

# **Sensorimotor Learning in Virtual Environments**

*Spencer Ryan Wilson*

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
**Doctor of Philosophy**  
of  
**University College London.**

Sainsbury Wellcome Centre for Neural Circuits and Behaviour  
University College London

February 23, 2024



I, Spencer Ryan Wilson, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.



# **Abstract**

My research is about stuff.

It begins with a study of some stuff, and then some other stuff and things.

There is a 300-word limit on your abstract.



# Impact Statement

UCL theses now have to include an impact statement. (*I think for REF reasons?*) The following text is the description from the guide linked from the formatting and submission website of what that involves. (Link to the guide: <http://www.grad.ucl.ac.uk/essinfo/docs/Impact-Statement-Guidance-Notes-for-Research-Students-and-Supervisors.pdf>)

The statement should describe, in no more than 500 words, how the expertise, knowledge, analysis, discovery or insight presented in your thesis could be put to a beneficial use. Consider benefits both inside and outside academia and the ways in which these benefits could be brought about.

The benefits inside academia could be to the discipline and future scholarship, research methods or methodology, the curriculum; they might be within your research area and potentially within other research areas.

The benefits outside academia could occur to commercial activity, social enterprise, professional practice, clinical use, public health, public policy design, public service delivery, laws, public discourse, culture, the quality of the environment or quality of life.

The impact could occur locally, regionally, nationally or internationally, to individuals, communities or organisations and could be immediate or occur incrementally, in the context of a broader field of research, over many years, decades or longer.

Impact could be brought about through disseminating outputs (either in scholarly journals or elsewhere such as specialist or mainstream media), education, public engagement, translational research, commercial and social enterprise activity, engaging with public policy makers and public service delivery practitioners, influencing ministers, collaborating with academics and non-academics etc.

Further information including a searchable list of hundreds of examples of UCL impact outside of academia please see [https://www.  
ucl.ac.uk/impact/](https://www.ucl.ac.uk/impact/). For thousands more examples, please see [http:  
//results.ref.ac.uk/Results>SelectUoa](http://results.ref.ac.uk/Results>SelectUoa).

# **Acknowledgements**

Acknowledge all the things!



# Contents

<b>1</b>	<b>Introduction &amp; Aims</b>	<b>17</b>
<b>2</b>	<b>Theoretical Lit Review</b>	<b>29</b>
2.1	Minimum Intervention Principle . . . . .	29
2.2	Model-based vs. Model-free . . . . .	29
2.2.1	Existing Models of Motor Control and Adaptation . . . . .	32
2.2.2	Adaptive Linear Quadratic Control . . . . .	33
2.2.3	Motor Adaptation . . . . .	33
2.2.4	Two-rate models . . . . .	33
2.3	The Heterarchical Motor System . . . . .	33
2.4	Motor Units to Muscles . . . . .	34
2.4.1	Coordinative Structures . . . . .	36
2.4.2	Fractionating Structures . . . . .	37
2.4.3	Supraspinal Motor Maps . . . . .	38
2.4.4	John Rothwell and Jens Bo Nielson: Voluntary Control . . . . .	43
2.5	Mathematical Theories of Motor Control and Learning . . . . .	44
<b>3</b>	<b>Experimental Literature Review</b>	<b>47</b>
3.0.1	Control . . . . .	47
<b>4</b>	<b>Experimental Methods</b>	<b>67</b>
4.1	Methodological Aims . . . . .	69
4.2	Recording Setup . . . . .	70
4.3	Raw Data . . . . .	70
4.4	Natural Movement Task . . . . .	70
4.5	Calibration Task . . . . .	75
4.5.1	Does the Calibration Task Encourage Exploration? . . . . .	76
4.6	Decoder Fitting . . . . .	76

4.7	Target Task . . . . .	80
4.8	Data Validity . . . . .	83
4.9	Activity Filter . . . . .	84
<b>5</b>	<b>Basic Results</b>	<b>89</b>
5.1	Does Subject Performance Improve over Trials? . . . . .	91
5.1.1	Hits over Targets . . . . .	97
5.2	Do Subject Attributes Predict Performance? . . . . .	98
5.3	Does the Decoder Predict Performance? . . . . .	101
5.4	Does Task Space Variance Decrease over Trials? . . . . .	104
<b>6</b>	<b>More Results</b>	<b>113</b>
6.1	Structure of the Data Manifold . . . . .	115
6.1.1	Is EMG Lognormal? . . . . .	115
6.2	How Does PCA Dimensionality Compare Across Tasks? . . . . .	120
6.2.1	UMAP . . . . .	121
6.2.2	Toy GMM model demonstration X-pattern of trajectories . .	121
6.3	Subspace Confinement . . . . .	122
6.4	Gaussian Mixtures . . . . .	125
6.4.1	GMMs vs PCA . . . . .	125
6.4.2	Example GMMs . . . . .	127
6.4.3	GMM Pseudo-derivative . . . . .	130
6.4.4	GMM Effective Rank . . . . .	131
6.4.5	GMM Entropy . . . . .	133
6.4.6	Log Transforming the Data . . . . .	134
6.4.7	Influence of Natural Repertoire . . . . .	134
<b>7</b>	<b>CHAPTER NAME</b>	<b>137</b>
7.1	GMM Null Space . . . . .	139
7.2	Nullspace Ratio of Task Error . . . . .	141
7.3	Nullspace Optimization Solutions . . . . .	143
7.4	Classifying Nullspace Noise . . . . .	146

<b>8 Conclusions &amp; Next Steps</b>	<b>151</b>
8.1 Conclusions . . . . .	153
8.2 Next Steps . . . . .	153
8.2.1 EMG Hardware . . . . .	153
8.2.2 Eye Tracking . . . . .	153
8.2.3 Preprocessing . . . . .	154
8.2.4 Calibration . . . . .	154
8.2.5 Task Design and Data Collection . . . . .	155
8.2.6 Modeling Control . . . . .	155
8.3 Modeling Learning . . . . .	157
8.4 Open Questions . . . . .	158
8.5 Model-based optimal control & LQR . . . . .	159
8.6 Inverse Optimal Control . . . . .	159
<b>Bibliography</b>	<b>161</b>



# List of Figures

1.1	Voder	20
1.2	Basmajian experimental method	23
2.1	Low-variance PCs contain task-relevant information	35
2.2	Comparing the work of Strick and Graziano	40
2.3	Motor system schematic	42
4.1	Experimental setup prototype	71
4.2	Experimental setup	72
4.3	Electrode layout with forearm and hand muscles	73
4.4	Example raw EMG data	74
4.5	Movement task visual feedback	74
4.6	Calibration task visual feedback	76
4.7	Example subject decoder	80
4.8	Center-hold, Reach-out task visual feedback	81
4.9	Data filtering steps	81
4.10	Point mass position trajectories in two-dimensional task space during the center-hold, reach-out task with 32 targets spaced evenly around the unit circle (shown with red borders). Color corresponds to target numbers, with target zero located at (0,1). Target order was randomized. Training was conducted over 3 blocks each with 32 trials, 1 trial per target.	82
5.1	Hit counts over trials	92
5.2	Reach time over trials	93
5.3	Trajectory length over trials	93
5.4	Trajectory segments over trials	94
5.5	Trajectory segment definition	95
5.6	Hits versus reward	96

5.7	Reward versus learning rate . . . . .	96
5.8	Hit and reward histograms . . . . .	97
5.9	Hits over targets . . . . .	97
5.10	Target hit Tukey test . . . . .	98
5.11	Arm versus reward . . . . .	99
5.12	Experiment day versus reward . . . . .	99
5.13	Hour of day versus reward . . . . .	100
5.14	Hours of sleep vs reward . . . . .	100
5.15	Subject group comparisons . . . . .	101
5.16	Example decoder directions in 2D. . . . .	102
5.17	Decoder normality testing . . . . .	103
5.18	Decoder axis similarity versus reward . . . . .	104
5.19	Mean trajectories . . . . .	106
5.20	Rotating trajectories . . . . .	107
5.21	Trajectory variance over blocks . . . . .	108
5.22	Trajectory variance statistics . . . . .	109
6.1	Pairplot of calibration data . . . . .	115
6.2	Pairplot of log-transformed calibration data . . . . .	116
6.3	Explained variance over tasks . . . . .	120
6.4	Explained variance Tukey test . . . . .	121
6.5	Explanatory Mixture Model . . . . .	122
6.6	GMMs and PCA Covariance . . . . .	125
6.7	Subject 1 trial GMMs overlayed on trajectories . . . . .	127
6.8	Subject 35 trial GMMs overlayed on trajectories . . . . .	128
6.9	Example prior GMMs . . . . .	129
6.10	Example GMMs for subjects 1 and 14 . . . . .	129
6.11	Example trial GMMs for subjects 37 and 42 . . . . .	130
6.12	Mean GMM differences over subjects . . . . .	130
6.13	Wasserstein distance between GMMs . . . . .	131
6.14	Effective rank of GMMs . . . . .	131

6.15	Effective rank . . . . .	132
6.16	Entropy of subject GMMs . . . . .	133
6.17	Successive GMM entropy differences . . . . .	133
7.1	Nullspace activity for GMMs . . . . .	139
7.2	Example Error . . . . .	141
7.3	Error Nullspace . . . . .	142
7.4	Computed solutions . . . . .	144
7.5	Distance from computed solutions . . . . .	145
7.6	regression scores . . . . .	146
7.7	classification scores . . . . .	146
7.8	pca 2d full . . . . .	147
7.9	pca 2d null . . . . .	147



# List of Tables

4.1	List of movement commands . . . . .	75
5.1	Statistics of performance regression . . . . .	97



## **Chapter 1**

# **Introduction & Aims**

*Movement is nothing but the quality of our being.*

— Sunryu Suzuki, *Zen Mind, Beginner's Mind*

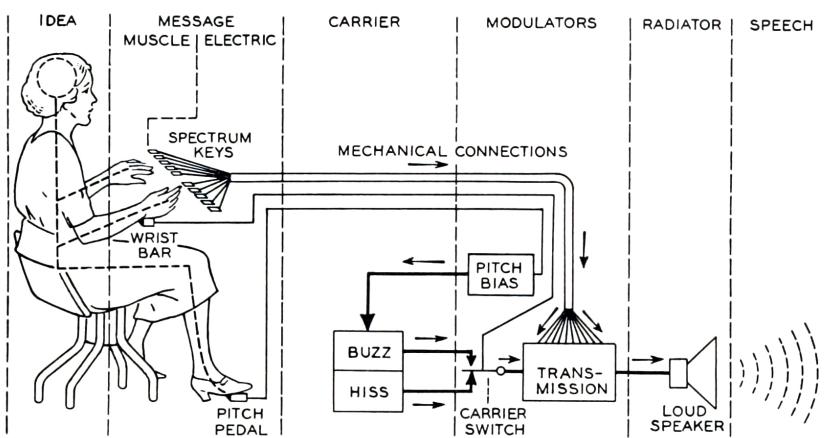


- Why do this? Why fund or support this? Broadly, why is this topic interesting?
- isn't motor just spinal modules being played like keys?
- Why motor? Why be interested in this?
- motor actions are complex and extended across time and space
- no, it's all about the dynamics— the rich correlations between movement variables, prediction, feedback...
- What are our goals? What are we trying to achieve? Why are we setting up this task?
- Voder
- Basmajian
- RL muscle work
- Bernstein
- 

In 1939, a Bell Labs researcher, Homer Dudley, engineered a demonstration device he named the “Voice Operation Demonstrator” or “Voder” for the New York World’s Fair[?]. Dudley explains:

In the Fair demonstration, . . . [the operator replies to an announcer’s] questions by forming sounds on the voder and connecting them into words and sentences. She does this by manipulating fourteen keys with her fingers, a bar with her left wrist and a pedal with her right foot. **This requires considerable skill by the operator.**

While Dudley’s work served as an engineering investigation to reproduce speech sounds mechanically in line with Bell Telephone’s goals of transmitting such sounds along their wires, he effectively invented a speech prosthetic controlled by the body



(a)



(b)

Figure 1.1: Longer voder caption

rather than the muscles of the tongue and larynx<sup>1</sup>. By reconfiguring the inputs from the muscles usually used for speech production to the many muscles controlling the hand, arm, and feet of the Voder’s operator, Dudley created an experimental environment for the operator, now subject, to solve a redundant motor task. To achieve success in their task, the operator must develop “considerable skill” as Dudley points out. It is noted elsewhere that there were only a few Voder operators, and they each trained for dozens of hours to reproduce their answers in the form of intelligible speech using the device.

That famous quote from Feynman’s blackboard: “What I cannot create, I do not understand”.

The experimental environment that Dudley created in 1939 is exactly the kind of study we attempt to create in this project. The chief purpose of this thesis is *to construct an experimental platform on which a suite of motor skill learning experiments can be devised and recorded*. By developing and sharing this platform, we hope to advance our understanding of trial-to-trial motor skill learning through the combination of theory and experiment, beginning a process of data generation to yield insights into how human beings are able to solve motor reconfiguration problems in the spirit of Dudley’s Voder.

The long term goal of the research direction suggested here is to develop tasks which ask subjects to produce a variety of movements in response to a variety of goals and perturbations. This will allow us to study the computations that humans use in everyday tasks to solve the motor problems they face. This stands in contrast to many repetitions of the same movements. However, we wish to validate our experimental setup on classical tasks as a stepping stone to tasks with greater variety. The dream is to have a task where:

Todorov: we know how to design and interpret experiments that involve many repetitions of the same movement however there is limited role for online optimization in that context. instead **we need experiments where subjects are required to come up with new movements all the time. how can we get experimenters to do such experiments? show cool movies of robots doing cool things, and hopefully**

**get the experimenters excited.”** (todorov online optimization slides)

We take considerable inspiration from work like Basmajian in the 1960s, where basic research on the limits of human motor control were being done, attempting to understand if humans could learn, with feedback, to fire single motor unit action potentials. Variability in performance between subjects varied widely, with some subjects being unable to perform at all, while others able to fire several motor units alone and in concert.

In the 1960s and 70s, there was a surge of interest in this quantum of behavior. In their 1962 experiment, Harrison and Mortenson provide real-time auditory and visual feedback to several subjects attempting SMU activation in the *tibialis anterior* muscle using EMG signals via surface and needle electrodes. They claim that auditory feedback is more helpful for learning compared to visual feedback and that some form of “external” feedback is essential. They report a great variability in response with respect to ability: “a considerable degree of ‘mental concentration’ is needed for independent contractions” [?]. They write that “As many as six individual motor units were recruited and identified in some experiments” though this was an extreme level of competency among their subjects.

Over the last few decades, there has been considerable amount of work done to untangle the abilities of the motor system to flexibly control the body including through optimal control theory[?], reinforcement learning in continuous action spaces[1], and detailed physiological studies[2]. Despite a wealth of painstaking and impressive work, a holistic understanding of the computations underlying the construction of skilled movement remains an incredibly exciting and fruitful direction of research. Our aim is to progress understanding of skilled movement by studying the solutions produced by human subjects to motor tasks in dynamically rich, yet experimentally manipulable, virtual environments. Our goal is to reverse-engineer the ability to acquire and perform novel motor skills.

Humans produce a great variety of movements every day, often without conscious thought. For example, movements like bringing a cup of coffee to our lips for a sip are generally out of reach for state-of-the-art robotic systems. We claim that

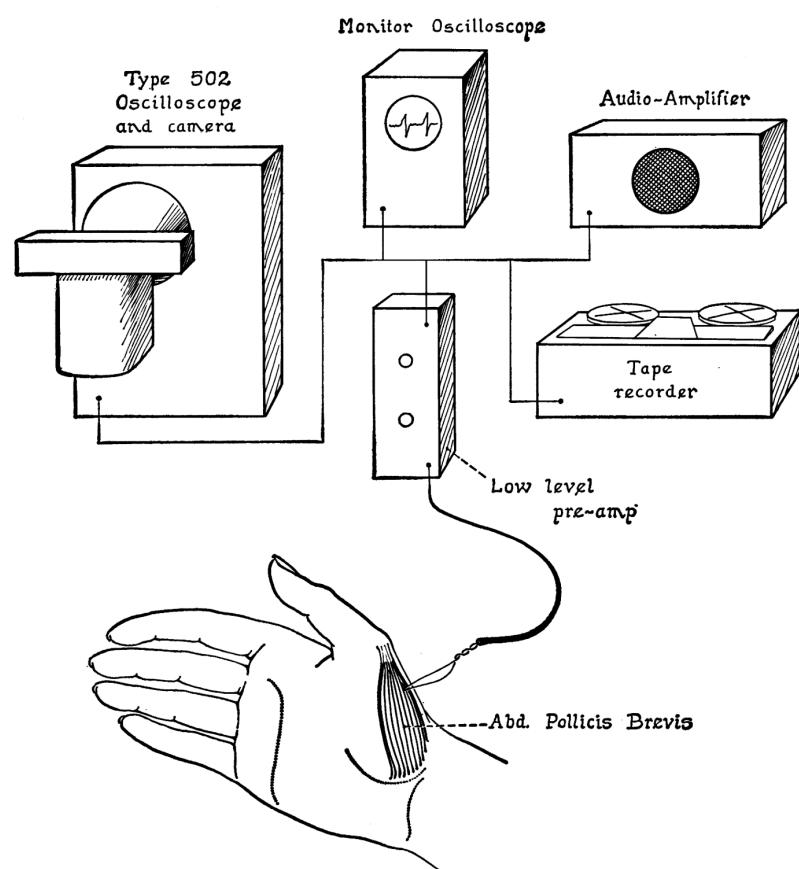


Fig. 2. Technique of recording from abductor pollicis brevis.

2 AUGUST 1963

Figure 1.2: CAPTION

this “motor gap” between biological and artificial motor systems is due to a lack of *dexterity*. Soviet neuroscientist Nikolai Bernstein defined dexterity as the ability to “find a motor solution in any situation and in any condition[3]”. The crux of this definition is the flexibility of such solutions. This flexibility, or robustness<sup>2</sup>[4], is the ability to optimize internal parameters in response to external perturbations and adapt to new information to achieve the goals of an ongoing plan.

Using data from our experimental setup, we wish to understand both how the structure of muscle activation variability evolves during skill acquisition and how the motor system constructs skilled movement through the composition of component muscle activations. To begin, we review a sampling of current motor physiology research relevant to dexterous motor computations in ???. In ??, we cover our prototype hardware and experiments. With inspiration from physiology and our experiments, we hope to make progress in modeling sensorimotor control and learning in our experimental setup. We cover preliminary work in this direction in ??, and discuss possible future directions in Section 8.2.

Why can’t robots move like humans? What is special about human movement?  
Why is movement a hard problem? What is hard about it?

Hans Moravec’s eponymous paradox states that it is easier to generate artificially intelligent performance on tasks we think of as intellectually challenging, such as chess, than to provide a machine with faculties we take for granted, such as movement. Moravec’s Paradox, for example, encourages us not to look past the complex computations generated by the human motor system. Following Moravec, this work focuses on what is arguably the most advanced control apparatus in the known universe: the human movement machine.

The interesting problem here is coordination of a redundant system to produce dexterous solutions— we want to solve the redundancy problem and produce solutions that are robust to external perturbations and sensitive to new information. While a robot may be able to move a cup of coffee to a precise location in space, its solution is often found to be brittle in a new context, or unable to generalize to the movement of new objects. We define a skill as a behavior that involves dexterity in

Bernstein's sense. The use of a tool such as a screwdriver is an example of a motor skill. We define a task as the production of skilled movement in a particular context. Driving a screw in a particular posture using a particular screwdriver is an example of a task.

“Why” is a question for evolutionary biologists

“How” is a question for engineers

Humans have extraordinary motoric ability, we want to understand why. They display traits like Robustness, Flexibility, Generalization, Composition

Why is this the most interesting problem? Because we are so far from creating machines that can move like we do, which means we are very far from understanding what elements of the body and brain are sufficient or necessary for human-level motor control.

The physiology of the human hand is unique from an evolutionary perspective, so we focus on use of the hand.

We want to set up a specific experiment will help us to track muscle-level changes

We need to get a little messy, collect and analyze data in a decidedly exploratory manner to then inspire hypotheses that can be modeled. We want these hypotheses to be inspired by theoretical work in control and learning theory.

To do so, we will with the conceptual language of the experimental sensorimotor learning community together with the language of the control theory and reinforcement learning community, as each of these communities shares a common goal of understanding the computation underlying the production of skilled movement. Making a step towards building a bridge between these communities.

Call to action:

The processes by which biological control solutions spanning large and continuous state spaces are constructed remain relatively unexplored. Future investigations may need to embed rich dynamical interactions between object dynamics and task goals in novel and complex movements [@McNamee2019].

To explore dexterous movement, we will leverage recordings of muscles controlling the hand as a readout of flexible motor behavior. This is a step beyond recording hand kinematics as electromyography provides a physiological output of the nervous system. Surface electromyography recordings taken from the forearms controlling subjects' dominant hands allows us to track the sequential selection of muscle activations during both skill acquisition and subsequent performance of that skill to achieve desired goal. As we are interested in subjects' abilities to acquire new skills, we design tasks that require subjects to use available, but uncommon, motor activations. We then track the selection and execution of these activation during virtual tasks. Preliminary work in this direction is described in @sec:experiment.

