disputed. The Main Hypothesis is unsupported.

Chapter 15

Exit, Dependence and the Displacement of Governance

15.1 The Interface as a Small Institution

An interface is a small institution. It fixes what can be said to a machine, in what form, at what cost and subject to whose approval, and the person who uses it is governed by those fixings in a way that is none the less real for being embodied in plastic and software. Lessig's observation that the architecture of software regulates conduct as effectively as law does is the general form of the point [10], and a platform is an interface scaled up until it governs a marketplace, a community or a population. The Interface Elimination Thesis has a political cousin, which holds that governance, like the interface, is an intermediary that technology can remove, so that a person equipped with the right tools can exit the authority of the territorial state and become independent of authority as such. The cousin shares the family fault. The question it raises is the one the preceding chapters have asked of the button: when a mechanism is removed, where does its function go?

15.2 The Thesis in Its Own Terms

The thesis has a documented intellectual history, and fairness requires that it be stated before it is examined. Davidson and Rees-Mogg argued in 1997 that the viability of the nation-state depends on its ability to identify, locate, tax and regulate economically productive individuals, that digital, mobile and cryptographically protected economic activity weakens that ability, and that governments would accordingly be obliged to compete for mobile capital in place of exercising unquestioned territorial authority [8]. Thiel wrote in 2009 that he no longer believed freedom and democracy to be compatible, that the education of the electorate had become a fool's errand for libertarians, and that the hope for freedom lay in technological frontiers, which he identified as cyberspace, outer space and the sea [13]; he later answered objections to his remarks on suffrage in the same forum. Srinivasan proposed in 2022 the network state, a community that forms online around a shared purpose, acquires territory

and seeks diplomatic recognition [12], and he acknowledges the earlier book as an antecedent.

Three kinds of claim are easily confused in this territory, and the chapter relies on only one of them. Intellectual affinity, which is a similarity of argument; documented influence, which requires evidence that one author read and was moved by another; and retrospective historical reconstruction, which assembles a narrative from both. Some recent polemical writing, including the introduction to *The Nerd Reich*, asserts the second of these: that the 1997 book had a profound effect on Thiel's later trajectory. That is a causal proposition requiring biographical and documentary evidence, and it is not established by similarity of ideas alone. The two readings of the program, as liberation of the individual from coercive institutions and as the foundation of a new oligarchy, are not logically incompatible, since a technology can decentralize authority relative to the state while concentrating effective power among those who control the infrastructure. This chapter asserts affinity of structure between the thesis and the Interface Elimination Thesis and makes no claim about influence or about the motives of any person. The projects associated with these writers are also not ideologically identical: some are explicitly anti-state, some propose to replace existing institutions, and some seek influence over the state itself. A framework that treated them as a single doctrine would be a framework of the kind this book exists to examine.

15.3 Displacement of Dependence

Let $D(G, T)$ denote a population's dependence on territorial governance as a function of technological capacity T, and let $P(T)$ denote its dependence on private technological authorities such as platforms, payment networks, identity providers and the maintainers of decoders. The sovereign-individual thesis asserts, in effect, that $\partial D/\partial T < 0$. Total dependence is

$$D_{\text{total}}(T) = D(G, T) + P(T),$$

and the following proposition records that the sign of the first term does not determine the sign of the total.

Proposition 15.1 (Non-Determination of Total Dependence)**.** *There exist smooth D and P with $\partial D/\partial T < 0$ and $P'(T) > 0$ for which D_{total} is strictly decreasing, and others for which it is strictly increasing.*

Proof. Take $D(T) = d_0 e^{-T}$ and $P(T) = p_0(1 - e^{-T})$ with $d_0, p_0 > 0$. Then $D' < 0$ and $P' > 0$, and $D'_{\text{total}}(T) = (p_0 - d_0)e^{-T}$, which is negative when $p_0 < d_0$ and positive when $p_0 > d_0$. □

The proposition is as slight as it looks, and it is the first logical problem in the sovereign-individual argument: escaping one authority is not equivalent to becoming independent of authority. Whether technological decentralization lowers total dependence is a question about the magnitudes of D and P, which cannot be read off from the fact that one of them falls.