The overarching goal of this task is to track learning in a new movement contingency. Subjects have no prior knowledge related to the task, and must explore to find solutions, while experiencing constraints of their fitted decoder/task mapping. We want to follow their learning progress statistically to gain an insight into how subjects are learning to deal with their new environment.

# Notes

1. For the armchair historians, see Wolfgang von Kempelen's (inventor of the "mechanical Turk" chess-playing automaton) "speaking machine" from 1780 as an earlier example of such experiments. Note that Von Kempelen spent twenty-odd years developing his speaking machine, whereas this thesis only required a mere six.
2. Kitano defines robustness as "the maintenance of specific functionalities of the system against perturbations, and it often requires the system to change its mode of operation in a flexible way". He claims that robustness requires control, alternative mechanisms, modularity and decoupling between high and low level variability.



## Chapter 2

# Theoretical Lit Review

*Even a simple movement is a global body event.* — Bizzi & Ajemian, 2020

*An interesting open question is how to relate trial-to-trial dynamics of learning to asymptotic predictions regarding optimal adaptation.* — Todorov, 2007

## 2.1 Minimum Intervention Principle

1. System Identification: learning a transition function  $p(y_t|x_t, u_t)$  - How do you learn the unknown observation model from data?
  1. Policy Optimization - Once dynamics are learned (or at least stable?), how do we form a policy that is generalizable to new tasks under these dynamics? - This is the control problem.

It's safe to assume that these processes are happening in parallel. Because we have complete and arbitrary control over the observation mapping, we can ask the subject to interact through a dynamic that is intuitive (informative prior) or unintuitive (uninformative or inhibitive prior). Each scenario, we hypothesize, will elicit different strategies for learning and control.

## 2.2 Model-based vs. Model-free

think there are a number of unique things to motor control: - We can use this framework to distinguish between skill learning (increasing policy complexity) and adaptation (no change in policy complexity). - In sensorimotor control, 'compression' may be more about being sloppy on the sensory end of things, rather than having a compact policy to store and retrieve. - We can speculate that practice and expertise

as convincing your brain to commit to a more complex policy (or maintain commitment to a current level of complexity). - Certain movement disorders could be viewed as involving information bottlenecks

default policy as habits [@pirayLinearReinforcementLearning2019]

info bottleneck / policy compression[@gershmanRewardcomplexityTradeoffSchizophrenia2020]

”optimal compression means knowing the state probabilities” having some form of a model – marginal state distribution (mean state occupancies?) SR row sums?

”the memory demand of policies acts as an information bottleneck in action selection”

↳ The hierarchical organization typical of earlier sensory areas is not adhered to everywhere. On the contrary, the anatomy of associative areas and prefrontal cortex suggests a more ”democratic” organization, and processing appears to take place in webs of strongly interacting networks (8). Decisions to act and the execution of plans and choices could be the outcome of a system with distributed control rather than a single control center. Coming to grips with systems having distributed control will require both new experimental techniques and new conceptual advances. Perhaps more appropriate metaphors for this type of processing will emerge from studying models of interacting networks of neurons. [@sejnowskiPerspectivesCognitiveNeuroscience1988]

From the physiology, we see that the motor system is highly distributed and constructs action based on a variety of state dependence. The theoretical question becomes *when does it make computational sense to construct action by composing control policies rather than selecting or tuning a single policy?* When is policy arbitration computationally advantageous?

Human movement is ultimately the result of the activation and contraction of muscle fibers, and movements lie on a spectrum between reflexive and volitional. The supramuscular circuitry which determines the degree of volition we ascribe to movement, where volitional movement relies on supraspinal (though not necessarily conscious) processes. The human hand is a unique evolutionary invention that

underlies our ability to perform various skills in a range of tasks— movements that are decidedly volitional. It could be argued that the hand is in fact a crucial aspect of humanness. It is thought that the human cerebellar and neocortices evolved reciprocally to expand and support the computational burden of increasingly complex motor tasks such as tool-making and language production[REF?]. The hand is the pinnacle of dexterity and, as such, it is a fruitful testbed for studying the computations and circuitry that drive dexterous movement. A detailed physiological review of the hand and its relation to skilled movement is described in ??.

Our theory of neural control of the hand is approximately: control is composed of a number of overlapping cortical controllers— these receive input from goal-oriented centers as well as a plethora of ongoing contextual, perceptual information. Control is thus modulated by these inputs, adjusting “online” to disturbance. Cortical controllers are massively redundant; they contain all available information about the context of an ongoing task, branching to an array of downstream spinal centers as well as converging to individual spinal innervations. Our hypothesis is that subjects will use their vast repertoire of pre-existing control schemes/movements/controllers/patterns/activations until they find a pattern that increases their success, upon which they will “hone” this scheme by refining the discovered movement. This hypothesis predicts an exploratory, or “search”, period of the task, followed by (or overlapping with) an exploitative or “honing” period as subjects settle on a motor solution. Our work is to highlight the statistical differences and attributes of these two sub-activities in our task, and explain how these activities relate to subjects’ natural hand movements and to theoretically optimal learning dynamics.

**The motor learning field does not yet possess an adequate computational model for practice-induced increases in motor acuity.** (Krakauer Motor Learning 2019)

## 2.2.1 Existing Models of Motor Control and Adaptation

### Optimal Feedback Control

Stephen Scott review – <https://www.nature.com/articles/nrn1427> look at the bullet points there, relate to our experiment

OFC is the best we got for motor coordination, but there's no adaptation or learning

The control setup writes a cost, environment has some dynamics.

What is changing in this scenario? What is being learned? What information is used to do this learning?

Which model variables correspond to muscles? Movements? What does the resultant feedback controller compute? How does this relate to cognition?

This model is lacking in . . .

A key paper is Valero-Cuevas 2009 which recording EMG from the seven muscles driving the finger in a force-feedback task. The authors found support for the “minimum intervention principle” [@Valero-Cuevas2009].

nonlinear iLQG models

Noise in OFC

Intuitive Example of the OFC framework

Error-based Adaptation

## **2.2.2 Adaptive Linear Quadratic Control**

### **2.2.3 Motor Adaptation**

State-space Models of Motor Adaptation

Example Models

Critique

### **2.2.4 Two-rate models**

## **2.3 The Heterarchical Motor System**

Motor circuits that support online control are highly distributed, including many cortical and subcortical regions. Many different regions likely contribute to each process, and a single brain region may participate in multiple processes.<sup>38</sup> Thus, it is highly unlikely that a single brain region is exclusively responsible for a single parameter or control process. For example, the control policy in which motor commands are regenerated based on the present state of the body likely involves cortical, brainstem, and spinal processing, because each level contributes to a certain degree to the final pattern of muscle activity during a motor action. The final motor command only emerges at the motoneuronal level (the final common path)<sup>51</sup> as some descending projections synapse directly onto motor-neurons.<sup>52</sup> This distributed and hierarchical organization means that the contribution at the highest cortical level will not simply reflect the pattern of muscle activity for a motor action. Thus, temporary deactivation in any of these brain regions involved in feedback control could lead to impairments that look like a reduction in control policy gain.

## 2.4 Motor Units to Muscles

The quantum of motor output is the motor unit (MU), defined as a single motoneuron axon and the set of junctions the terminals of its axon branches form with one or more muscle fibers. The MU provides the motor system with spatial redundancy at the muscle level: multiple muscle fibers contract due to a single alpha motoneuron (AMN) spike in the spinal cord's ventral horn, and multiple AMNs may overlap in their innervations. The forces produced by motor units span several orders of magnitude, though most units produce very small forces. Here we find temporal redundancy: in order to produce movements, MUs combine to generate a range of forces[@fuglevandMechanicalPropertiesNeural2011]. Since the innervation ratios of muscles in the forearm and hand are relatively small compared to more proximal muscles (which contain thousands of MUs), the logarithmic recruitment and redundancy of motor units enables the hand to produce movements with very fine spatiotemporal resolution.

Muscle fibers are contained within muscle compartments, and each muscle may have one or more compartments. The fingers of the hand are extended by the extensor digitorum (ED) which contains four compartments, one for each of the tendons the muscle produces. Each tendon connects to the three metaphalangeal joints of each digit. The fingers are flexed by two muscles, the flexor digitorum superficialis (FDS) and the flexor digitorum profundus (FDP). Like the ED, these muscles produce four tendons, one to each finger from each of their four compartments. As such, one must coactivate these agonist and antagonist muscles in order to extend or flex a single finger in isolation[@fuglevandMechanicalPropertiesNeural2011]. Adduction and abduction of the fingers is produced by the 19 intrinsic muscles of the hand, each of which has their origin and insertion points within the hand itself[@vanduinenConstraintsControlHuman2011]. The intrinsic muscle tendons form a kind of network around each of the digits. The human hand, thumb, and forearm system contains more than 30 muscles and at least 20 degrees of freedom are theoretically available for actuation. However, due to biomechanical coupling, the effective degrees of freedom is presumably less than 20.

This structure exists in order to facilitate the acquisition of new skills and the generalization of existing skills to new contexts. While the anatomy of the hand and forearm presents constraints on movement, the system remains capable of producing an incredible variety of movement patterns[@yanUnexpectedComplexityEveryday2020;@Basmaj]. The structure of the neuromuscular system that underlies this variety offers many clues as to the relevant computations required for dexterous movement. In {+@fig:low\_variance\_PCs}, Yan et al. show how even low-variance principle components of joint kinematics during object grasping and ASL signing display correlational structure and not merely noise. That is, the production of hand movement is highly task-specific, where individual tasks are linked to bespoke muscle activations patterns.

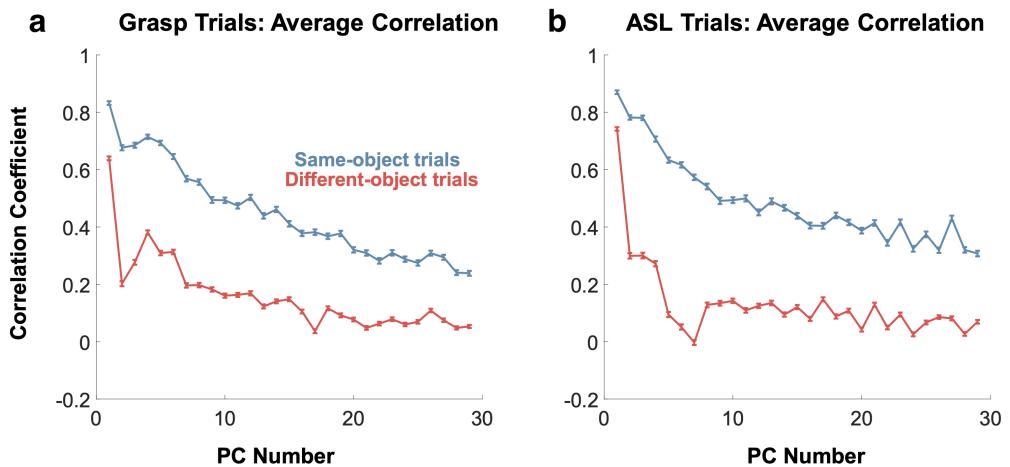


Figure 2.1: Taken from Yan et al. 2020. Plots show mean correlations between hand joint kinematic trajectories during grasp trials with the same (blue) and different (red) objects (a) and ASL signs (b) projected onto the same principle components. Correlations are averaged across 8 subjects. Within-object and within-sign correlations are systematically higher than their shuffled counterparts. Error bars denote SEM. This data supports the idea that low-variance components of kinematics data contain task-specific structure rather than merely reflecting noise. This is encouraging for our experiments, which hope to extend this idea into careful analyses of task specific features of EMG data across learning and in response to perturbations.

### 2.4.1 Coordinative Structures

Many studies have contributed to the concept of synergies as a hard-wired organizing feature of the motor system[@mussa-ivaldiMotorLearningCombination2000,@DAvella2003]. However, these works tend to extrapolate from non-primate preparations, particularly in the frog, and use tasks which are inherently low-dimensional to explain covariance structure in primate and human kinematic and electromyography data[@giszterMotorPrimitivesNew2015;@gao2017]. That said, it would be foolish to deny the existence of synergistic muscle coactivation even at the structural level. Careful studies of force control by the fingertips present a complex story of dimensionality of control in this regime[@raczSpatiotemporalAnalysisReveals2013]. Constraints exist in the architecture of the hand as well as its control system, though we maintain that concept of synergies, especially in the context of dexterous movement, is often presented as an oversimplification rather than a mere simplification. We believe the story of the hand is more complex.

Studies have attempted to quantify the number of effective degrees of freedom of the hand with various methods. This has primarily been taken to be the number of linear features which contain a desired level of the original signal variance, where the signal is the joint angles of the hand engaged in various behaviors[@Ingram2009;@TodorovDimensionality2005]. These methods have resulted in roughly 8 linear features of hand kinematics to solve a variety of tasks, with subtleties found in inter-task and inter-subject variations. Note that the motor repertoire is hardly high-dimensional when compared to the dimensionality of the visual feature extraction system[@yanUnexpectedComplexityEveryday2020]. A recent study found that low-variance linear, kinematic components displayed significantly higher correlation within condition (e.g. grasp of a specific object) than across condition. This suggests that these components carry task-dependent information rather than condition-independent, task-irrelevant noise[@yanUnexpectedComplexityEveryday2020]. This suggests that the control of the hand is more nuanced than a set of fixed synergies.

What Bizzi and colleagues call “the problem of supraspinal pattern

formation”—how synergies are activated through time— we argue, in the context of hand control, is not simplified by the existence of hard-wired or soft-wired synergies[@bizziMotorPlanningExecution2020]. Rather, the CNS produces control signals in a range of contexts and in response to continually changing task demands. Rather than the CNS “simplifying movement” through synergetic action, it is more likely that hand synergies fall out of a optimization strategy which trades off effort and accuracy where effort may, in part, correspond to independent control of individual control dimensions. In this view, synergies, hard-wired or not, reflect the statistics of the environment in which movement is constructed[@brutonSynergiesCoordinationComprehensive2018]. If we limit ourselves to synergetic control, then we have simply passed the problem to a lower-dimensional one of the same fundamental nature. Neural control of the hand likely contains a spectrum of modularity in order to maintain its role as a flexible instrument. Synergetic action is one end of this spectrum resulting from the computations inherent to, along with the structures of the human movement machine.

## 2.4.2 Fractionating Structures

Just as many muscle fibers may be innervated by a single AMN, up to thousands of neurons contact single AMNs through monosynaptic corticospinal, or corticomotoneuronal (CM), connections and other descending pathways through elaborate spinal circuitry. The hallmark of CM connections in particular is their influence over multiple muscle compartments as well as multiple muscles, though typically agonist or antagonist sets[@cheneyFunctionalClassesPrimate1980]. This may seem counter-intuitive as a means to produce individuated movement, but experimental evidence in primates has shown that the convergence of many CM collateral fibers onto single AMNs driving the distal muscles in particular can produce a fine grading of activity over motor units driving the distal joints. CM cells also appear to play a role in the inhibition of antagonist muscles prior to contractions required for movement [@griffinMotorCortexUses2020]. These findings confirm theories about the excitatory and inhibitory role of these connections dating back decades, and combine to suggest that variables encoded in cortical ensembles are more complex

than kinematics or dynamics alone[@cheneyFunctionalClassesPrimate1980].

The CM tract thus acts in coordination with synergistic muscle activations of the hand to achieve control that is balanced between modularity and flexibility. Findings suggest that there is a bipartite structure in human motor cortex driving dexterous control of the distal part of the upper limb which, it has been suggested, evolved under pressure to quickly generalize between tasks. This work argues that these two streams of hand control, namely “fractionated” and “synergistic” control, may interact to produce versatility, and balancing these subsystems may be a key part of the optimization function when learning new skills[@Rathelot2009;@griffinCorticotoneronalCellsAre2015;@Takei2017]. This dualism is likely not rigidly dichotomous, but rather a spectrum of overriding fractionation (so-called “New M1”) atop a phylogenetically older system of synergistic action[@dumCorticospinalSystemStructural2011]. Griffin and colleagues found that CM cells are functionally tuned to a muscle’s mode of activity (agonist, antagonist, fixator) to “bypass spinal cord mechanisms and sculpt novel patterns of motor output that are essential for highly skilled movements”[@griffinCorticotoneronalCellsAre2015]. The hypothesis stemming from the previously described work is that CM connections override the “consolidated” patterns putatively generated via spinal interneuron circuitry. The setup devised in our work aims to measure fractionation by tracing motor unit correlations across learning. Whether fractionation in our experiments is due to the CM pathway can only be speculation, but our work may provide direction for future studies pairing intracortical recordings with careful electromyography.

### 2.4.3 Supraspinal Motor Maps

It is known from recent work that primary motor cortex (M1) is not an isolated movement-generating dynamical system, but rather a node in the network of a feedback-modulated, distributed movement machine[@sauerbreiCorticalPatternGeneration2019]. Thinking of the structural architecture of M1 as an input-driven system with outputs along a spectrum of modularity from synergistic to fractionated, we can ask what kind of functional architecture might have evolved in the neuromuscu-

lar controller? Graziano and colleagues found that 500ms electrical stimulation to M1 reliably produced stereotyped movements in primates[@graziano2006]. These movements appeared to produce goal-oriented actions pulled out of other contexts such as bringing food to the mouth, and seemed to be arranged on the cortical sheet topographically in terms of spatial endpoints rather than as a humunculus. Graziano refers to this as the cortical “action map”, that these stimulations tapped into the control mechanisms of the primate’s motor system[@grazianoIntelligentMovementMachine2009]. These results has recently been confirmed by optogenetics work in marmosets and macaques [@ebina2019;@watanabeForelimbMovementsEvoked2020].

The motor map concept suggests interpreting activity in M1 as a field of feedback control microcircuits, integrating and transforming inputs, both internal and external, to sculpt ongoing movement[@wiltschkoMappingSubSecondStructure2015]. This is in accordance with the idea that there is a structural hierarchy in M1 covering a spectrum of movement modularity. These ideas together form a picture of the motor system as a structural scaffold upon which behaviorally relevant feedback mappings from cortex to the spinal cord are continuously activated and modulated based on information and estimates about the periphery. In this view, the encoded variables of interest depend on the goals, context, and perturbations of the intended movement. {+@fig:strick\_graziano} shows Graziano et al.’s stimulation results, what might be termed a functional view of the cortical motor system, next Strick er al.’s described above clarifying the structural view of modularity in this system.

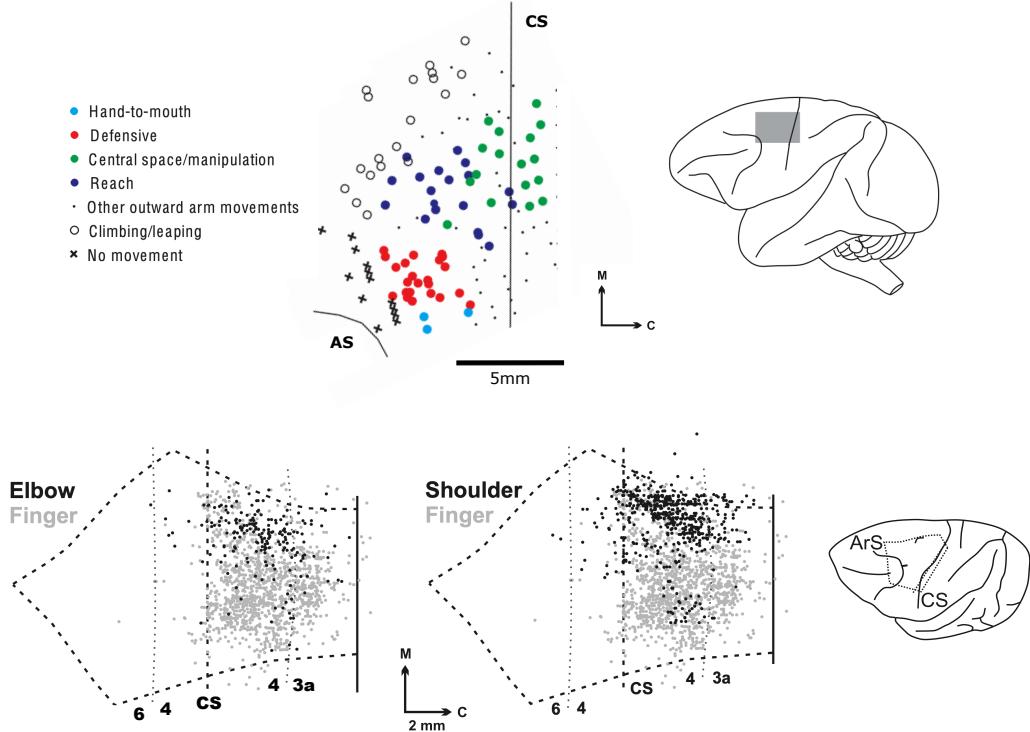


Figure 2.2: Similarities between electrical stimulation on behavioral timescales and rabies tracing identification of CM cells. CM cells are largely confined to the caudal half of M1, while this region tends to evoke complex manipulatory movements when electrically stimulated. (Top Left) Corticomotoneuronal (CM) cells traced using rabies from muscles of the elbow and finger. (Top Right) CM cells traced using rabies from muscles of the shoulder and finger. (Bottom) Complex movements evoked by 500ms electrical stimulation pulse trains. Adapted from Graziano 2005 and Rathelot et al. 2009[@graziano2005; @Rathelot2009].

Graziano writes:

“The usefulness of a feedback-dependent mapping from cortex to muscles is that it can in principle allow neurons in motor cortex to control a diversity of movement variables, such as direction, speed, hand position, or posture that transcend a fixed pattern of muscle activation. If the network receives feedback information about a specific movement variable, then it can learn to control that variable.”

Muscle activity is, in this sense, a readout from a network transforming state-dependent inputs into movement goals. Rather than choosing muscle patterns in reconfigurable blocks, it creatively constructs and sculpts movement. The hierar-

chy of the motor system may not be rigidly organized around a particular set of variables. As shown in {+@fig:motor\_system}, many loops exist connecting cortex with the spinal cord, the cerebellum, the basal ganglia, and the sensorimotor periphery. Each of these loops contributes information for the flexible activation of the relevant action maps. Put simply, prevailing evidence suggests that cerebellar loops provide predictive state information while basal ganglia loops provide state and/or action value information. Taken together, this work provides an image of the incredible complexity which generates dexterous movements of the hand. This is the foundation on which we can work to build experiments which elucidate the computations involved in the production of skilled movement. We aim to connect our results back to what is known about the system we are attempting to reverse-engineer in order to inspire future inquiries into the inner workings of the movement machine.

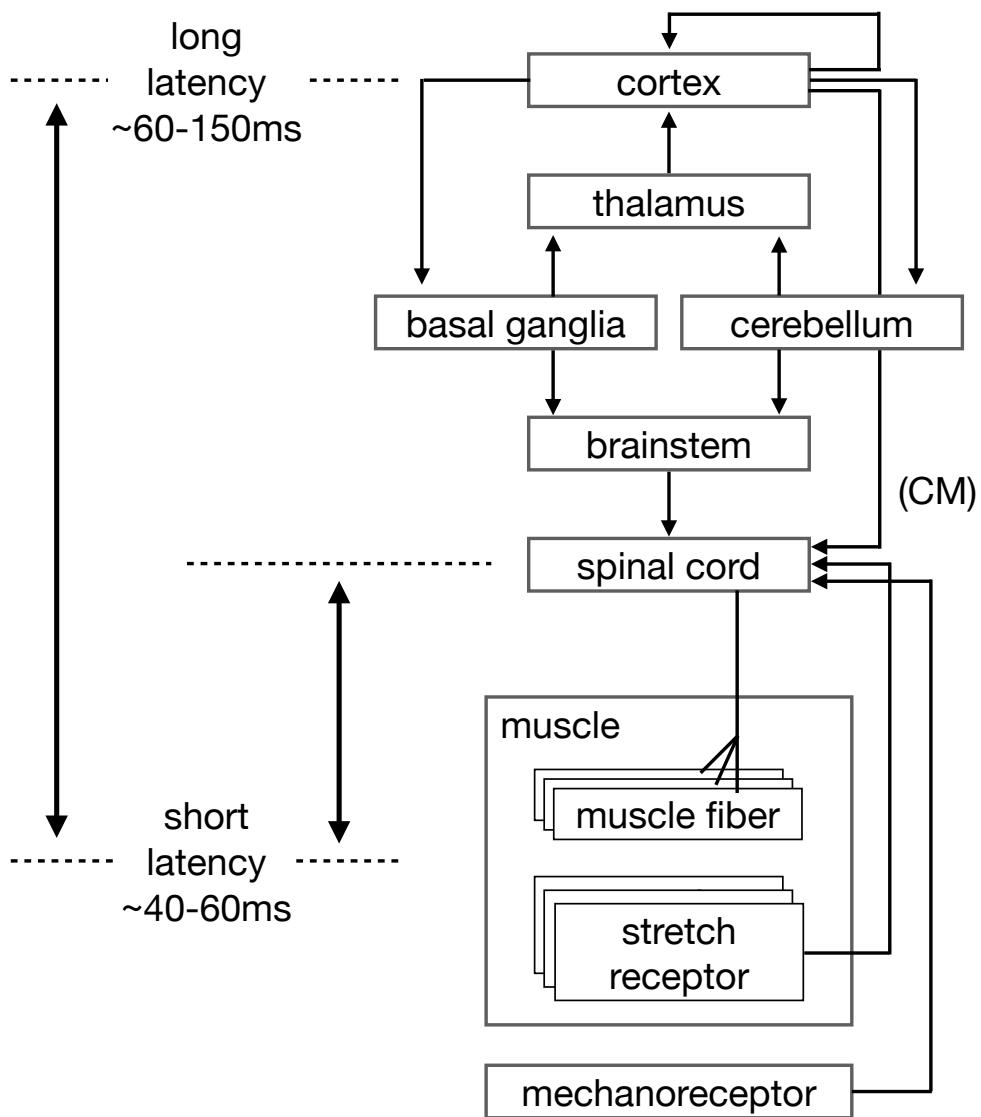


Figure 2.3: Overview sketch of the motor system depicting the the redundancy of the system both hierarchically (multiple muscle fibers are innervated by the same motor neuron, many motor neurons innervate the same muscle) as well as heterarchically (parallel spinal, corticomotoneuronal, cerebellum, basal ganglia feedback loops). Parallel reflex responses can be classified as long latency (approximately 60-150ms) and short latency (approximately 60ms). We hope to consider the parallelism and redundancy of the motor system to inspire our data analyses and models of motor computation.

#### **2.4.4 John Rothwell and Jens Bo Nielson: Voluntary Control**

In the vast majority of cases, cortical inputs first contact interneurones which then relay the commands to motoneurones. Since the same interneurones also receive continuous input from sensory receptors (and hence might be thought to participate in spinal reflexes) as well as from interneurones from other parts of the spinal cord, this means that by the time cortical input reaches motoneurones it has been filtered by multiple lower level systems. In higher primates and in man, cortical input can access some motoneurones via a special direct pathway (the corticomotoneuronal pathway), which is often supposed to play a critical role in volitional movement. However, even if this input is strong (and there is little comparative evidence on this) the excitability of motoneurones will have been biased by the multiple other inputs that each one receives. Thus, even this connection does not guarantee the brain a straightforward control of muscle.

We argue that distributed cortical projections allow for flexibility of connections between muscle representations, and therefore are critical to the flexibility of movements unrestricted by postural demands. Physiologically, the cortex is the main gateway for visual inputs to enter and influence motor control. This is particularly relevant during reaching with the arm and during the swing phase of gait for the leg. In both cases, the limbs are relatively free from feedback control from gravitational and contact force sensors, and can therefore be driven to a large extent by visual inputs.

Posture and contact dictates much of corticospinal function, and we would expect that these demands influence the architecture of the underlying circuitry. “Conscious” control is likely simply the availability of visual and propriospinal sensation in the course of the movement.

## **2.5 Mathematical Theories of Motor Control and Learning**

# Notes

3. In a classic study, Basmajian and colleagues showed that it is possible to activate single motor units in the thumb abductor.



## Chapter 3

# Experimental Literature Review

*We have some idea as to the intricate design of the puppet and the puppet strings, but we lack insight into the mind of the puppeteer. —*

Emilio Bizzi & ??? Ajemian ??? 2020

!– 1. System Identification – learning a transition function  $p(y_t|x_t, u_t)$  - How do you learn the unknown observation model from data?

1. Policy Optimization - Once dynamics are learned (or at least stable?), how do we form a policy that is generalizable to new tasks under these dynamics? - This is the control problem.

It's safe to assume that these processes are happening in parallel. Because we have complete and arbitrary control over the observation mapping, we can ask the subject to interact through a dynamic that is intuitive (informative prior) or unintuitive (uninformative or inhibitive prior). Each scenario, we hypothesize, will elicit different strategies for learning and control.

### 3.0.1 Control

Is LQR (as it's claimed to be) a reasonable model for feedback control and error reduction + variability prediction for dimensionality reduction-based motor interface

(task reads out from D muscles, find modes of that data; do PCA to get K < D dimensions, controller only responds to motion in those K directions)—does behavior + motor activity follow LQR? this question has already been asked, but it hasn't been asked for this kind of high-to-low dim mapping. It's been asked in tasks where muscles haven't been directly in control (Bolero 2009). Todorov: do a task, look at muscle signal. Muscles that aren't necessary for task have higher variability b/c they're not being optimized for task (but does't introduce perturba-

tions). Also see Loeb (2012) for a negative result saying that muscle coordination is habitual rather than optimal, but it has issues (low # muscles). Can we replicate previous reaching optimality results in our set-up? What's unique about our set-up is the PCA/dimensionality reduction in muscle activity space. This is important because you can create arbitrary muscle-cursor mappings, so you have to learn a new skill/mapping. This is different than perturbing a fundamental movement and forcing adaptation, which is what has been previously done. For our task, the participants actually have to learn a new task/mapping, rather than just do what they already know and be robust to perturbations. We test the LQR hypothesis once they've learned the task, because LQR isn't a learning theory, it's a theory about optimal control. We can see if, once people learn a new skill, their behavior is optimal wrt LQR theory. If we establish this, then we can think about how this LQR model is actually learned (enter RL).

Our results are consistent with a recently described model in which an optimal feedback control policy is calculated independently for each potential target and a weighted average of these policies (that is, feedback gains) is computed at each point in time based on the relative desirability of each target<sup>50</sup>. Notably, this model, which pre-dicts averaging of feedback gains, can also account for spatial (that is, trajectory) averaging in go-before-you-know tasks. We submit that our result showing feedback gain averaging, coupled with previous work demonstrating trajectory averaging, provides strong support for the compelling idea that the CNS, under cases of target uncertainty, encodes in parallel multiple motor plans, along with their associated control policies, for competing action options. (Wolpert Nature 2018 competing control policies)

Nashed 2014 — short-latency R1 and long-latency R2 responses (60ms; 45–75ms; 75–105 ms) stretch responses R1 show dexterity (Andrew Pru 2019,2020) in holding and in reaching

[?] cooled areas of motor cortex and mapped them onto OFC— limitation of

this is the behavior, it's a postural maintenance task, not a reach or a complex movement...OFC is definitely the best model here. We should be mindful about modeling things that go beyond this model, they should in some way be able to reduce down to OFC in the limit of postural maintenance, perhaps...

### Adaptation of Reaching

Sinusoidal over trials visuomotor rotations allowed experimenters to disentangle perturbation error sequence over trials in “driven” and “undriven” dimensions by knowing the frequencies of the rotational perturbation. Explicit strategies were given by participants before moving, all errors based off this [?]. This is a very clever experimental design but leaves many questions. Particularly, implicit adaptation seems to be driven by a combination of performance error and sensory prediction error, but how can we tease this apart?

Although little is known about the specific error signals that drive these different processes, an intriguing possibility is that distinct components of implicit learning are driven by performance errors and sensory-prediction errors.

The convergent findings of the SEM and time lag analyses, based on the amplitudes and temporal structure of implicit and strategic adaptive responses, point to an implicit learning process that actively responds to compensate low-fidelity explicit strategy.

Our simulation also reproduced the results of the SEM analysis and the temporal lag analysis, which demonstrates that the low-noise process (which models implicit learning) lags behind and effectively compensates for the inappropriate behavior of the high-noise process (which models strategy), a result that can be predicted by mathematical derivation (Supplementary Math Note).

In our experiment, we might call all learning implicit, since there is so much unknown what would we constitute as a strategy? Think we need to make clear the difference between learning and adaptation. Adaption, as used here, is correction to

a perturbation over trials of a well-practiced movement. Learning is the discovery of unfamiliar movement patterns through trial and error.

An important idea in motor skill learning research is that motor learning proceeds from predominantly explicit to implicit states as a learner develops from novice to expert.

This is interesting, as in our experiment we might think the opposite? Depends on how we define explicit and implicit. Here the definition is loosely deliberate and effortful versus automatic and intuitive. Their point here might be mirrored in overall muscle activation, early on it's actively searching. But what for? And how? We might think of this as a foraging task, searching for information about the unknown mapping— movements as hypotheses.

- prisms
- rotations
- forcefield
- nothing — van Beers variability

The vast majority of research in motor learning studies this capacity through adaptation paradigms in which a systematic perturbation is introduced to disrupt a well-practiced behavior, such as point-to-point reaching. [?]

- classic reaching adaptation -> this is a different goal
  - shadmehr
  - krakauer
- unperturbed movements
  - van beers

[?]

## Arbitrary Visuomotor Mappings

There exist a handful of prior studies mapping EMG activity and finger joint angles directly to virtual stimuli, though few are focused on the learning process and none have the input dimensionality we aim to achieve in work proposed here.

[?]

The null space generated by our glove-cursor map had effectively 17 dimensions ( $19 - 2$ ). We observed a marked tendency of subjects to reduce the amount of motion in this null space (Fig. 4). The selective reduction of null-space motion is particularly important because it may reveal how the Euclidean metric of the task space (the monitor) is effectively “imported” into the coordination of hand. The tendency to generate finger motions with smaller null-space components suggests that the movements tend to remain confined to subspaces that are minimum-norm images of the cursor space. This observation provides us with further evidence that the motor system is effectively capturing the metric structure of the controlled space and that it uses this metric as a basis to form coordinated motions of the fingers. [?]

...we have observed a general tendency of subjects to reduce the amount of finger motion, again suggesting that they are learning trajectories, not just final positions or by points. Our data also show a strong and progressive decrease of movement variability from day to day along the entire motion. This is in sharp contrast with the hypothesis that, through practice, subjects learn to export increasing amounts of variability into the null space to achieve a less-variable task execution.

hand configurations and glove signals are related by a nonlinear isomorphism, rectilinear motions of the cursor on the monitor are not compatible with rectilinear motions in the space of finger-joint coordinates.

what is the natural space of motor activations? what do these spaces look like?  
how are they mapped?

There are several studies using non-EMG-driven sensorimotor mappings to study human motor control and learning.

- Remapping Hand Movements in a Novel Geometrical Environment  
<https://www.ncbi.nlm.nih.gov/pubmed/16148276>

Hinton, Fells palsy study

takehome: humans are really good at learning tasks like these, especially with their hands. this type of dexterity is specific to primates if not humans. let's use this ability to understand and try to model how this learning process unfolds.

***What does this give us that a force-field reaching task can't?***

[?]

## Skill Learning Tasks

- skill learning tasks
  - ball and cup
  - dart throwing tasks

## Learning in Cortical Interfaces

- cortical BMI work
  - Batista papers, lee miller papers
- speech learning – analogy to speech
- bird vocal learning
- we're doing the same experiment, at the muscle level
- try to convince why this is useful, but not too hard

## Skill Learning in Myoelectric Interfaces

“Motor learning explored with myoelectric and neural interfaces” (Nazapour, Jackson)

Experimental myoelectric and neural interfaces can create simplified sensorimotor worlds in which the map from motor commands to effectors can be precisely controlled. These abstract sensorimotor paradigms offer an opportunity to explore further fundamental motor learning questions that would otherwise be obscured by the anatomical and biomechanical complexity of the limbs. By introducing highly artificial and unusual sensorimotor mappings, we may ask whether the human motor system is constrained to naturalistic behaviors or can adapt to circumstances outside the normal ethological repertoire

biomimetic interfaces v. abstract interfaces

flexible use of divergent pathways to form new, task-specific muscle synergies

There is an approximately linear relationship between the rectified EMG and force under conditions of isometric muscle contraction.

several hundred trials of practice

the distal muscles are not limited to a small number of fixed synergies

Optimal Feedback Control is one of our most complete descriptions of the computations that must be performed by the motor system. But it says little about how minimum intervention policies are learned and implemented by neural elements. It remains to be seen whether its predictions can be reconciled with the neuroanatomical evidence for divergent pathways underlying muscle synergies.

We find that with training, comparable results can be obtained irrespective of which distal muscles are used to control the task.

task-specific correlation structure

[?] bypass synergies hard-wired into subcortical and spinal circuitry, allowing the motor cortex considerable fine control over specific mus-

cles. The distal bias of cortico-motoneuronal connections likely explains why these muscles are better suited to control of abstract myoelectric interfaces than proximal muscles.

A typical cell may exhibit post-spike effects in several forearm and hand muscles, which defines the muscle field for that cell (Fetz & Cheney, 1978; Jackson et al., 2003). Across the population a wide variety of different muscle fields are seen for individual corticospinal neurons.

We have already seen how redundancy in the mapping from muscles to movement can be exploited by distributing effort optimally across multiple muscles. We now find that redundancy in the mapping from cortical neurons with divergent muscle fields provides a further opportunity to distribute effort among multiple muscle synergies. The picture that emerges is one of hierarchical levels of redundancy within the motor system (Figure 4.6). At each level, convergence onto fewer dimensions in the level below means a multitude of redundant activity patterns are consistent with the desired goal. At the same time, divergence in the descending pathways allows behavior to be optimized for high-level task goals like accuracy. However, a high-dimensional control space is a double-edged sword; the flexibility to optimize the motor pattern for any abstract high-level goals comes at the cost of a large search space within which this pattern must be sought. The learning mechanisms that allow the motor system to achieve optimality are the subject of much current research.

---

Ctrl-labs motor BMI [?] – decoding motor units and tendon electric fields at the wrist, but nothing really about learning...

Due to the unstructured nature of the exploration period, we first decomposed motor unit firing rates into separate components via non-

negative matrix factorization (NMF) to identify groups of units that were often mutually active. We fixed this number of components to 3, aligning with the instructions given to the participant to ultimately select 3 representative motor units. [?]

NMF is  $X \approx WH$

Since the relative scales of the projections ( $W$ ) and its components ( $H$ ) are typically arbitrary, we resolved ambiguity by scaling each component to unit L2-norm and scaling its corresponding transformation by the appropriate reciprocal factor.

We trained 8 participants over 6 consecutive days using this system on a center-out task requiring both individual and simultaneous control of three motor units. We showed that participants demonstrated improvements in performance both within and across days. Through comparisons to isometric, ramp-and-hold contractions, we provide evidence that neurofeedback enabled participants to expand their ability to control individual motor units outside of naturalistic movement constraints.

Taken together, these results reveal the center-out task enabled both a significant, population-level increase in dimensionality relative to during stereotyped, isometric contractions and an increase in unexplained variability in the unselected motor unit population. [?]

Dimensionality here is the “participation ratio”:

The participation ratio (PR) was computed to quantify the dimensionality of the iEMG and firing rate data<sup>83–85</sup>. The PR is a metric computed on the covariance matrix of a feature and represents the approximated-dimensionality of the manifold spanned by that feature; a higher participation ratio means more principal components are needed to explain a given proportion of the feature’s variance.

“performance levels and rates of improvement were significantly higher for intrinsic hand muscles relative to muscles of the forearm.” [?]

## synergies and learning

the motor system can acquire new muscle synergies during motor skill learning, especially for skills that cannot be adequately accomplished by deploying preexisting synergies... The exact mechanism responsible for 581 this learning has not been established. But during early skill learning, the motor system may 582 discover the direction of synergy change by exploiting and modulating the intrinsic variability 583 of the synergies and of their temporal activations, and subsequently drive this change by 584 reinforcing the synergy patterns that lead to reward-producing actions (Cheung et al., 2020b) in 585 a manner analogous to how reinforcement learning relies on action exploration [?]

Berger et al. 2013

[?]

Berger et al.’s preprocessing - lowpassed butterworth at 5Hz - normalized to MVC calibration - periodic baseline noise subtraction - choosing synergies by uniformity in force direction (pretty arbitrary)

Using EMG in a learning experiment is not unheard of. Berger et al. 2013 use EMG with 13 muscles to test whether learning new synergy combinations for a task is more difficult than recombining existing synergies@Berger2013a. As we would expect, learning new synergy combinations is more difficult. I would argue that the demand in their “incompatible virtual surgeries” is too strict, that we need to more carefully design synergy perturbations to develop a model of learning in such a task.