15.4 Conflicting Claims

A second problem concerns the universality of sovereignty. Suppose that sovereignty means the capacity to decide the disposition of a resource free of any other agent's unilateral discretion.

Proposition 15.2 (Incompatibility of Universal Unilateral Sovereignty)**.** *If agents a and b each hold unrestricted unilateral authority over the disposition of a resource r and their desired dispositions are mutually exclusive, no feasible allocation satisfies both decisions.*

Proof. A disposition of r satisfying a's decision excludes b's by hypothesis, and conversely. At least one agent must therefore revise, subordinate, negotiate or abandon the original claim. □

The absence of a state does not remove the problem of conflicting claims. It changes the mechanisms through which they are resolved, which may be bargaining, arbitration, exclusion, force or an agreed rule system. The proposition is not a proof that the nation-state is necessary, and the book has no wish to prove it; it shows that unrestricted sovereignty cannot be universally realized under conflicting claims, and that a procedure for adjudication exists in every arrangement and is either public or private.

15.5 Formal Exit and Effective Exit

Hirschman distinguished exit, the withdrawal of a member from a relationship, from voice, the attempt to change it from within [9], and observed that the two interact: easy exit can discipline a provider and can also drain the pressure of voice. The distinction between exit that is formally available and exit that is effective is the center of the matter. Let user i derive value v_i from remaining with an institution, net of nothing, let the institution charge a fee f, let a_i be the value of the best alternative, and let C_i be the cost to i of switching, including the loss of relationships, reputation and accumulated history that are not portable. The user stays if and only if $v_i - f \geq a_i - C_i$. Writing $B_i = a_i - (v_i - f)$ for the benefit of leaving, effective exit is available exactly when $B_i > C_i$.

Proposition 15.3 (Rent from Exit Cost)**.** *The maximal fee that the institution can charge user i without inducing exit is $f_i^* = v_i - a_i + C_i$, which increases one-for-one with the switching cost C_i whatever the values of v_i and a_i.*

Proof. The user stays if and only if $f \leq v_i - a_i + C_i$, so the supremum of fees consistent with staying is $v_i - a_i + C_i$, and the derivative with respect to C_i is 1. □

The mere existence of an exit option therefore demonstrates nothing about independence. An institution that can raise switching costs can raise what it extracts by the same amount without providing any additional value, and an institution may become more extractive precisely because its operators know that users cannot afford to leave. It follows that revenue can be derived from continued participation even where the participants' own expected returns are negative, and that the territorial reach of the institution is irrelevant to the harm it can do

to those who depend on it. The practical corollary is the subject of Chapter 12: portability, which sets C_i toward zero, is a governance instrument. A standard that lets a person take skills, records and relationships elsewhere is the cheapest form of constitutional protection yet devised, and its absence is a decision about who rules.

15.6 The Historical Irony

There is a historical possibility here that deserves to be stated cautiously. Feudal arrangements often involved overlapping jurisdictions, privately exercised authority, unequal capacities for exit and obligations mediated through powerful intermediaries. Digital platforms can exhibit structurally similar characteristics without being feudal institutions in any literal sense. A platform may control access to markets, communication, identity, reputation, payment and dispute resolution, and its users may possess the formal freedom to leave while finding the cost prohibitive. The political ideal of technological sovereignty might then reproduce, in new materials, features of the order that information technology was predicted to transcend. The qualification is essential: this is a structural analogy and not an identification, and the analogy does not show that any actual platform is a fief.

The notion of domination supplies a useful criterion. Pettit's republican account locates unfreedom in the capacity of another party to interfere arbitrarily, that is, at will and without being forced to track the interests or the contestation of the person affected [11]. On this account the question to put to any governing arrangement, whether public or private, is not whether it is called a government but whether its power over those subject to it is arbitrary. The same criterion applies to states, which also exercise arbitrary power, and the chapter makes no claim that private institutions dominate more than public ones.

15.7 The Sovereignty Displacement Thesis

The foregoing results support a conditional statement of the thesis that is more modest than the Main Hypothesis and more likely to be correct.

Definition 15.4 (The Sovereignty Displacement Thesis)**.** A technological system that reduces dependence on one governing institution does not thereby reduce domination. Unless it also reduces arbitrary discretionary power, lowers effective exit costs and preserves meaningful opportunities for contestation, it may merely transfer governance from publicly accountable institutions to privately controlled infrastructures.