Berger et al. fit a muscle-space to force-space mapping  $H$  using a force-driven calibration task, and a synergy-space to muscle-space mapping  $W$  using NMF.

$$f = Hm$$

$$m = Wc$$

$\dim(m) = M$  muscles,  $\dim(c) = N$  synergies, and  $\dim(f) = D$  dimensions of task space where  $M > N > D$ . Because  $H$  and  $W$  are rectangular, they have at most rank  $D$  and  $N$ , and we constrain these matrices to be full rank. There are three key subspaces: the nullspace of  $H$  mapping muscle activations to 0, the column space or range of  $W$  mapping synergy activations to muscle activations, and the common subspace between these two. That is, there are synergy activations which generate muscle activations which lie in the null space of  $H$ . The paper uses this fact to develop mappings that specifically rotate muscle activations produced by synergies into the null space of  $H$  which were not there prior to rotation. The dimensionalities of these subspaces are defined:

$$\dim(\text{null}(H)) = M - D \quad \text{muscle vectors} \rightarrow 0$$

$$\dim(\text{col}(W)) = N \quad \text{synergy activations} \rightarrow \text{muscle subspace}$$

$$\dim(\text{null}(H) \cap \text{col}(W)) = N - D \quad \text{synergy activations} \rightarrow 0$$

In the paper, the authors find an orthonormal basis  $W_o$  for the range (column space) of the synergy weight matrix  $W$  (presumably using a QR factorization) and find the nullspace  $H_{null}$  of  $H$ . These computations are done presumably through QR factorizations (an orthonormal basis multiplied by a rotation and scaling) by finding  $Q$  in the first case and finding the latter  $M - D$  columns  $Q_2$  of  $Q = [Q_1 Q_2]$  which are  $H_{null}$  in the second case:

$$W = Q_W^{M \times M} R_W^{M \times N}$$

$$= \begin{bmatrix} Q_{W,1}^{M \times N} & Q_{W,2}^{M \times M-N} \end{bmatrix} \begin{bmatrix} R_{W,1} \\ 0 \end{bmatrix}$$

$$W_o = Q_{W,1}^{M \times N}$$

$$W_o^T = Q_{W,1}^{T, N \times M}$$

$$H_{M \times D}^T = Q_H^{M \times M} R_H = \begin{bmatrix} Q_{H,1}^{M \times D} & Q_{H,2}^{M \times M-D} \end{bmatrix} \begin{bmatrix} R_{H,1} \\ 0 \end{bmatrix}$$

$$H_{null}^T = Q_{H,2}^{M \times M-D}$$

To find each of the three subspace, we take the SVD of the composition  $W^T H^T$

$$W_o^T H_{null}^T = Q_{W,1}^{T, N \times M} Q_{H,2}^{M \times M-D}$$

$$= Q_W^T Q_H \dim(N \times M - D)$$

$$= U_{N \times N} \Sigma_{N \times M-D} V_{M-D \times M-D}^T$$

$$H_{null}^T V = W_o U \Sigma$$

Now we can pick out the three subspaces using the SVD

$W_c = W_o U[1 : N - D]$	synergy activations → muscle activations in task null space
$H_c = H_{null}[1 : N - D]$	synergetic muscle activations → 0
$W_{nc} = W_o U[N - D + 1 : N]$	synergy activations → nonzero muscle activations
$H_{nc} = H_{null} V[N - D + 1 : M - D]$	non-synergetic muscle activations → 0

To construct new mappings, the authors construct rotations to alter muscle ac-

tivation vectors by rotating them from  $W_{nc}$  and remaining in  $W_{nc}$  and from  $W_{nc}$  into  $H_{nc}$ . In the first case this alters the mapping by changing the effective muscle activations produced by the existing (learned) synergistic actions. That is, muscle activations putatively produced by synergistic action will be altered to produce different forces in task space (compatible rotations). In the second case, muscle activations putatively produced by existing synergistic action (via  $W$ ) will be mapped into the null space of  $H$  and produce zero force in task space (incompatible rotations).

A key critique of this paper is that such a transformation is too harsh. The compatible rotation allows you to recombine the same muscle patterns, the incompatible doesn't allow you to use existing coactivation patterns at all. The authors do see new synergies emerging even after their training session, consisting of:

- 16 trials of maximum voluntary contraction in 8 directions (calibration)
- 72 trials using force control (calibration)
- 24 trials familiarization
- 144 trials baseline
- 288 trials surgery
- 144 trials washout
- 144 trials baseline

After 288 trials subjects aren't able to complete the task for some movement directions.

Nazarpour 2012 J.Neuro

[?]

Flexible Cortical Control of Task-Specific Muscle Synergies <https://www.jneurosci.org/content/32>

Fig. 4A – cursor controlled muscles begin to dissociate from non cursor controlled muscles.

Feedforward processing to muscle fields / tunings in the presence of signal dependent noise

Feedback processing based on visual errors

## Radhakrishnan 2008

[?]

x Learning a Novel Myoelectric-Controlled Interface Task — Radhakrishnan, 2008 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2576223/>

proprioception is not required to learn nonintuitive MCI mappings

several hundred trials subjects learned pointing with six muscles

prism adaption requires active movement; efference copy implicated if proprioception doesn't seem to be required

control models Fig 10

## de Rugy 2012 - Habitual not Optimal

[?]

just because it's harder to adapt to incompatible surgeries doesn't mean that there are fixed synergies, it just means there are multiple timescales of adaptation available in the neural control hierarchy – diversity in the neural controller depending on context

learning inverse model may be separate from learning to optimize trajectories on top of that model – some type of “fine tuning”

skill acquisition (slow, constructing novel synergies) vs. motor adaptation (less slow, adapting existing synergy activations)

It's a good test, but it pushes the optimal control framework too hard? perhaps we need a model for what “good enough” is? If we penalize moving to a new controller from a previously optimized movement, the findings make sense. An optimal control model would predict the exactly optimal coordination patterns for the new scenario, it wouldn't say anything about adaptation from an old solution to a new one. This is why we need to develop a model of adaptation that formalizes this scenario not of kinematic perturbations (noise during movement), but to a drastic change in the plant itself (e.g. muscle failure).

## Mussa-Ivaldi 2019

Earlier theoretical work by Jordan and Rumelhart [?] considered how the learning of actions can be viewed as the concurrent learning of forward and inverse models of actions. **They introduced the concept of distal learning, where the learner has to find a mapping from desired outcomes to actions in order to achieve a desired outcome. To do so, the subject begins by forming a predictive forward model of the transformation from actions to distal outcomes. Such transformations are often not known a priori, thus the forward model must generally be learned by exploring the outcomes associated with particular choices of action. Once the forward model has been at least partially learned, it can be used to guide the learning of an inverse model that predicts the action needed to achieve the distal outcome.** Mussa-Ivaldi2019

Our findings are consistent with the hypothesis that learning proceeds through the concurrent evolution of coupled forward and inverse models of the body-to-object mapping established by the BoMI. Mussa-Ivaldi2019

Not being square, the matrix  $H$  does not have a unique inverse. But there exist infinite “right inverses” that combined with  $H$  yield the  $K \times K$  identity matrix in the task space of external control signals. Each such right inverse transforms a desired position of the controlled object into one particular set of values for the body signals. We consider users to be competent when they are able to move their body successfully in response to a presented target for the controlled object. Mathematically, we consider this as finding one right inverse  $G$  of the mapping  $H$ , out of a multitude of possible and equally valid choices. Mussa-Ivaldi2019

Gradient learning of a forward and inverse model (mapping):

$$\hat{H}_{n+1} = \hat{H}_n + \epsilon(p_n - H_n q_n) q_n^T$$

$$G_{n+1} = G_n - \eta \hat{H}_n^T e_n u_n^T$$

$$e_n = p_n - u_n$$

The comparison between model predictions and actual data in Fig 3 indicates that our proposed model of learning is sufficient to explain the data. However, the mechanism we propose is not necessary; we cannot rule out other possibilities, such as reinforcement learning. [?] This agreement between model and experimental results does not exclude the possibility of alternative learning mechanisms, such as a direct learning of the inverse model [?] or the use of reinforcement learning [?] to acquire an action policy that would play the role of the inverse model.

How do we break a simple gradient model? On a task that is more difficult? will learning take longer? - savings phenomenon -

Although the interface forward map is linear (Methods, Eq (5)), this is a many-to-one map admitting a multitude of inverses. This “redundancy” opens the possibility of successful linear and nonlinear inverse maps. Redundancy also leads to an important consideration about gradient descent learning. The reaching error surface in the space of the inverse model elements does not have a unique minimum, but a continuously connected set of minima corresponding to the null space of the forward map. In the metaphor of a skier descending from a mountain following the gradient, this space of equivalent inverse models corresponds to a flat elongated valley at the bottom of the mountain. Anywhere along the valley is a valid end to the ride, as it corresponds to a valid inverse model. The inverse model on which the steepest descent ends depends on the initial conditions, as predicted by the dynamical model (see Fig

3b–evolution of the norm of the inverse model error), as well as on the realization of the noise employed in any given simulation of the learning model.

Although the two-dimensional subspace formed by the first two PCs captured a large fraction of the total variance of body motions, it did not necessarily reflect the natural up-down/left-right orientation of the display monitor. Therefore, following calibration and PC extraction, there was a customization phase in which users were allowed to set the origin, orientation, and scaling of the coordinates in task space, based on their preference.

Subjects have prior knowledge of their directions in task space?

x 90% isn't enough – Follow-up on the previous paper – critiques “direct evidence” <https://www.biorxiv.org/content/10.1101/634758v1>

- Structured variability of muscle activations supports the minimal intervention principle of motor control <https://www.ncbi.nlm.nih.gov/pubmed/19369362>

This is the 7-muscle fine-wire version of the experiment, which we'll use to compare our findings about null and task space variability.



# **Notes**



## **Chapter 4**

# **Experimental Methods**

*We have some idea as to the intricate design of the puppet and the  
puppet strings, but we lack insight into the mind of the puppeteer.*

— Emilio Bizzi & ??? Ajemian  
???, 2020



## 4.1 Methodological Aims

what we want: continuous action space (difficult) adaptation vs. skill learning look different behavior changes over learning

We want the subject to embody, as much as possible, a infant-like state of ignorance, not intellectually, but within their bodies, a small-k knowledge of themselves.

- We think this will “stretch out” learning and let us look at the redundancy problem from a new perspective. We want to give subjects a more naturalistic environment in which to solve a motor problem. - This is at odds with typical motor tasks, gives us more control at the redundancy problem

- While motor control has many results, most of which focus on optimal control, we’re interested in motor learning— how can we stretch learning to get from “zero” (or do we ever start at zero?) to performant? How do we develop a new skill? How do we learn to control in a new context/environment/contingency? If subjects in motor tasks tend to reach certain definitions of optimal performance, what comes before optimality?

- What kind of task/experiment should we devise to tackle these questions? What unique vantage would this paradigm offer that existing paradigms do not?

To explore dexterous movement, we will leverage recordings of muscles controlling the hand as a readout of flexible motor behavior. This is a step beyond recording hand kinematics as electromyography provides a physiological output of the nervous system. Surface electromyography recordings taken from the forearms controlling subjects’ dominant hands allows us to track the sequential selection of muscle activations during both skill acquisition and subsequent performance of that skill to achieve desired goal. As we are interested in subjects’ abilities to acquire new skills, we design tasks that require subjects to use available, but uncommon, motor activations. We then track the selection and execution of these activation during virtual tasks. Preliminary work in this direction is described in ??.

The concept of the experimental setup is shown in Figure 4.1, where 32 monopolar electrodes are attached to a subject’s forearm to record muscle activity. The arm and hand are kinematically constrained in a custom fixture and mo-

tor activity is recorded during low-level isometric muscle contractions. The setup circumvents the limb biomechanics by mapping muscle output directly to virtual stimuli shown on a screen. By focusing on low-force, isometric contractions we intend to avoid complications due to artifacts in dynamic, high-force movements.

The concept of redundancy is common: in the human perception of color, we have a 3-dimensional plane of perception red, green, blue, while the space of physical colors is an infinite-dimensional spectrum. Thus, we can perceive some colors (e.g. yellow light vs. green+red light) with different spectra as identical, as they are identical when “projected” onto our perceptual plane (this is known as metamerism).

## 4.2 Recording Setup

Describe the experimental setup/environment, what is taking place in each session, protocol, software, hardware, how the signal is captured, physical setup

As far as we are aware, this setup is novel in combining a high number of channels with an abstract mapping. Learning experiments have used joint angles and a few muscles (typically movements of the wrist or pairs of thumb and intrinsic hand muscles), but none have taken a data-driven approach in constructing a virtual learning environment in the style of cortical BMI[?]. Our EMG recording setup is custom-built: the “Sessantaquattro” EMG amplifier was acquired from OT Bioelettronica, the electronic connector was designed in-house, the electrodes were acquired from Medkit UK, and the recording software was written in a mixture of Bonsai (C#) and Python. EMG is acquired at 2kHz sample rate with 24-bit precision. A clip of raw data is shown in Figure 4.4.

## 4.3 Raw Data

## 4.4 Natural Movement Task

In our first preliminary experiment, a single subject produced flexions and extensions of each finger in the recording setup without any kind of artificial feedback. One trial was collected per finger movement in three blocks per session and one

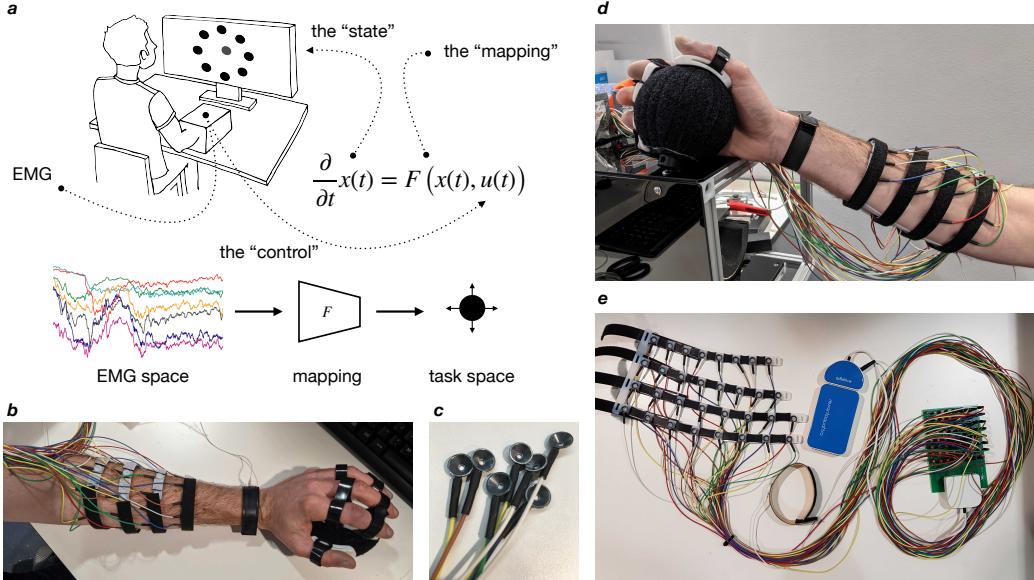


Figure 4.1: (a) Graphic depicting the closed-loop EMG interface concept in a center-hold, reach-out type task. The multidimensional EMG signal is transformed online through a mapping  $F$  from EMG electrode space to a lower dimensional task space, creating an experimentally controllable redundancy problem for the subject. In experiments shown here the task space is two-dimensional, though the EMG interface can be extended to tasks with higher-dimensional inputs. The subject's arm and hand are constrained during the experiment to ensure isometric contractions. (b) First prototype of custom recording hardware consisting of four bands of eight electrodes each, and a spherical hand constraint. Our recordings are 32 channel monopolar recording with reference electrode at the wrist. (c) Example cup-style monopolar recording electrodes, 5mm in diameter. (d) Side view of the recording hardware. Also pictured is the arm restraint frame to ensure isometric contractions. The frame obscures the subject's arm from view and contains adjustable elbow and wrist rests. (d) Recording hardware shown off the arm with wireless amplifier and connection board.

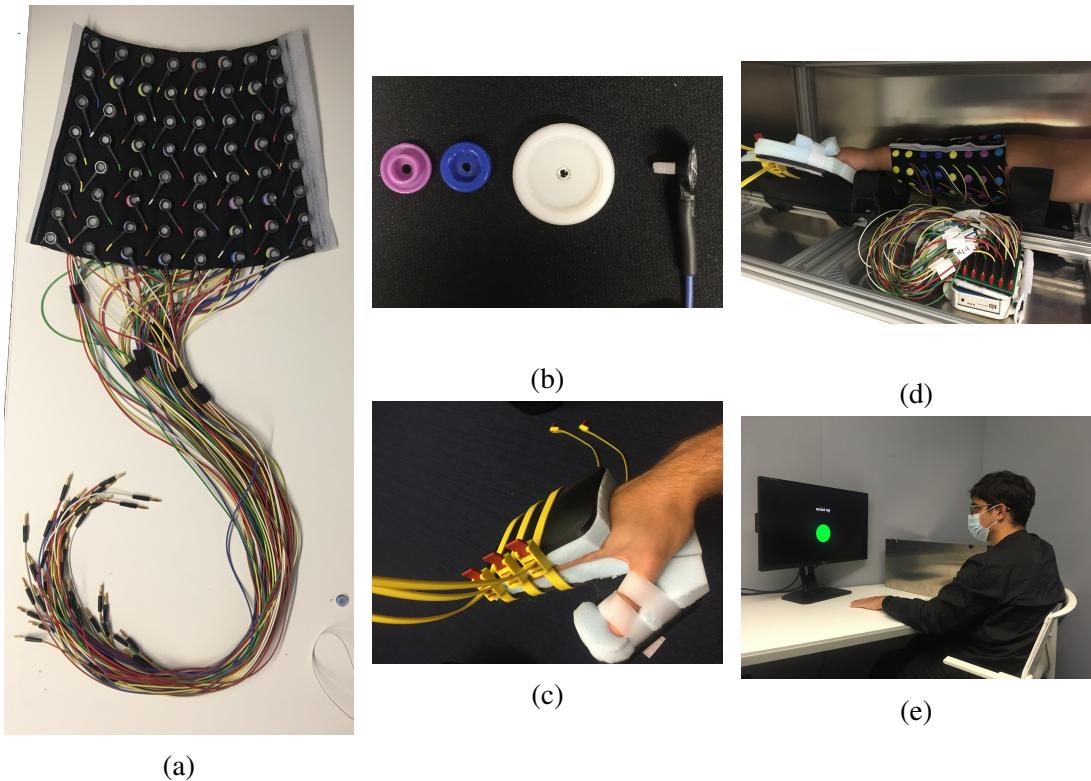


Figure 4.2: (a) (b) (c) (d)

session per day over five days for a total of fifteen trials per finger movement. The purpose of this experiment is to determine the robustness over trials and sessions of EMG features for a simple low-contraction movement, as well as to determine the level of noise and artifacts in the data. Such baseline measurements are important to properly decompose variability due to electrode placement and exogenous noise from behavioral and physiological variability in order to ensure reproducibility of our results. Additionally, this baseline task may prove useful as a benchmark for later tasks in terms of testing analysis and decomposition techniques. A plot of all 32 channels for a single trial after preprocessing is shown in ??.

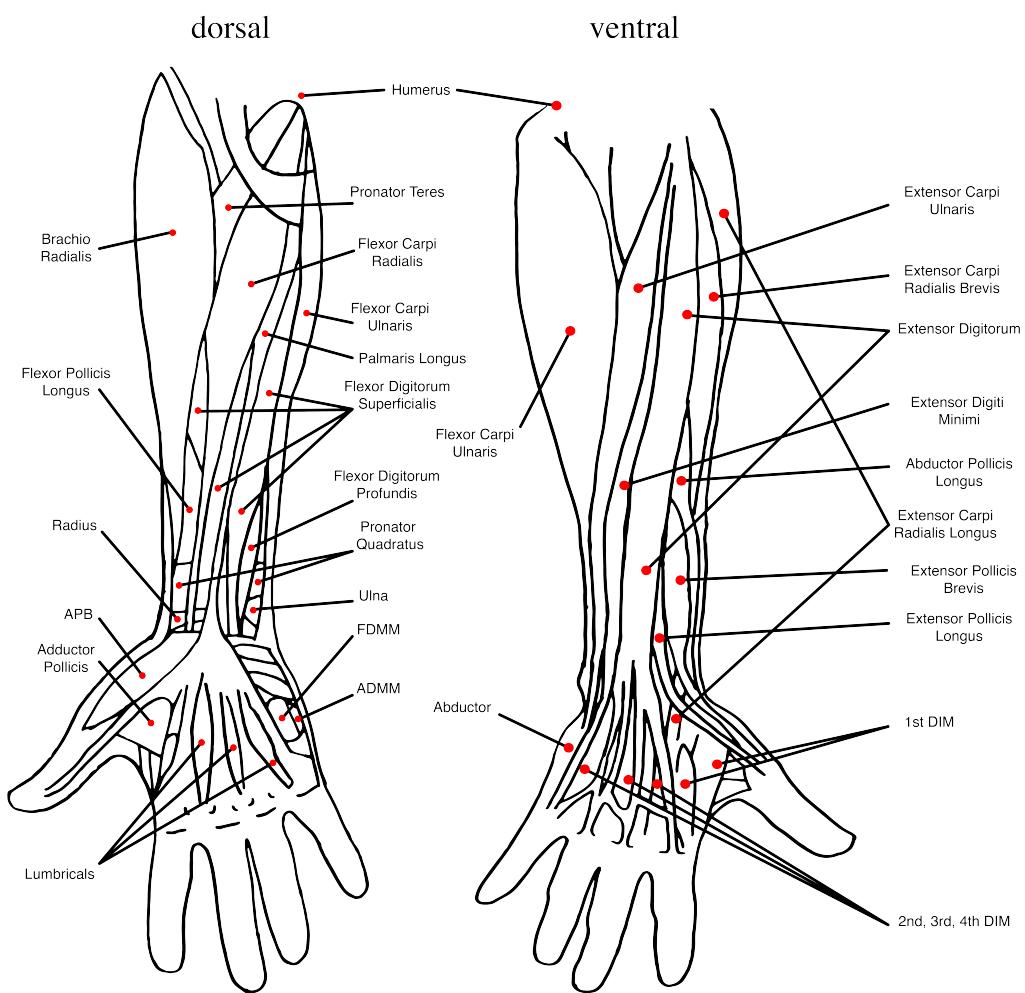


Figure 4.3

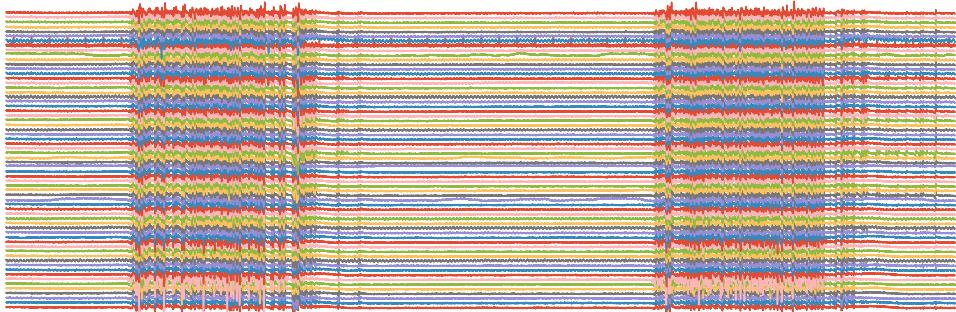


Figure 4.4: Notice: highly correlated across channels, artefacts and noise.  
10 seconds of raw 64-channel EMG data taken during a minimal finger flexion trial.  
Note that some channels include a nontrivial amount of line noise. This noise will  
be drastically reduced through changes being made in the next recording hardware  
revision which include shielding, shorter cables, and better cable routing. Note that  
some channels (e.g. channel 24) show very low noise and putative single motor unit  
action potentials can be seen on many channels.

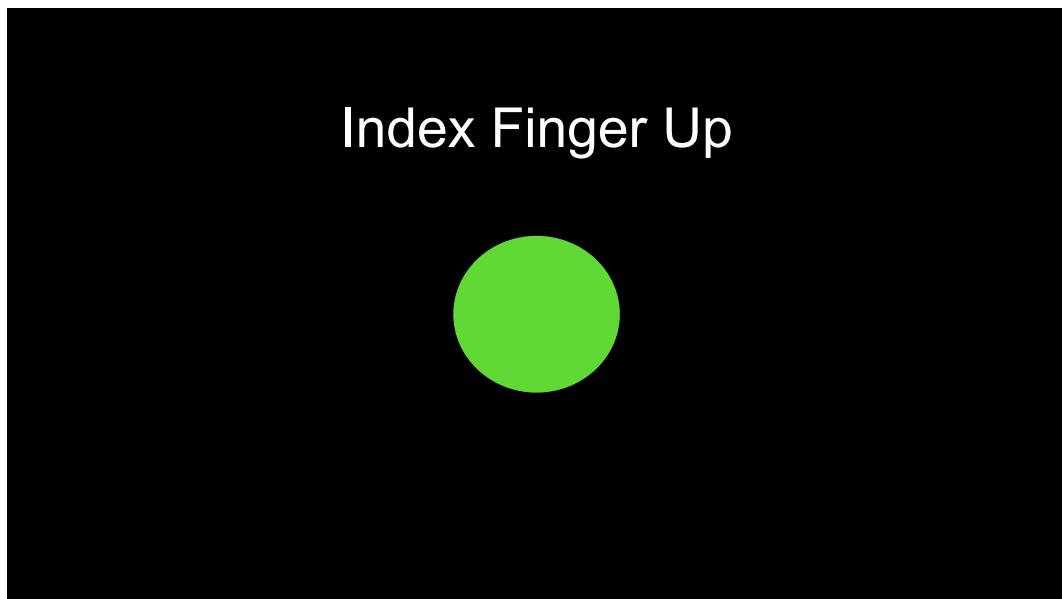


Figure 4.5

Table 4.1: List of movement commands issued during the “Natural Movement” task.

Movement
index finger down
index finger up
middle finger down
middle finger up
ring finger down
ring finger up
little finger down
little finger up
hand open
hand closed
wrist up
wrist down
wrist in
wrist out

## 4.5 Calibration Task

Goal: What is the best calibration task to find the boundaries of the available EMG space?

Does this work?? Can we quantify whether people are actually exploring the space?

To preprocessing the data, simple filtering and rectification were applied, as is commonly done in the literature[?]. As shown in ??, here we apply highpass filtering at 40Hz to remove any low-frequency oscillations and DC offsets, rectification and lowpass filtering at 5Hz to extract what is typically associated with a force readout of the EMG signal in the case that electrodes are positioned over the belly of a single muscle. These filter parameters were chosen by visual comparison across a range of values. While these preprocessing steps are in accordance with the literature and yield a signal with frequencies on a behavioral timescale though, as discussed in Section 8.2, preprocessing of raw EMG signals is an area worth investigating the development and application of more advanced methods.

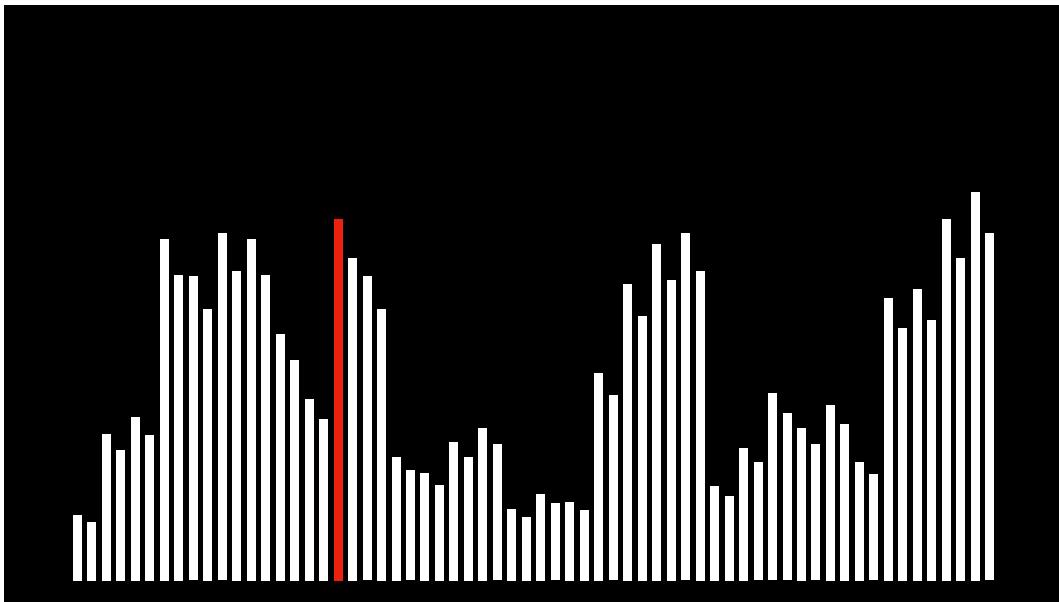


Figure 4.6: Calibration task visual feedback

#### 4.5.1 Does the Calibration Task Encourage Exploration?

Compare to the Movement task! Calibration will show higher variance (maybe look at the determinant of the covariance? across subjects?)

Are people “better” than random EMG at the calibration task? How do we define the “null” hypothesis here, what is comparable random EMG? (E.g. sample from a distribution with the same statistics)

Hypothesis: subjects aren’t moving randomly in this task. They are goal-oriented and are using feedback to solve this search problem.

## 4.6 Decoder Fitting

**How is the decoder derived? Why this way? What does this mean geometrically?**

Electrode data from a single trial of a single session is held in a data matrix  $X$  ( $n_{\text{electrodes}}, n_{\text{samples}}$ ), and we wish to find a latent weight matrix  $W$  ( $n_{\text{electrodes}}, n_{\text{components}}$ ) which reconstructs  $X$  by projecting latent trajectories

$H$  (n\_components, n\_samples) into electrode space:

$$X = W \cdot H$$

$H$  is the activity of the latent processes, and  $W$  is the mixing matrix. The columns of  $W$  are the principal vectors spanning the latent subspace in electrode space. If we have new samples, we can project these new points onto this subspace:

$$h_{new} = W^T \cdot w_{new}$$

To justify this decomposition, we have to make some assumptions about the nature of the EMG signal, namely that the signal is linear instantaneous (each EMG sample can be instantly mapped to control space). The other assumption is that the basis  $W$  should be orthonormal, that the columns of  $W$  are orthogonal with unity norm. This ensures that the left inverse  $W^{-1}$  is equal to the transpose  $W^T$  such that:

$$\begin{aligned} X &= W \cdot H \\ W^{-1} \cdot X &= H \\ W^T \cdot X &= H \end{aligned}$$

See \*Muceli 2014\* for use of the Moore-Penrose pseudoinverse in place of the transpose when the columns of  $W$  do not form an orthonormal basis. This would be the case for NMF. Is there a factorization that produces nonnegative, orthogonal coordinates? Or is the pseudoinverse okay? I will need to test this.

Stated in an information theoretic way, we want to minimize the reconstruction loss  $\mathcal{L}$  for our derived encoder-decoder pair  $(E,D)$ . We're decoding high dimensional activity into its latent dimensions, and encoding back into the high dimensional space.

$$\min_{E,D} \mathcal{L}[X - EDX] \quad (4.1)$$

This way, forget about orthonormality and solve for an encoder and decoder directly. That is,  $E \neq D$  is perfectly acceptable.

Each row of  $D$  might be called a \*\*spatial filter\*\*, a linear combination of electrode activities into a surrogate, hopefully more intuitive space.

$$WH = X$$

$$TxK \times Kx64 = Tx64$$

transformation x components = data

factors x weights = data

inverse:

$$W = XH^{-1}$$

$$W^T = H^{-T}X^T$$

$KxT = Kx64 \times 64xT$  We're not regularizing here, so the cost function for the fit looks like:

$$\min ||X - WH||_{loss}^2$$

The mapping between EMG and task space  $M$  can be written as

$$M = \begin{bmatrix} \tilde{M} & \tilde{M} & \tilde{M} & \tilde{M} \end{bmatrix} \quad (4.2)$$

where  $\tilde{M}$  consists of 8 equally spaced directions, one for each “column” of the 4 EMG electrode bands around the subject's arm:

$$\tilde{M} = \begin{bmatrix} 0 & 0.71 & 1 & 0.71 & 0 & -0.71 & -1 & -0.71 \\ 1 & 0.71 & 0 & -0.71 & 1 & -0.71 & 0 & 0.71 \end{bmatrix} \quad (4.3)$$

A graphic showing the mappings from electrodes to force directions is shown in ???. While there are 8 possible force vectors the subject can modulate by controlling the electrode activity on each of her 8 columns, the EMG mapping is ultimately a projection onto the 2D plane. Since the EMG signal is nonnegative, the subject could technically modulate just four modes of electrode activity, the minimum number needed to span the task space, to reach all 32 targets. This mapping was chosen in order to provide a simple starting point to explore the virtual EMG task. There are no added environmental dynamics, only a redundant readout of force, and the signal is processed online in the same manner as in the offline analyses. This task or a variant we think can serve as a foundation upon which we can build more complex mappings and virtual environments. Task space trajectories from each block are shown in ???. The fraction of trials resulting in hold timeouts, reach timeouts, and target hits are shown over blocks in ???.

**Run NMF on the second calibration and compare – does similarity here correlate with performance?**

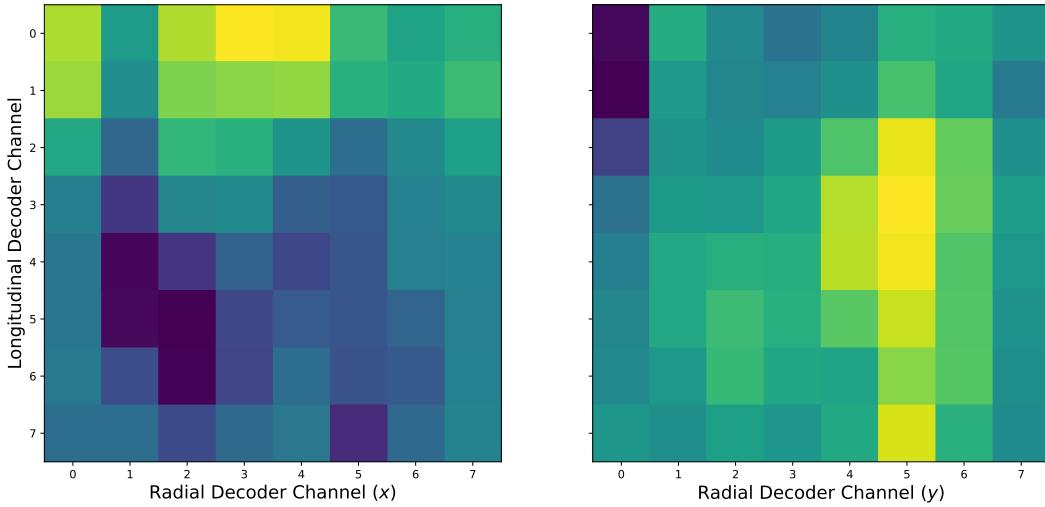


Figure 4.7: Example EMG-to-force decoders from a single subject. The “center hold, reach out” task works by mapping 64 channels of EMG activity from subjects’ forearms to a 2-dimensional force vector, a component acting in the  $x$  and  $y$  directions within the task’s linear dynamics. Depicted here are the two 64-dimensional “decoders” arranged as the EMG electrodes were arranged on subjects’ arms (along the arm, longitudinally, and around the arm, radially). The left plot shows the  $x$  force decoder, and the right plot the  $y$ .

## 4.7 Target Task

“I would be most interested to hear how u are thinking of approaching the analysis. I.e. you have a bunch of channels, movements, tasks what is the workflow to get from that raw data into something manageable/useful?”

What does the basic data look like, discuss the shape of the data we have 46 Subjects x 12 targets x 45 trials x 5k timepoints x 64 channels + 2 x 14 x Movements task + 2x Calibration task

In this task, the 32-dimensional EMG electrode activity vector is mapped to a 2D force acting on a point mass shown on the screen. The mapping  $M \in \mathbb{R}^{2 \times 32}$  maps 8 “columns” each consisting of 4 electrodes placed in a line down the length of the forearm each to one of 2D root of unity. Each column of electrodes is thus mapped to one of 8 two-dimensional force vectors. In this experiment, the point mass has zero inertia and zero friction and as such displays a direct, though redundant, read-out of the EMG signal. The task asks the subject to reach one of 32 equally spaced targets on each trial. Subjects must hold in the center of the task space for a des-

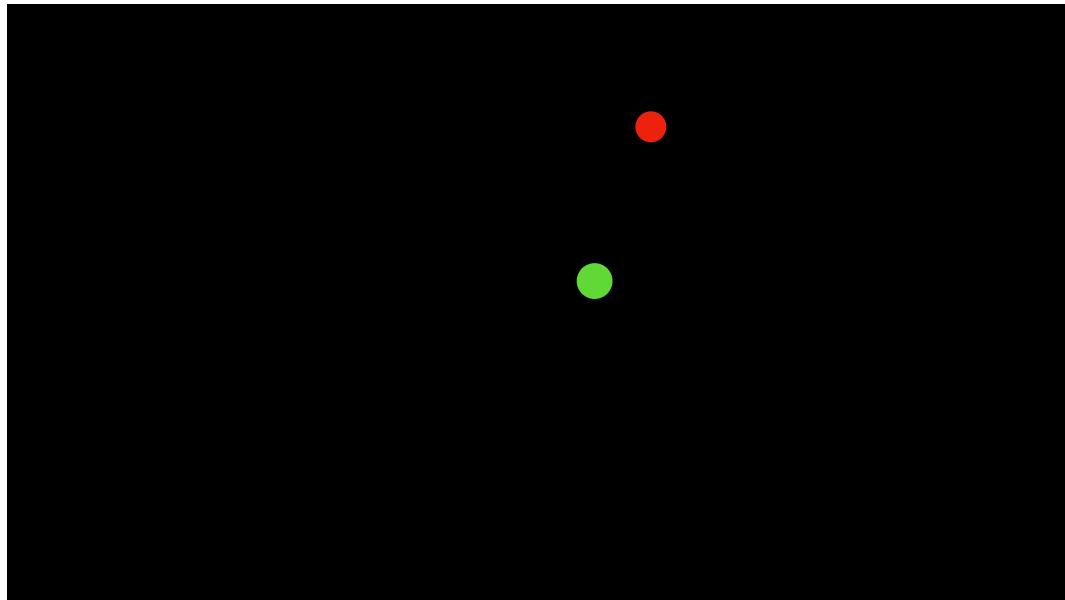


Figure 4.8

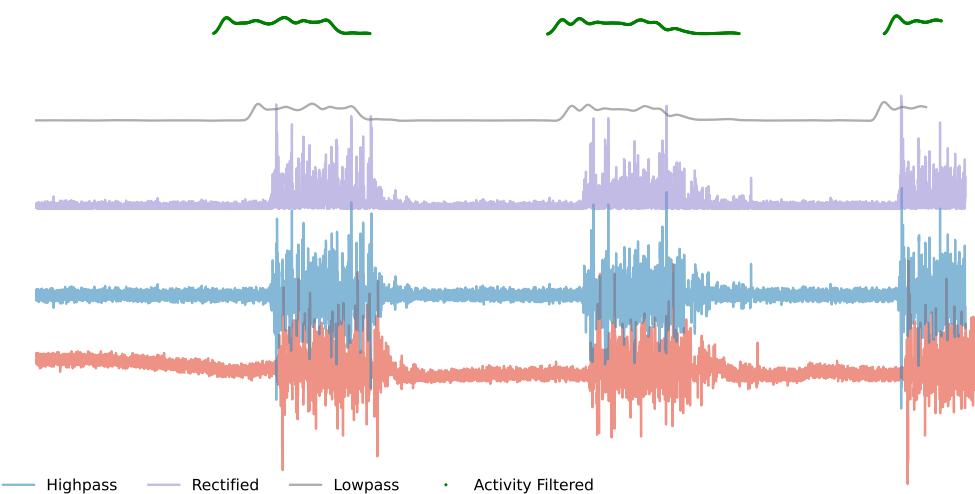


Figure 4.9: Data from a single trial showing each step of preprocessing. The prototype preprocessing pipeline is highpass at 50Hz, rectification, and lowpass at 5Hz. The within-trial, per-channel means are subtracted from each trial. This matches what is typically done in the literature to find a correlate of intended force[@sangerBayesianFilteringMyoelectric2007;@churchlandNeuralPopulationDynamics2012a;@c]. There is significant room for improvement on this workflow, as discussed in the text.

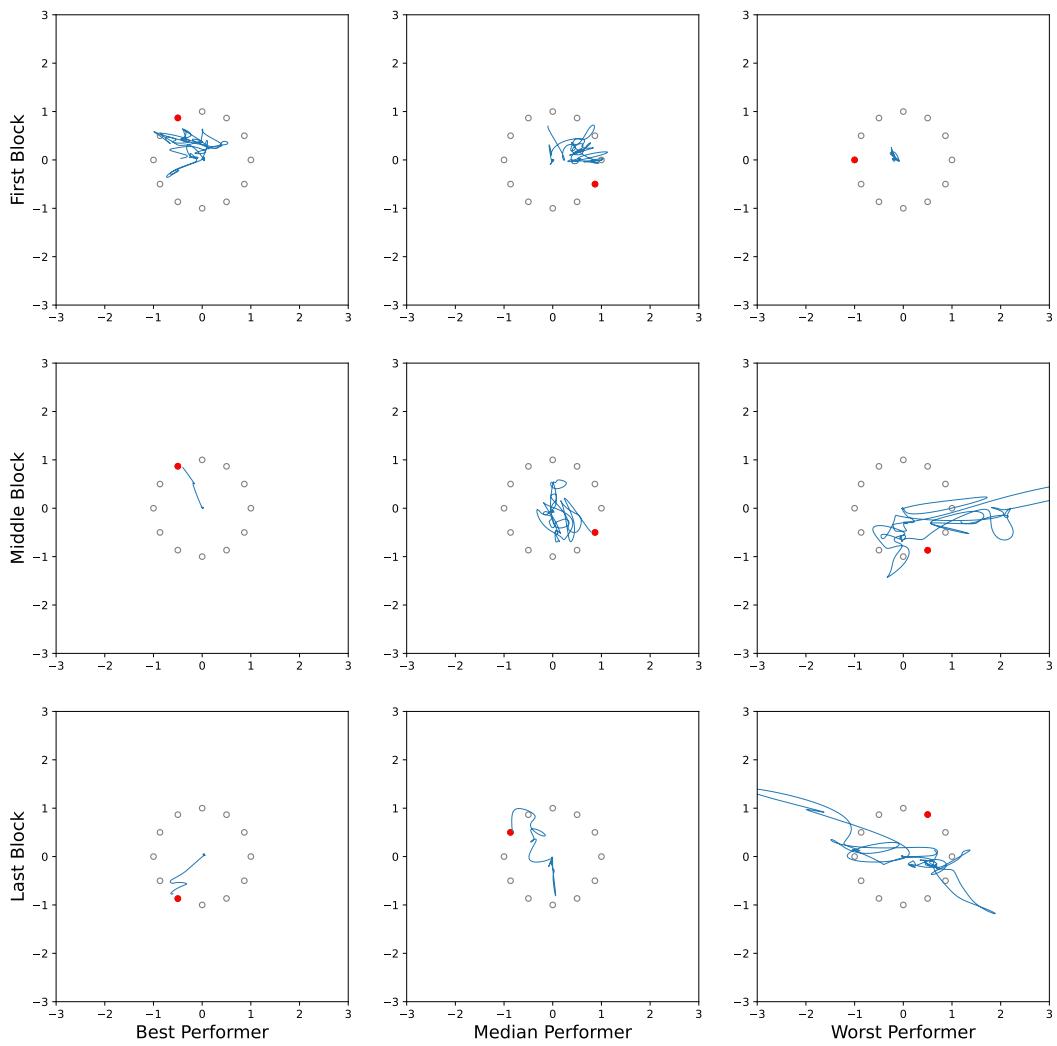


Figure 4.10: Point mass position trajectories in two-dimensional task space during the center-hold, reach-out task with 32 targets spaced evenly around the unit circle (shown with red borders). Color corresponds to target numbers, with target zero located at  $(0,1)$ . Target order was randomized. Training was conducted over 3 blocks each with 32 trials, 1 trial per target.

gnated period of time, after which the target appears. Subjects then have a time window reach the target. Data from one subject was recorded for three blocks in one session. Each block consisted of 32 trials, one per target in a randomized order.