The thesis is conditional, and it is not an assertion that every private institution is more dominating than every state. A decentralized arrangement could genuinely improve autonomy if it offers interoperable alternatives, low switching costs, distributed control and enforceable protections against arbitrary exclusion, and some such arrangements exist. Whether a given project satisfies the conditions is a question to be answered by examining the project: by measuring the distribution of switching costs among its participants, by identifying the procedures available for contesting its decisions, and by recording how often and on what

grounds participants have been excluded without recourse. These are observable, and the thesis is falsified for a given project by finding that it satisfies the conditions, which the book neither asserts nor denies of any.

A tension in the position of its proponents deserves to be named without being resolved. If technology allows genuine exit from the state, it is not obvious why those who hold that view should invest substantially in influencing state institutions. There are practical answers, and the strongest is that states regulate the conditions of exit itself, through taxation, capital controls, recognition and the law of contract, so that a rational exit strategy includes a political strategy for keeping exit available. The tension is accordingly a feature of the position and not a contradiction in it.

15.8 The Interface Once More

The return to the interface is direct. To the extent that the maintenance, the parameters and the support decision lie with a vendor, which is an institutional risk and not a necessary property of neural interfaces, a person who has eliminated the button depends on whoever maintains the decoder, whoever holds the parameters that map that person's neural activity to commands, and whoever decides whether the device will be supported next year (Chapter 12). In the vocabulary of this chapter the dependence D on the physical intermediary has fallen and the dependence P on the vendor has risen, and Proposition 15.1 says that the sign of the total is not given. The switching cost C_i of a surgically implanted, individually calibrated device is among the highest imaginable, and by Proposition 15.3 the rent extractable from such a user rises accordingly. The right not to execute (Chapter 13) is the user's voice, the interval in which a person can contest a command before it takes effect, and an open standard for decoders is the user's exit. Governance, like the interface, cannot be eliminated by being relocated to a place where it is hard to see. The absurdity is not that technologists want freedom from government. It is that an attempt to abolish governance by reconstructing its essential functions under proprietary terms of service, while declaring that governance has been eliminated, is the same mistake as the Imagined Button on a larger scale.

Standing of the Main Hypothesis. Chapter 15 has proved that the sign of one term of total dependence does not fix the sign of the sum, that unrestricted sovereignty cannot be universal under conflicting claims, and that the rent extractable from a user rises with the user's switching cost. It has stated a conditional thesis and has said what would test it. It has not shown that any particular project fails the conditions, that private institutions dominate more than public ones, or anything whatever about the motives of the persons whose writings it has described. The Main Hypothesis is not advanced, and the chapter has widened the territory over which it fails to be advanced.

Chapter 16

A Modest Proposal for the Imagined Button

16.1 The Device

A natural culmination of the program of interface elimination can now be described. A user wishing to issue a command imagines pressing a button. The implant detects the neural correlates of the imagined press, applies a threshold, and so concludes that a press has occurred. The machine, having no button of its own, displays a virtual button on the screen, and the user is then asked to confirm that the imagined press was intentional, by imagining pressing the displayed button. The implant decodes the second imagined press, the machine executes the command, and the developers announce that humanity has been liberated from the tyranny of the physical button.

Definition 16.1 (The Imagined Button System)**.** The Imagined Button System consists of an implant that decodes attempted presses, a display that echoes the decoded press as a virtual button, and a second decoding of an attempted press that serves as confirmation, after which the command is executed.

16.2 Dominance of the Button

Compare the system with a physical button for a user able to press it. The comparison assumes that both observe the same deliberate act, the attempted press A: the button observes it through the contact, which we idealize as exact, and the implant observes it through its neural correlates N. This is an assumption and not a fact, since an imagined or attempted press need not be the same act as a physical press, and the theorem concerns only the case in which it is.

Assumption 16.2 (Idealized Button)**.** The physical button senses the user's attempted press exactly, has no calibration burden and carries no surgical risk. Both systems observe the same attempted press, produced with the same act-generation latency, and the stages that the Imagined Button System adds (decoding, echo and confirmation) have nonnegative latency.