A graphic showing the mappings from electrodes to force directions is shown in ???. While there are 8 possible force vectors the subject can modulate by controlling the electrode activity on each of her 8 columns, the EMG mapping is ultimately a projection onto the 2D plane. Since the EMG signal is nonnegative, the subject could technically modulate just four modes of electrode activity, the minimum number needed to span the task space, to reach all 32 targets. This mapping was chosen in order to provide a simple starting point to explore the virtual EMG task. There are no added environmental dynamics, only a redundant readout of force, and the signal is processed online in the same manner as in the offline analyses. This task or a variant we think can serve as a foundation upon which we can build more complex mappings and virtual environments. Task space trajectories from each block are shown in ???. The fraction of trials resulting in hold timeouts, reach timeouts, and target hits are shown over blocks in ??.

## 4.8 Data Validity

Show that we can reproduce the data offline

Compare calibration to movement in terms of variance, rank, some metric of dispersion in EMG space. How uniformly can subjects reach their EMG space? E.g. compared to a gaussian with the same statistics? The calibration task defines the decoder, and thus it is worth confirming that it is successful in encouraging subjects to explore their EMG space. We can compare their activity to the natural movement datasets as a baseline.

Hypothesis: Variance in the calibration datasets is higher than in the natural movement dataset, even taking all natural movements together. This implies that the calibration dataset encourages exploration of a wider EMG space than that of natural movement.

We used NMF to produce decoders from one of the calibration task datasets.

If we run NMF on the second calibration task dataset and compare the results to what was used in the task, are the results similar? This would confirm that the EMG space was adequately captured. We can define a distance metric between the two decoders and test if this correlates with performance. Differences in the decoders implies that our experimental decoder did not capture the EMG space adequately enough to provide a reasonable learning challenge. Hypothesis: we will see similar results from the second NMF run, which will reject the hypothesis that our decoder does not well-represent subjects' EMG movement space. Hypothesis: inter-decoder distance within subjects will correlate negatively with performance (different decoders, less performant). This implies that the decoder's capturing of whole EMG space, opposed to a subspace, enables subjects to perform better on the task. We expect a crossover point here— you want an EMG search space large enough to encompass the task itself, but small enough not to provide too great of a search task for subjects.

## 4.9 Activity Filter

Explain log filtering here?

Work back through the activity filter thing to make sure it makes sense statistically and we can defend it

Norm of EMG → histogram per trial → look at the distribution of this, should we log-transform it into a rough gaussian? → what is the variance cutoff we should use here? Mean  $\pm$   $N\sigma$ , what is  $N$  to reject outliers? What do we do if the data isn't gaussian even after log transforming?

Once we have our cutoff, we mask emg leaving only samples that fit within our defined activity window

Visualize norm histograms to check for outliers, weirdness

Explain the filtering process, both the data and the activity filtering. Talk about the log-normal transformation.

Remove offsets from EMG – show test Capture “active” parts of EMG signal based on spatial norm – show test compared densities of movement before and after

filtering? Remove artifacts and outliers from EMG – Mahalanobis compared to active mean – show test Use these “masks” on subsequent analyses



# **Notes**



## Chapter 5

# Basic Results

*Perception is not something that happens to us, or in us, [...] It is something we do.*

— Alva Noë, Perception in Action

*To view skilled performance as being the product of underlying component processes is to see a learning curve as a macrocosm of many individual learning experiences. Performance at one point in time reflects what has been learned at some previous time that is able to impact on performance at the moment. Thus, although a learning curve may be viewed as reflecting an incremental improvement process that leads to a smooth transition from novice to expert performance, it is actually a summary of the operation of a vast number of component processes, each with their own improvement functions, and each with varying histories of application with or without success.*

— Speelman & Kirsner, Beyond the Learning Curve

*Because some tasks require that certain components be mastered before other components can be performed or even attempted, there is likely to be a typical developmental sequence in acquiring certain skills. Furthermore, if substantial work is required for component processes to reach a level of mastery before higher level tasks can be attempted, then there is likely to be some period of consolidation in the acquisition of the higher level skill. That is, there will be periods in which there appears to be little progress being made in performance of the task. Closer scrutiny, however, may reveal that performance is improving, but only on lower level components of the task. According*

*to this view of skill acquisition, then, the processes that underlie plateaus in learning curves may also underlie the stages that are characteristic of cognitive development*

Piaget, 1953

*Learning is not like a coin, which remains physically whole even through the most infamous transactions; it is, rather, like a very handsome dress, which is worn out through use and ostentation. —*

Umberto Eco, The Name of the Rose

## 5.1 Does Subject Performance Improve over Trials?

Are subjects actually learning the task? Hypothesis: Yes! Across various measures of performance, the majority of subjects learn to achieve the task's goals in the allotted time period.

Is the task difficult enough to pose a learning challenge for subjects? Hypothesis: yes, some subjects struggle to learn the task, some subjects learn very quickly, while the majority of subjects are between these two extremes. Learning curves are “stretched” across the allotted learning period.

Is there variability across subjects in terms of performance? Hypothesis — yes! And now we want to understand what correlates with particular subject performance

we see an interaction between learning systems – cortex, cerebellum, basal ganglia – strategy / unsupervised prediction / supervised, RL / policies? this leads to multiple learning rates, fast and slow learning, discontinuities in learning

skill learning will lead to hierarchy, and different levels of the hierarchy may be learned at different times by different means/systems

what do our learning curves look like? plateaus? discontinuities?

what you're learning is dimensions of goal-relevance. as these dimensions are learned, we see changes in variability and in performance

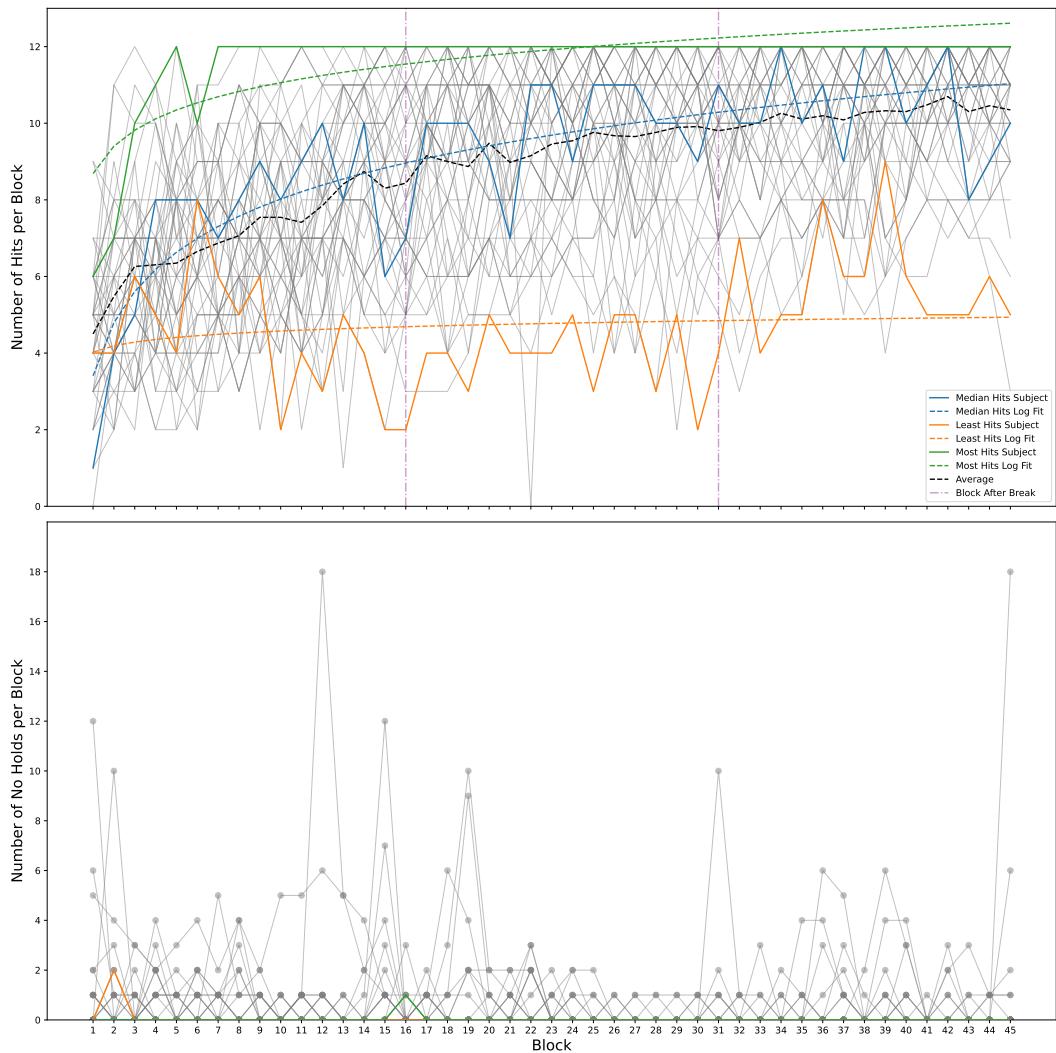


Figure 5.1: Task outcomes over blocks. Outcomes across all 45 blocks of 12 trials (targets) in the “center hold, reach out” task. Subjects with the most, median, and least hits are shown in green, blue, and orange respectively. All other subjects are show in gray. From top to bottom: The percentage of “Hits” within each block, the percentage of “Misses” (timeouts), and “No Holds” (subject unable to quiet forearm muscle activity initially).

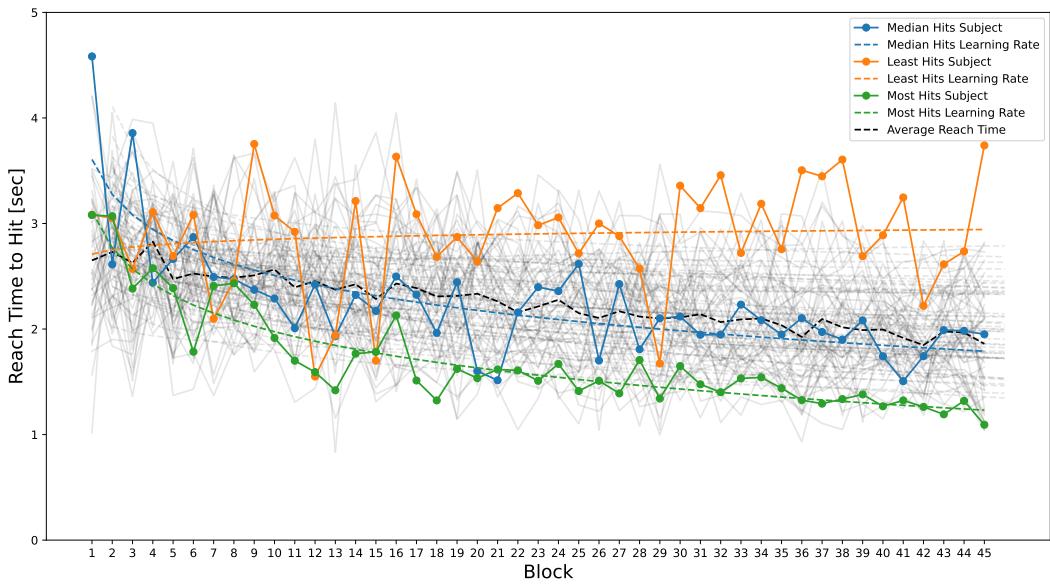


Figure 5.2

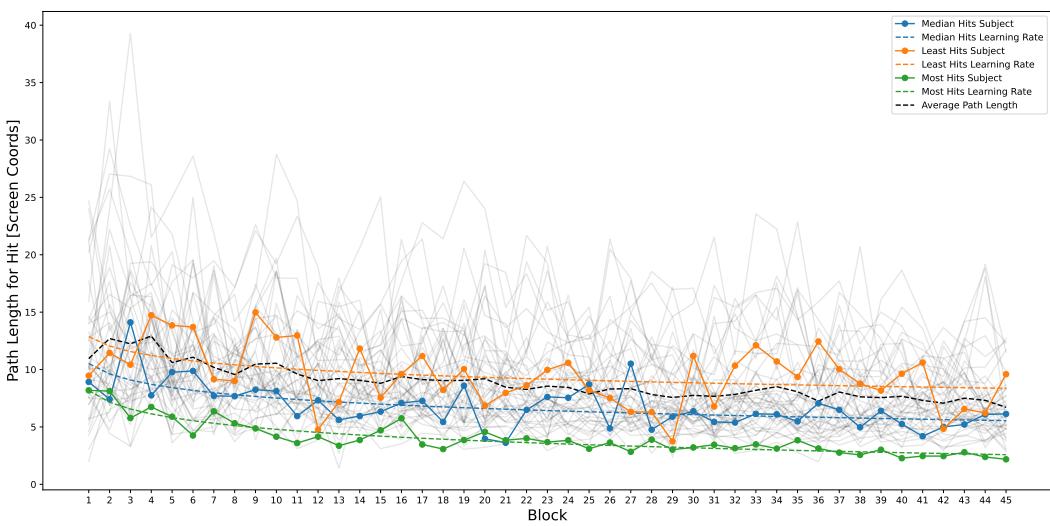


Figure 5.3: `return np.sum (np.sqrt (np.sum (np.diff (a, axis=0)**2, axis=1)))`

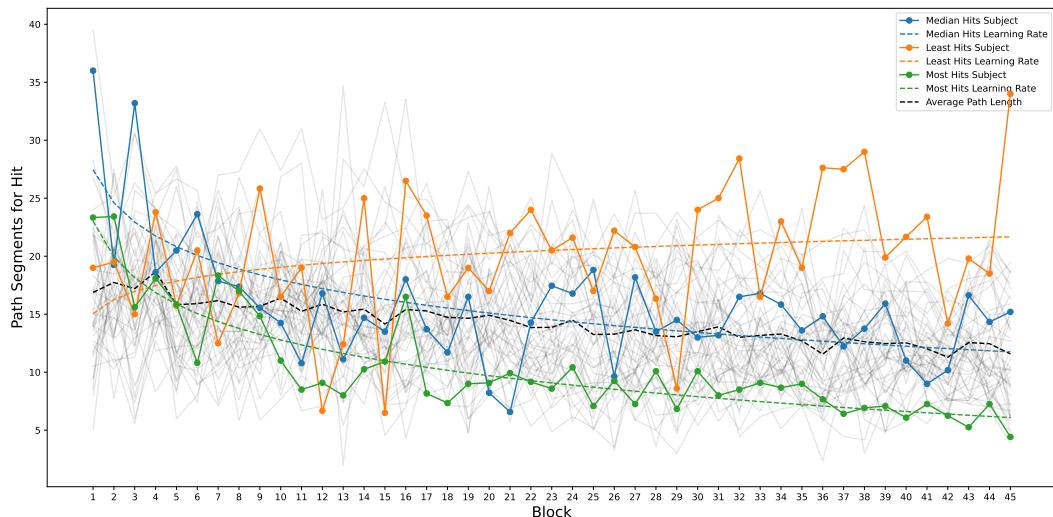


Figure 5.4: This connects to the idea of freezing degrees of freedom. We expect to see degrees of freedom in the task space, quantified by segments, to decrease over trials.

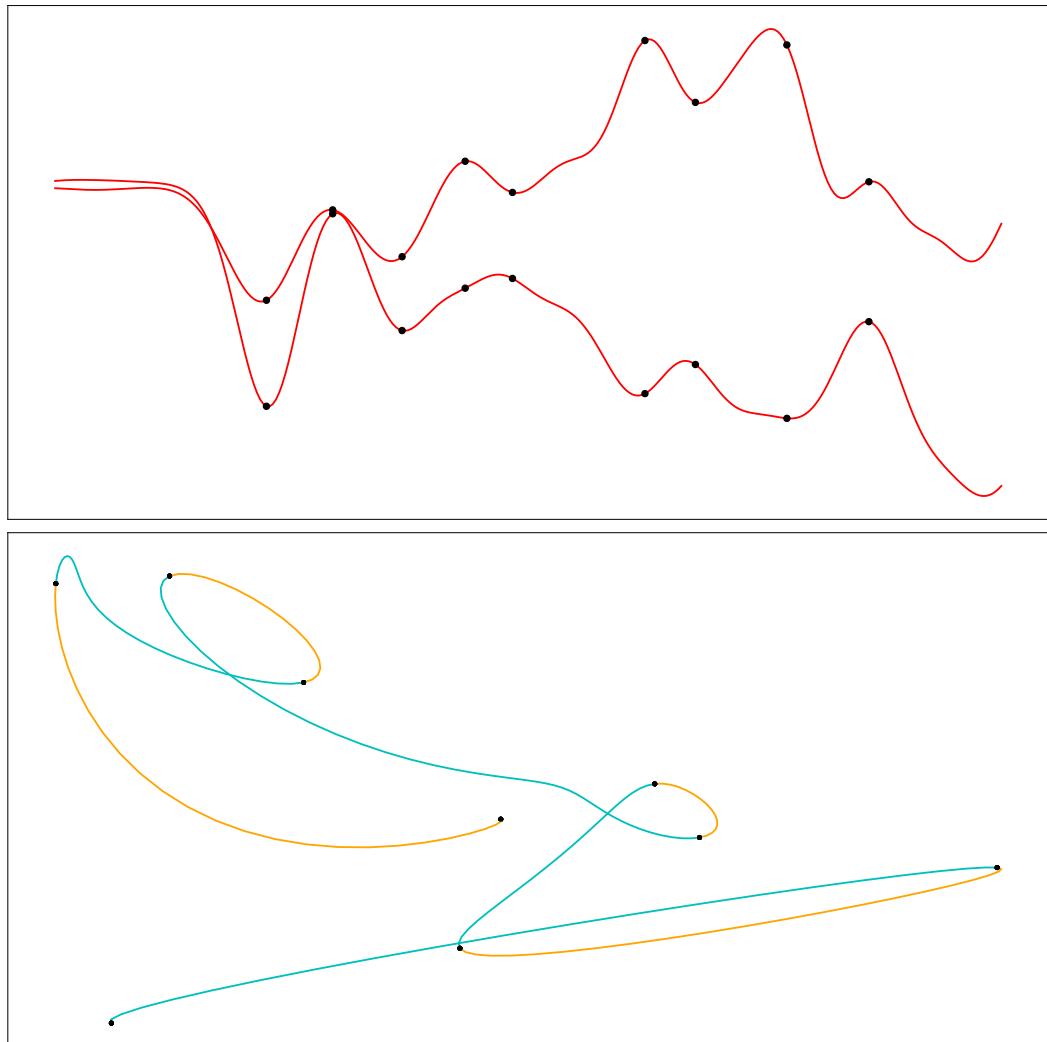


Figure 5.5: zero crossings are diff (sign (diff (x))), then filter crossings that are within a time interval and above a minimum velocity as a percentage of the max

Reward is defined over  $K$  trials and samples or timepoints  $T$  as:

$$r = \frac{1}{KT} \sum_k^K \sum_t^T \|\vec{g}_k - \vec{x}_{t,k}\|^2. \quad (5.1)$$

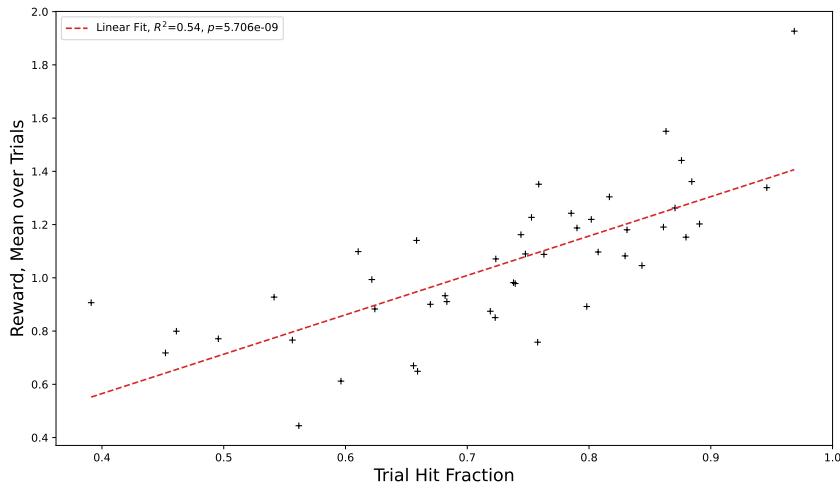


Figure 5.6: Fractions of trial outcome types for each block of the center-hold, reach-out task. Hit fraction increases on each trial, suggesting the beginnings of task learning. Hold timeout failure increases over trials as well, perhaps suggesting increased baseline excitation of muscles during the hold period. Part of learning to activate certain muscle modes is learning to inhibit others.