Theorem 16.3 (Dominance of the Button)**.** *Under Assumption 16.2, for a user able to press the physical button, the Imagined Button System delivers no more information about the intended command than the button, and is no better than the button in latency, calibration burden or risk.*

Proof. For information, the chains $\Theta \to A \to \widehat{\Theta}_B$ and $\Theta \to A \to N \to \widehat{\Theta}_{IB}$ are Markov, with $\widehat{\Theta}_B$ a deterministic function of A by the assumption of exact sensing. By Theorem 3.1 applied twice, $\mathsf{I}(\Theta; \widehat{\Theta}_{IB}) \leq \mathsf{I}(\Theta; A) = \mathsf{I}(\Theta; \widehat{\Theta}_B)$. For latency, the system executes the stages of imagined press, decoding, echo, second imagined press and second decoding in series, so its latency is the common act-generation latency plus nonnegative added stage latencies, and is at least that of the single press. For calibration burden and risk, the idealization gives the button zero in both, and the implant cannot have less than zero. □

The Imagined Button System is therefore dominated by the button in every component of J except the one that does not appear in it, namely the number of physical intermediaries visible on the desk, on which it is admittedly superior. The theorem concerns a user able to press, and it says nothing about a user who cannot, for whom the comparison does not arise and the imagined button is the only button available. That is the clinical case, and the book has already declined to speak against it.

16.3 A Note on Imagery

The instruction to imagine a press presupposes a capacity that varies among persons. Some people report no voluntary mental imagery whatever [43], and it is the attempted movement and not the visual image that motor decoders in current clinical use rely upon [41]. The device as described in the proposal is accordingly addressed to a particular kind of imaginer, and the careful designer would specify attempted action and not imagined sight. The distinction is a reminder that a command channel is a convention between a person and a machine, and that conventions presuppose competences that must be taught and may not be universal.

Standing of the Main Hypothesis. Chapter 16 has proved that the Imagined Button System is dominated by the button along every dimension of the cost functional for a user able to press. This is the closest the book has come to a result against the Interface Elimination Thesis, and it is a result about a device that nobody has proposed. The Main Hypothesis, which concerns the thesis and not the proposal, is unsupported in exact proportion to the thesis's distance from the proposal, which the author estimates, in lieu of a number, to be considerable.

Chapter 17

What Has Not Been Shown

The book has proved a number of propositions, and it is proper to record what they come to. A system that issues more than one command must distinguish among its observations, and must provide the further mechanisms of execution selection and error control (Theorems 1.2 and 5.1). A decoder cannot recover information that its observations do not contain, recording volume is not command information, and the ceiling on command information is set by the user (Theorems 3.1 and 3.2). The removal of an interface neither entails nor excludes a reduction of total work (Proposition 4.4). Every discretization has a boundary, hysteresis tames the boundary, and perfect classification is unavailable when observation distributions overlap (Theorems 6.1 and 6.3 and Proposition 6.2). Authorization cannot be recovered from signals that do not contain it, and it is a property of protocols (Theorem 8.2 and Proposition 8.4). Repeated confirmation helps under independence and fails under correlation (Proposition 9.1 and Theorem 9.2). A drifting channel requires maintenance at a rate bounded away from zero (Proposition 10.2). Delay has positive value when second thoughts are common (Proposition 13.2). Dependence removed from one authority may reappear as dependence on another, and the rent extractable from a user rises with the cost of leaving (Propositions 15.1 and 15.3). A device that asks the user to imagine a button, display it and confirm it is dominated by a button (Theorem 16.3).

None of these results establishes the Main Hypothesis. They do not show that neural interfaces are useless, that interface elimination is impossible, that any neural interface performs worse than any conventional one at any task, or that the Interface Elimination Thesis is untenable as distinct from unsupported. Nearly every result holds for conventional interfaces as well as neural ones, which is the property that makes the results correct and the property that makes them incapable of deciding the comparison. The strongest and most testable hypothesis in the neighborhood of the book is that conventional interfaces frequently outperform neural alternatives for users with intact hands, because conventional interfaces exploit deliberately constructed, observable and comparatively stable distinctions. That hypothesis could be established only by empirical comparison, and no deduction in the preceding chapters does the work of an experiment.