Eq. (5.1)

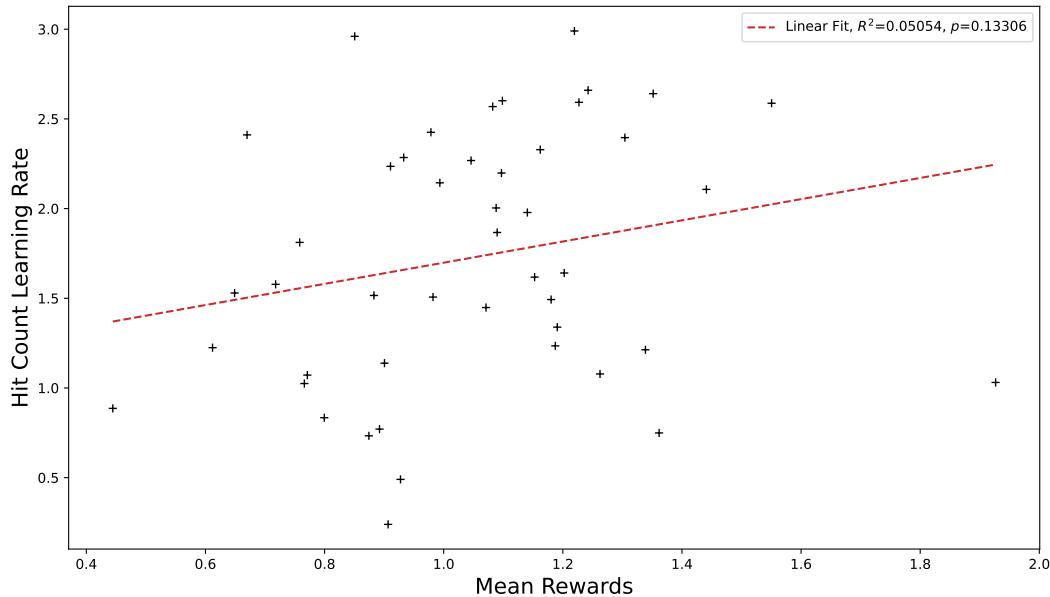


Figure 5.7: No correlation!

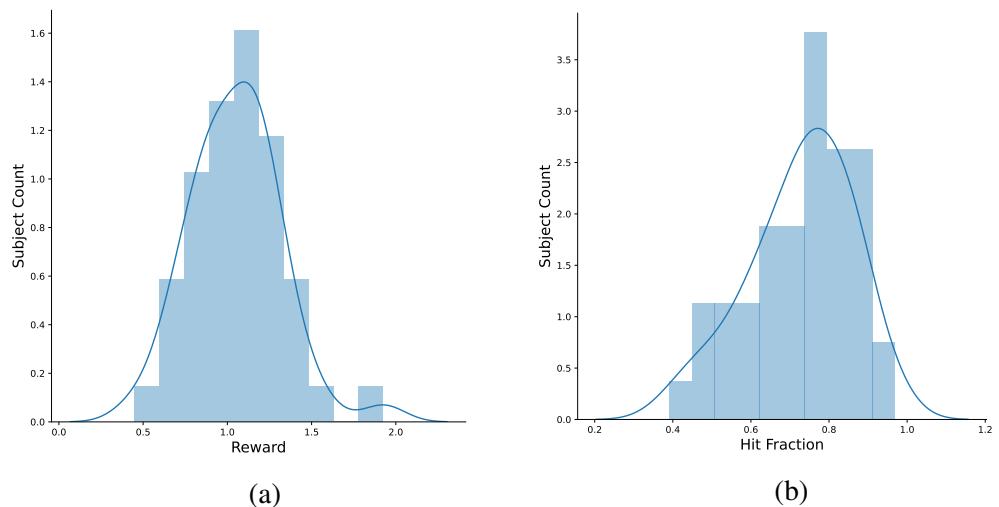


Figure 5.8: Hits and rewards histograms over subjects

Table 5.1: Statistics of performance regression

$p(\text{Hit Learning Rate, Reward})$	0.177
$p(\text{Reach time Learning Rate, Reward})$	0.215
$p(\text{Path length Learning Rate, Reward})$	0.786
$p(\text{Segment Learning Rate, Reward})$	0.072
Adjusted R-squared	0.058
Prob (F-statistic)	0.170

### 5.1.1 Hits over Targets

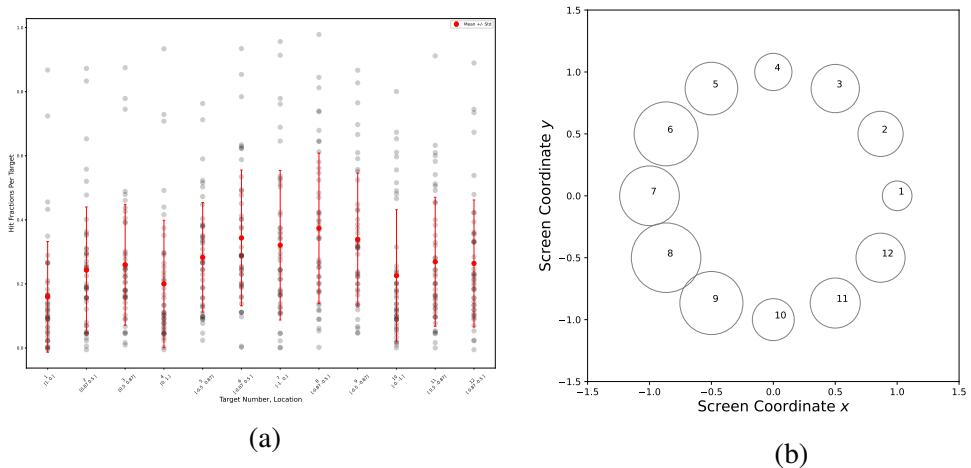


Figure 5.9: Hits over targets

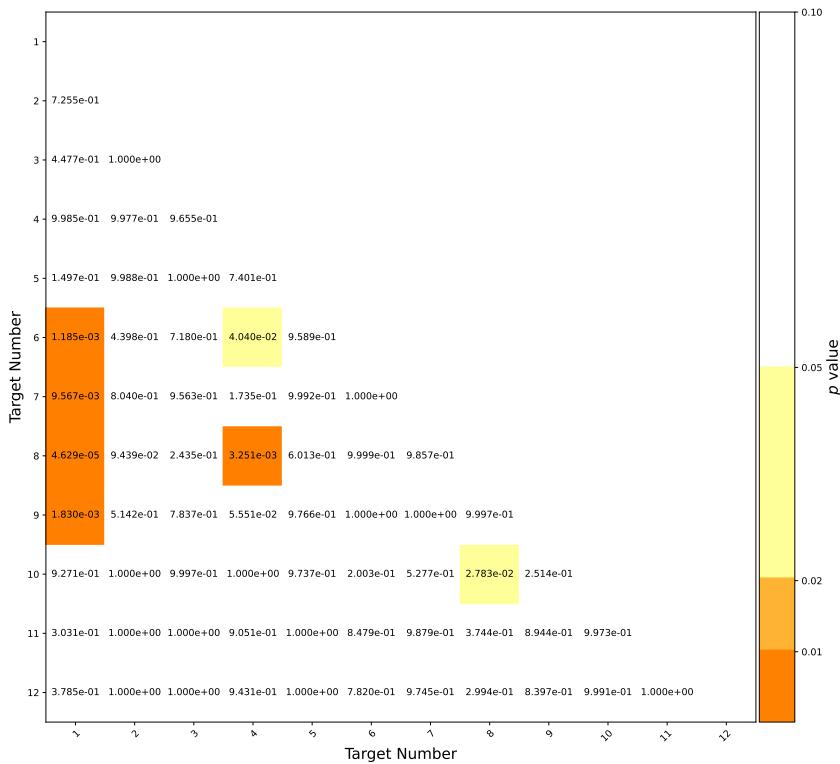


Figure 5.10: ??

Figure 5.10 Figure 5.10

## 5.2 Do Subject Attributes Predict Performance?

Look at self-reported form response data (this should show that there aren't obvious confounds, as far as we can tell?) Arm size Time of day Time order Max "force" – sum of the channel variances? Caffeine vs. non Instrument players vs. non

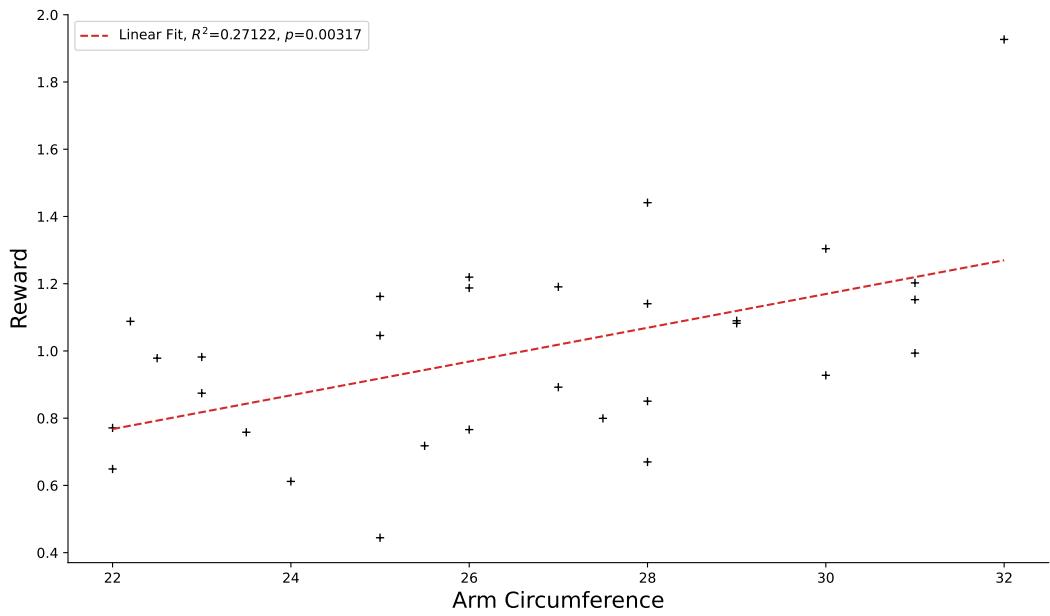


Figure 5.11

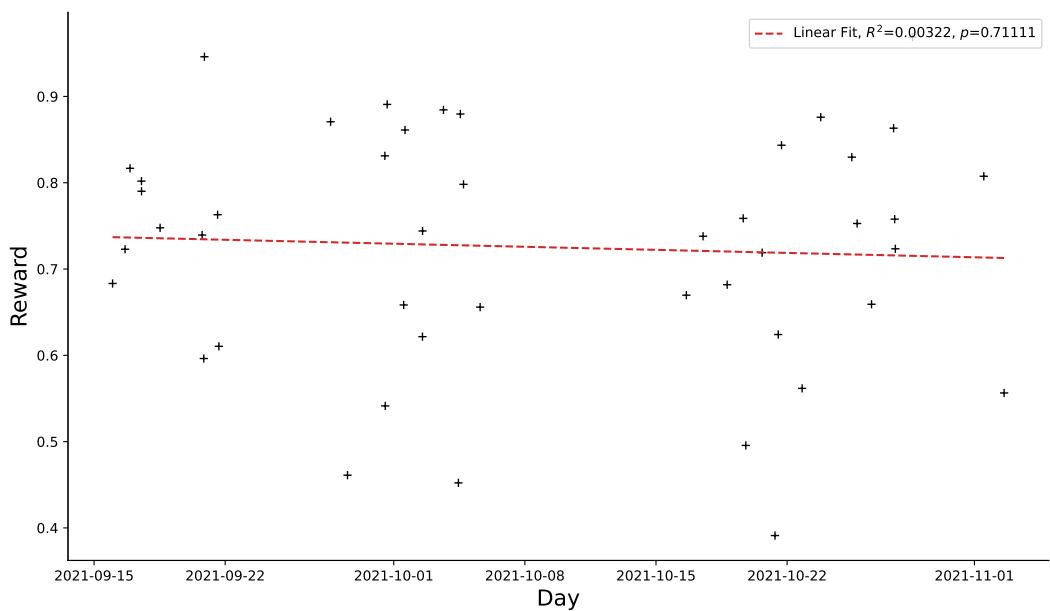


Figure 5.12

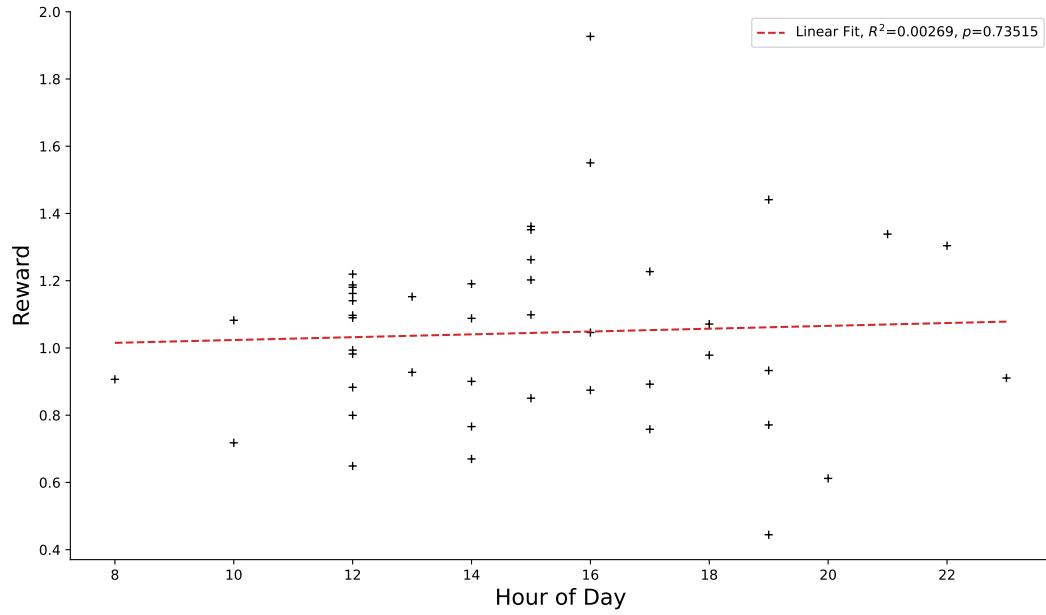


Figure 5.13

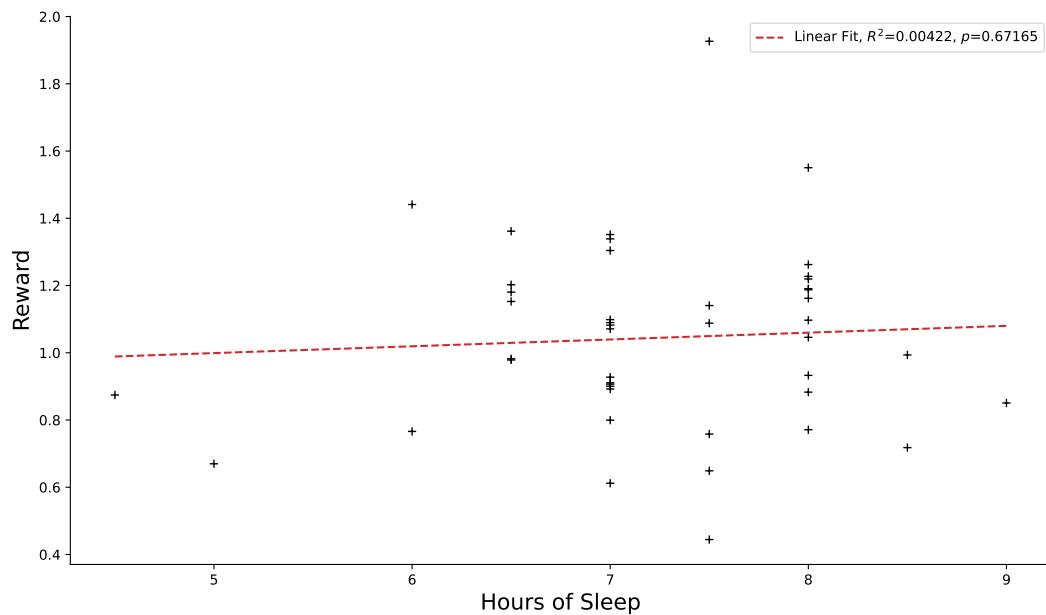


Figure 5.14

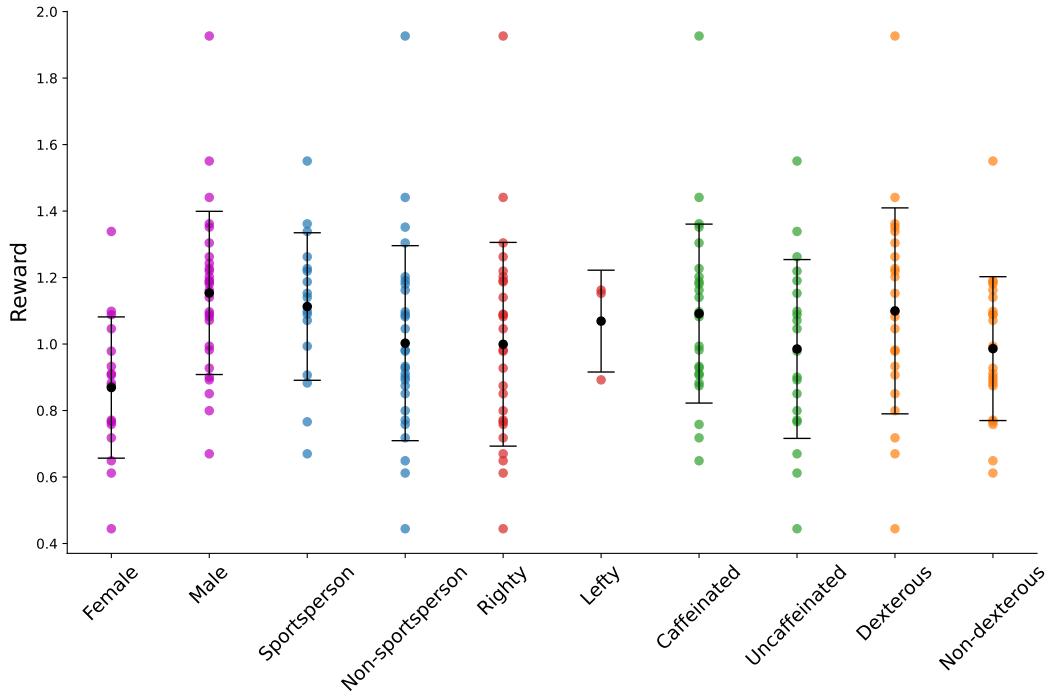


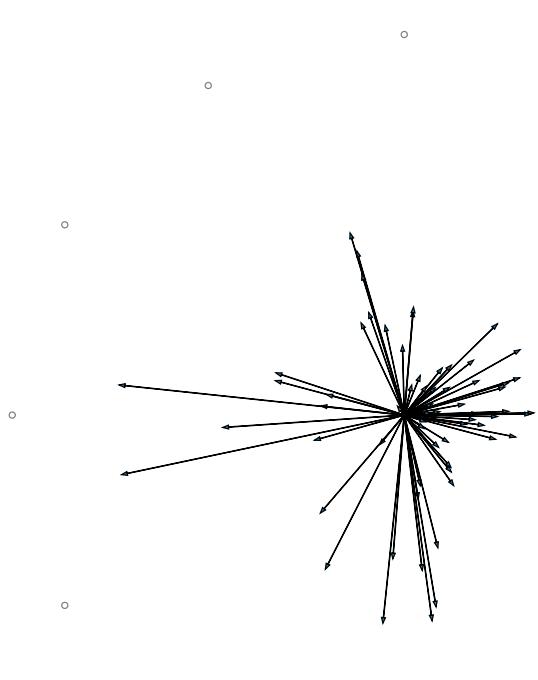
Figure 5.15

### 5.3 Does the Decoder Predict Performance?

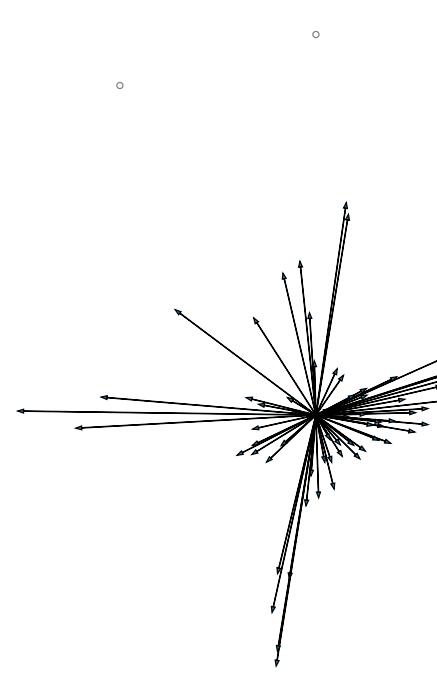
Digging into the decoder

Do the directions of the decoder correlate with performance? Hypothesis: yes, decoder directions which happen to be oriented towards targets will make the learning problem easier, and thus correlate positively with performance.