What the book has done, then, is to describe the position of someone who has been told that a device does away with the interface, and who responds by asking where the distinction,

the interpretation, the authorization and the correction have gone. They have gone where they had to go. A system measures a biological signal, estimates a latent intention, applies a decision threshold, checks for error, requests confirmation and finally generates a discrete command, and having thereby reconstructed every logical property of the button, declares the button obsolete. The declaration is not false; it is a statement about which part of the arrangement is visible. The conclusion of the book is that the delay, discretization and friction of conventional interfaces are not defects awaiting abolition but, in many cases, the means by which a person keeps the distinction between what is possible, what is considered and what is done. That conclusion is offered as an observation about design and not as a result, and the Main Hypothesis remains, as it has remained on every page, a hypothesis.

Appendix A

Standard Results Used

The results of information theory used in the text are standard and are collected here with the one proof that the main text omitted. The chain rule for mutual information, the data processing inequality and Fano's inequality may be found in [6], and Fano's original statement in [16]; the foundational definitions are due to Shannon [35].

Proposition A.1. *For equal priors on $Z \in \{0,1\}$ and conditional distributions P_0, P_1 of the observation, the minimum classification error over all decoders is $\frac{1}{2}\big(1 - \mathrm{TV}(P_0, P_1)\big)$.*

Proof. Let p_0, p_1 be densities of P_0, P_1 with respect to a common dominating measure μ. A decoder is a measurable set A on which it outputs 1, and its error is $\frac{1}{2}P_0(A) + \frac{1}{2}P_1(A^c) = \frac{1}{2}\left[1 - \big(P_1(A) - P_0(A)\big)\right]$. The difference $P_1(A) - P_0(A) = \int_A (p_1 - p_0)\, d\mu$ is maximized by $A = \{p_1 > p_0\}$, where it equals $\mathrm{TV}(P_0, P_1)$. Hence the minimum error is $\frac{1}{2}\left[1 - \mathrm{TV}(P_0, P_1)\right]$. $\square$

Appendix B

Notation and Glossary

Alphabet. A finite set $\mathcal{C}$ of commands containing the null command $\varnothing$. **Button.** An interface in which the user's deliberate act produces a discrete event directly (a source quantizer in the sense of Chapter 2). **Decoder.** The map from an observation, with a bounded history, to an executed command. **Interface.** Any mechanism producing an observation from the user's activity and an executed command from the observation. **Main Hypothesis.** The claim, unproved by this book, that the Interface Elimination Thesis is logically untenable. **Interface Elimination Thesis.** The claim that the work of distinction, interpretation, authorization and correction performed by conventional interfaces is dispensable overhead. **Standing.** The statement closing each chapter, recording the unchanged condition of the Main Hypothesis. **Symbols.** Θ intended command; $\widehat{\Theta}$ executed command; N neural observation; Z authorization indicator; I mutual information; H entropy; TV total variation distance; J interaction cost with components E, L, U, K, R for error, latency, effort, calibration and maintenance, and risk.

Appendix C

A Statement Concerning Assistive Neural Interfaces

Interfaces that decode neural activity to restore communication and control to people with paralysis, anarthria or loss of limb are a distinct matter from the thesis examined in this book. They address a real loss, they replace a channel that no longer functions, and they have produced results of a kind that would have been regarded as implausible a generation ago [19, 39, 26, 40, 25]. The propositions proved here apply to them as they apply to all interfaces, and what they imply is that these systems must perform, and evidently do perform, the work of distinction, interpretation, authorization and correction, and that the achievement of the engineers consists in having done it. Nothing in the book is a reason to doubt the value of that achievement. The Interface Elimination Thesis is a different matter, because it extends a justified result for those who have lost a channel to a prediction about those who have not, and the extension is what the book has examined and has not refuted.

Bibliography

[1] Ashby, W. Ross. *An Introduction to Cybernetics*. London: Chapman and Hall, 1956.

[2] Austin, J. L. *How to Do Things with Words*. Oxford: Clarendon Press, 1962.

[3] Card, Stuart K., Thomas P. Moran, and Allen Newell. *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1983.

[4] Carroll, Lewis. "What the Tortoise Said to Achilles." *Mind* 4(14): 278–280, 1895.

[5] Collinger, Jennifer L., Brian Wodlinger, John E. Downey, Wei Wang, Elizabeth C. Tyler-Kabara, Douglas J. Weber, Angus J. C. McMorland, Meel Velliste, Michael L. Boninger, and Andrew B. Schwartz. "High-Performance Neuroprosthetic Control by an Individual with Tetraplegia." *The Lancet* 381(9866): 557–564, 2013.

[6] Cover, Thomas M., and Joy A. Thomas. *Elements of Information Theory*. 2nd ed. Hoboken, NJ: Wiley-Interscience, 2006.

[7] Dretske, Fred. *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press, 1981.

[8] Davidson, James Dale, and William Rees-Mogg. *The Sovereign Individual: How to Survive and Thrive During the Collapse of the Welfare State*. New York: Touchstone, 1997.

[9] Hirschman, Albert O. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press, 1970.

[10] Lessig, Lawrence. *Code and Other Laws of Cyberspace*. New York: Basic Books, 1999.

[11] Pettit, Philip. *Republicanism: A Theory of Freedom and Government*. Oxford: Oxford University Press, 1997.

[12] Srinivasan, Balaji S. *The Network State: How to Start a New Country*. Self-published, 2022.

[13] Thiel, Peter. "The Education of a Libertarian." *Cato Unbound*, April 13, 2009.

[14] Driscoll, Laura N., Noah L. Pettit, Matthias Minderer, Selmaan N. Chettih, and Christopher D. Harvey. "Dynamic Reorganization of Neuronal Activity Patterns in Parietal Cortex." *Cell* 170(5): 986–999, 2017.

[15] Elster, Jon. *Ulysses and the Sirens: Studies in Rationality and Irrationality*. Cambridge: Cambridge University Press, 1979.

[16] Fano, Robert M. *Transmission of Information: A Statistical Theory of Communications*. New York: MIT Press and John Wiley, 1961.

[17] Fitts, Paul M. “The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement.” *Journal of Experimental Psychology* 47(6): 381–391, 1954.

[18] Hick, William E. “On the Rate of Gain of Information.” *Quarterly Journal of Experimental Psychology* 4(1): 11–26, 1952.

[19] Hochberg, Leigh R., Mijail D. Serruya, Gerhard M. Friehs, Jon A. Mukand, Maryam Saleh, Abraham H. Caplan, Almut Branner, David Chen, Richard D. Penn, and John P. Donoghue. “Neuronal Ensemble Control of Prosthetic Devices by a Human with Tetraplegia.” *Nature* 442: 164–171, 2006.

[20] Ienca, Marcello, and Roberto Andorno. “Towards New Human Rights in the Age of Neuroscience and Neurotechnology.” *Life Sciences, Society and Policy* 13: 5, 2017.

[21] Jarosiewicz, Beata, Anish A. Sarma, Daniel Bacher, Nicolas Y. Masse, John D. Simeral, Brittany Sorice, Erin M. Oakley, Christine Blabe, Chethan Pandarinath, Vikash Gilja, Sydney S. Cash, Emad N. Eskandar, Gerhard Friehs, Jaimie M. Henderson, Krishna V. Shenoy, John P. Donoghue, and Leigh R. Hochberg. “Virtual Typing by People with Tetraplegia Using a Self-Calibrating Intracortical Brain-Computer Interface.” *Science Translational Medicine* 7(313): 313ra179, 2015.

[22] Lampson, Butler W. “Hints for Computer System Design.” *ACM SIGOPS Operating Systems Review* 17(5): 33–48, 1983.

[23] Libet, Benjamin, Curtis A. Gleason, Elwood W. Wright, and Dennis K. Pearl. “Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential).” *Brain* 106(3): 623–642, 1983.

[24] MacKenzie, I. Scott. “Fitts’ Law as a Research and Design Tool in Human-Computer Interaction.” *Human-Computer Interaction* 7(1): 91–139, 1992.

[25] Metzger, Sean L., Kaylo T. Littlejohn, Alexander B. Silva, David A. Moses, Margaret P. Seaton, Ran Wang, Maximilian E. Dougherty, Jessie R. Liu, Peter Wu, Michael A. Berger, Inga Zhuravleva, Adelyn Tu-Chan, Karunesh Ganguly, Gopala K. Anumanchipalli, and Edward F. Chang. “A High-Performance Neuroprosthesis for Speech Decoding and Avatar Control.” *Nature* 620: 1037–1046, 2023.

[26] Moses, David A., Sean L. Metzger, Jessie R. Liu, Gopala K. Anumanchipalli, Joseph G. Makin, Pengfei F. Sun, Josh Chartier, Maximilian E. Dougherty, Patricia M. Liu, Gary M. Abrams, Adelyn Tu-Chan, Karunesh Ganguly, and Edward F. Chang. “Neuroprosthesis for Decoding Speech in a Paralyzed Person with Anarthria.” *New England Journal of Medicine* 385(3): 217–227, 2021.

[27] Musk, Elon, and Neuralink. “An Integrated Brain-Machine Interface Platform with Thousands of Channels.” *Journal of Medical Internet Research* 21(10): e16194, 2019.

[28] Neuralink Corporation. Public Progress Updates on the Telepathy Clinical Trial. Company communications, 2024 to 2026.

[29] Norman, Donald A. *The Psychology of Everyday Things*. New York: Basic Books, 1988.

[30] Perge, János A., Mark L. Homer, Wasim Q. Malik, Sydney S. Cash, Emad Eskandar, Gerhard Friehs, John P. Donoghue, and Leigh R. Hochberg. "Intra-Day Signal Instabilities Affect Decoding Performance in an Intracortical Neural Interface System." *Journal of Neural Engineering* 10(3): 036004, 2013.

[31] Quine, Willard Van Orman. *Word and Object*. Cambridge, MA: MIT Press, 1960.

[32] Saltzer, Jerome H., David P. Reed, and David D. Clark. "End-to-End Arguments in System Design." *ACM Transactions on Computer Systems* 2(4): 277–288, 1984.

[33] Schmitt, Otto H. "A Thermionic Trigger." *Journal of Scientific Instruments* 15(1): 24–26, 1938.

[34] Schurger, Aaron, Jacobo D. Sitt, and Stanislas Dehaene. "An Accumulator Model for Spontaneous Neural Activity Prior to Self-Initiated Movement." *Proceedings of the National Academy of Sciences* 109(42): E2904–E2913, 2012.

[35] Shannon, Claude E. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27(3): 379–423 and 27(4): 623–656, 1948.

[36] Simon, Herbert A. *The Sciences of the Artificial*. 3rd ed. Cambridge, MA: MIT Press, 1996.

[37] Soon, Chun Siong, Marcel Brass, Hans-Jochen Heinze, and John-Dylan Haynes. "Unconscious Determinants of Free Decisions in the Human Brain." *Nature Neuroscience* 11(5): 543–545, 2008.

[38] Strickland, Eliza, and Mark Harris. "Their Bionic Eyes Are Now Obsolete and Unsupported." *IEEE Spectrum*, February 2022.

[39] Willett, Francis R., Donald T. Avansino, Leigh R. Hochberg, Jaimie M. Henderson, and Krishna V. Shenoy. "High-Performance Brain-to-Text Communication via Handwriting." *Nature* 593: 249–254, 2021.

[40] Willett, Francis R., Erin M. Kunz, Chaofei Fan, Donald T. Avansino, Guy H. Wilson, Eun Young Choi, Foram Kamdar, Matthew F. Glasser, Leigh R. Hochberg, Shaul Druckmann, Krishna V. Shenoy, and Jaimie M. Henderson. "A High-Performance Speech Neuroprosthesis." *Nature* 620: 1031–1036, 2023.

[41] Wolpaw, Jonathan R., Niels Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, and Theresa M. Vaughan. "Brain-Computer Interfaces for Communication and Control." *Clinical Neurophysiology* 113(6): 767–791, 2002.

[42] Yuste, Rafael, Sara Goering, Blaise Agüera y Arcas, Guoqiang Bi, Jose M. Carmena, Adrian Carter, Joseph J. Fins, Phoebe Friesen, Jack Gallant, Jane E. Huggins, and others. "Four Ethical Priorities for Neurotechnologies and AI." *Nature* 551: 159–163, 2017.

[43] Zeman, Adam, Michaela Dewar, and Sergio Della Sala. "Lives Without Imagery: Congenital Aphantasia." *Cortex* 73: 378–380, 2015.