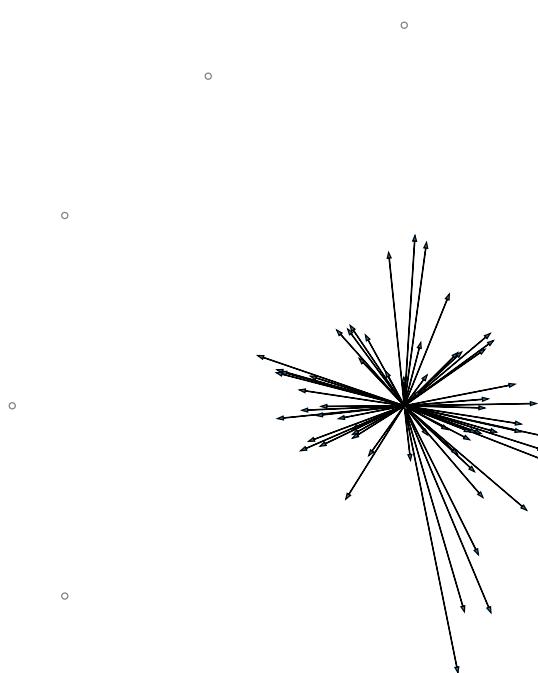
If the decoder directions are more/less aligned to the target directions, do subjects explore less/more? Is exploration (e.g. EMG variance) higher in subjects with misaligned decoders? Hypothesis: exploration will be lower in subjects with target-aligned decoders. Subjects make fewer incorrect movements, and thus are induced to explore less.



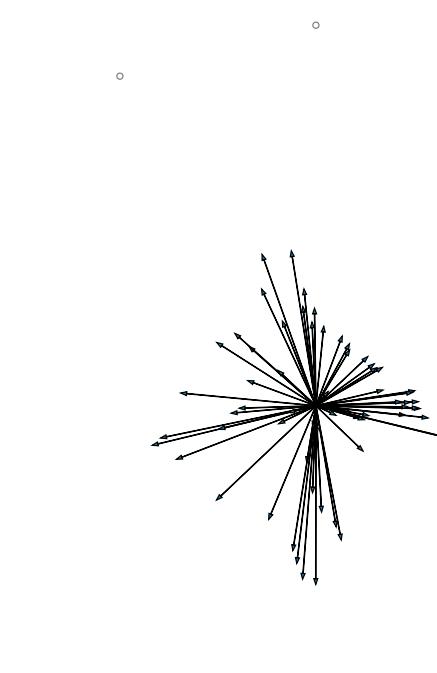
Subject 4



Subject 12



Subject 28



Subject 37

Figure 5.16: Example decoders directions in 2D.  
106

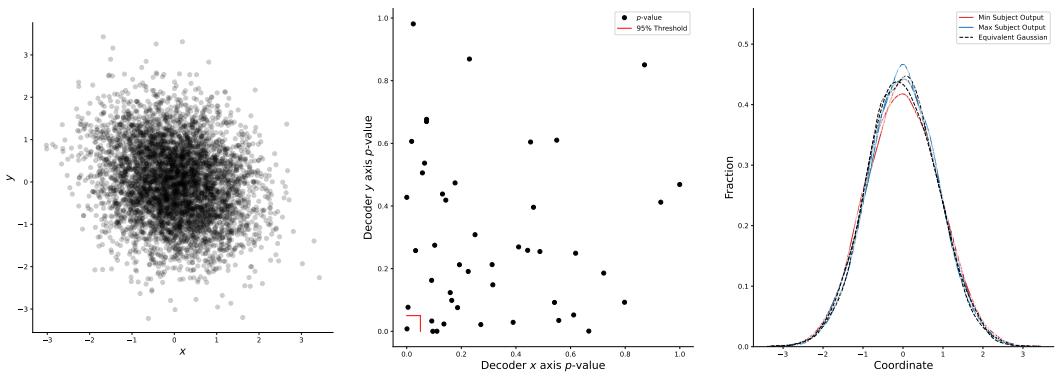


Figure 5.17: Testing the effect of the decoder on uniform data.

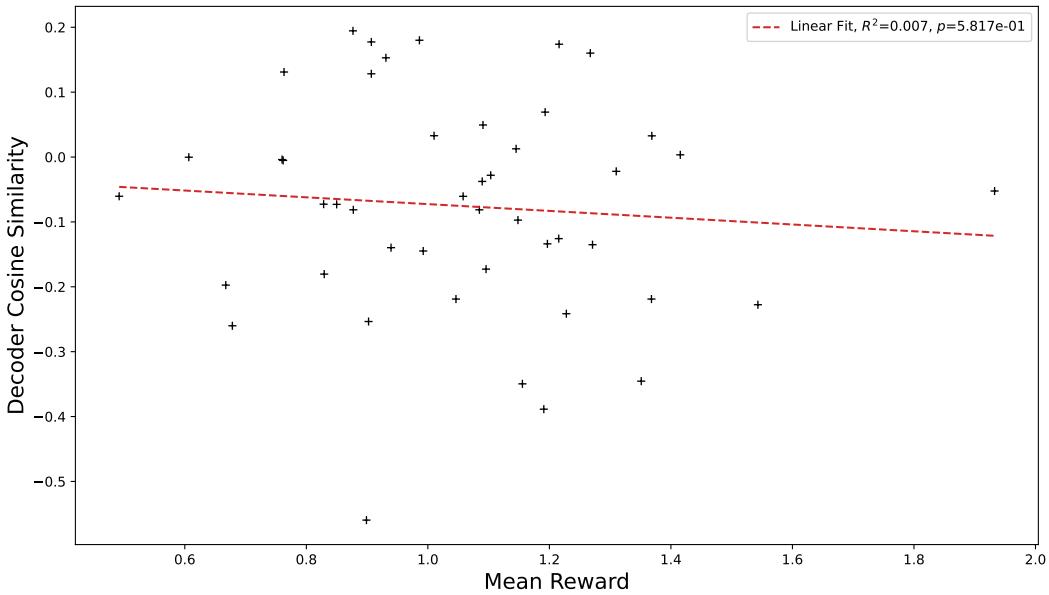


Figure 5.18: Decoder cosine similarity. EMG-to-force decoders are computed using a calibration dataset collected before the “center hold, reach out” task. 4 “modes” of EMG activity are extracted from the dataset using non-negative matrix factorization. These 4 modes are then subtracted in pairs to yield the two 64-dimensional EMG-to-force mappings, shown in `{+@fig:low_variance_PCs}`. The left plot shows the cosine similarity of the  $x$  and  $y$  EMG-to-force decoders. A cosine similarity of 1 means the two vectors are parallel, producing identical forces (in the respective directions) for the same EMG activity ( $F_x = F_y$ ). A cosine similarity of -1 means the vectors are antiparallel, producing equal but opposite forces in the two directions ( $F_x = -F_y$ ). A cosine similarity of 0 means the decoder directions are orthogonal; e.g. producing a force in the  $x$  direction with a certain EMG activity produces no force in the  $y$  direction. Plotted across subjects, we see a range of decoder similarities, providing a variety of task contingencies. The rightmost plot asks whether cosine similarity is predictive of task success, in terms of the numbers of “Hits”. We find no significant correlation, implying that the decoder cosine similarity, in the range we tested, does not predict task success. Task success, therefore, likely relies on an alternative task variable.

## 5.4 Does Task Space Variance Decrease over Trials?

Trajectory viz, discussion, stats (task space only)  
 Look at histograms over trials for x and y – where are people moving over time?  
 Look at trajectory means, trajectory distros  
 Show 2D histograms, mean trajectories  
 What must be happening in the EMG space to get these trajectories? This is all “on manifold” – what’s happening off-manifold?  
 Histos to show if any subjects are outliers that we should ignore  
 Who are

the “best” subjects in terms of trajectories? Note these as a subgroup that we can zoom in on later.

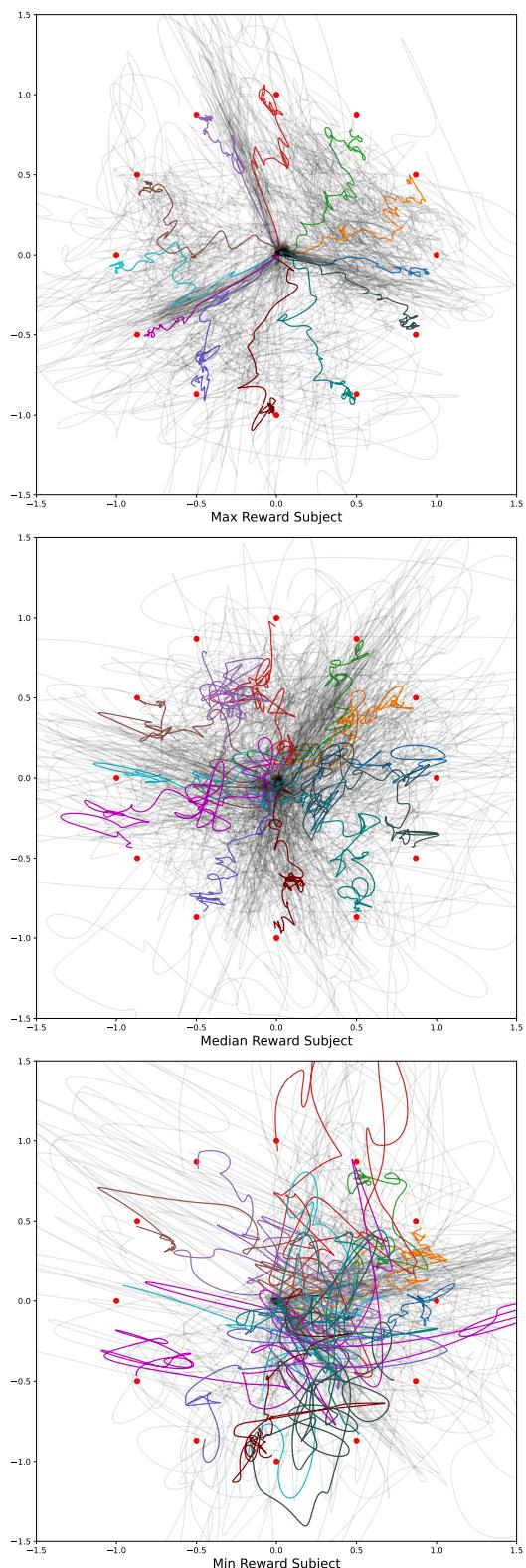


Figure 5.19: ??

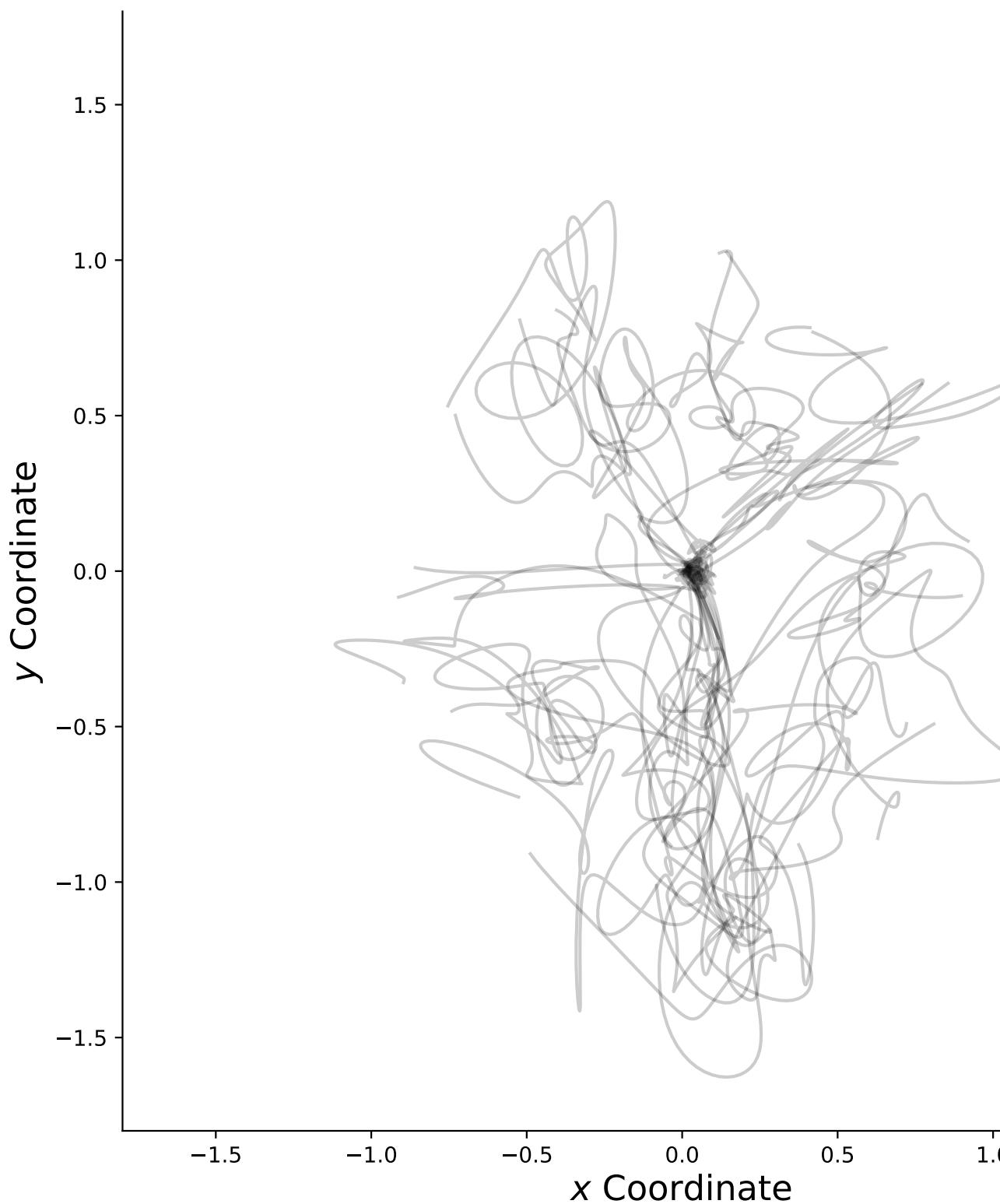


Figure 5.20: ??

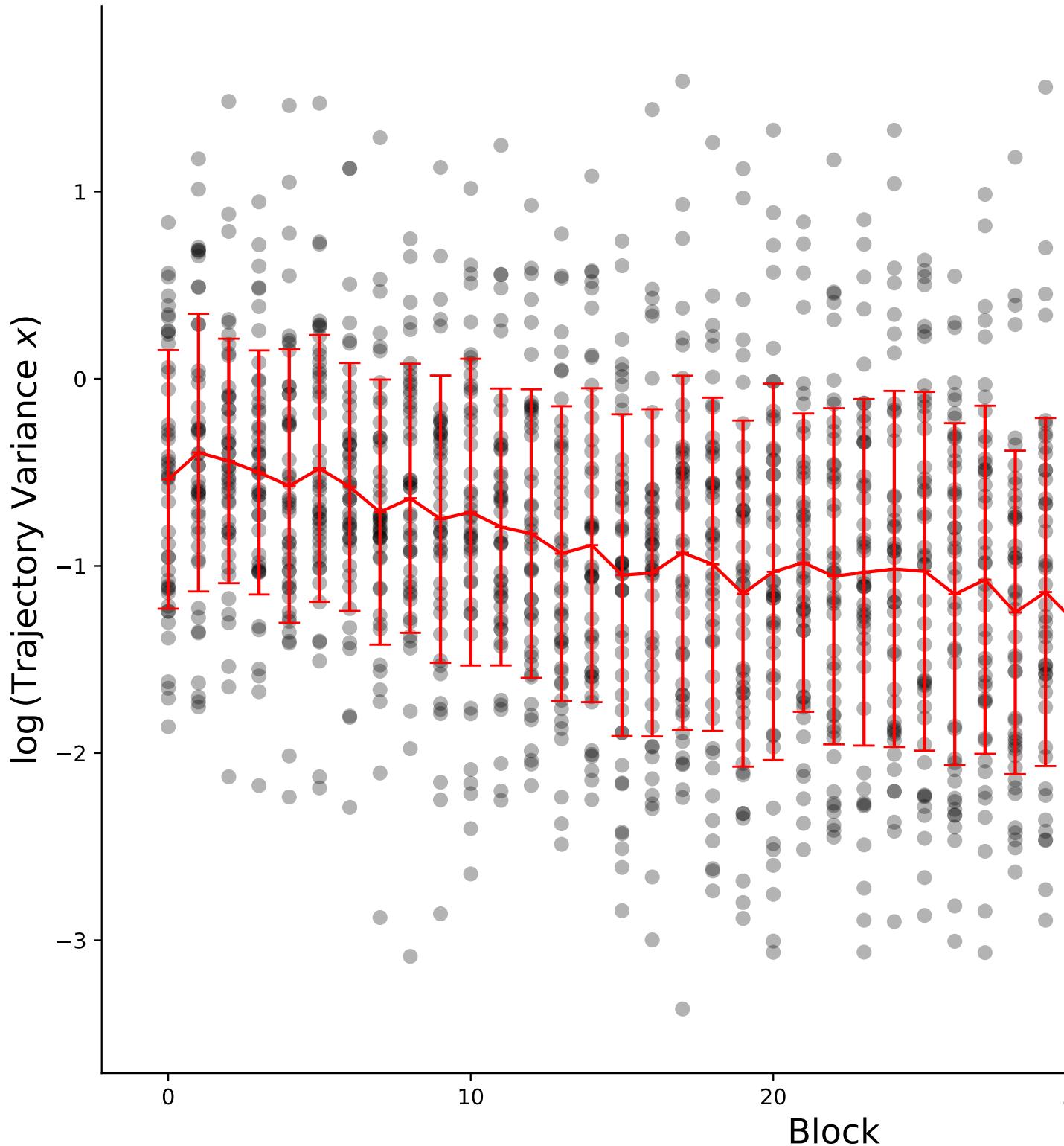


Figure 5.21: Plotting the log as variance tends to decrease exponentially

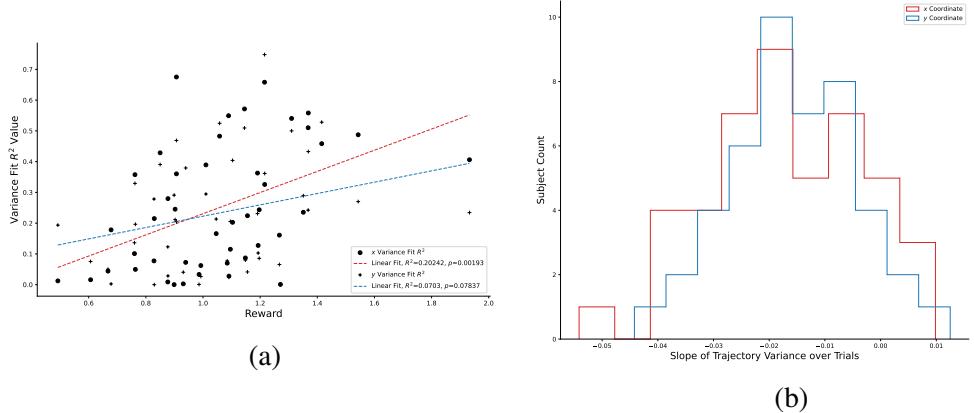


Figure 5.22: Trajectory variance statistics



# **Notes**



## Chapter 6

# More Results

*My view is that language and the hand have a certain common agenda—that is, they enable us to **grasp** things: to pin them down and make them useful. And we cannot deny that they have done that in spades. They have helped us to use the world and, by doing so, to develop many of the things of which we are most justly proud, the fruits of civilization.*

— Iain McGilchrist, *Ways of Attending*, 2016



## 6.1 Structure of the Data Manifold

### 6.1.1 Is EMG Lognormal?

Brief intro into log-normality of the data

Also show correlations between channels and introduce the idea of subspace confinement

Motivate the shape of the manifold discussion next

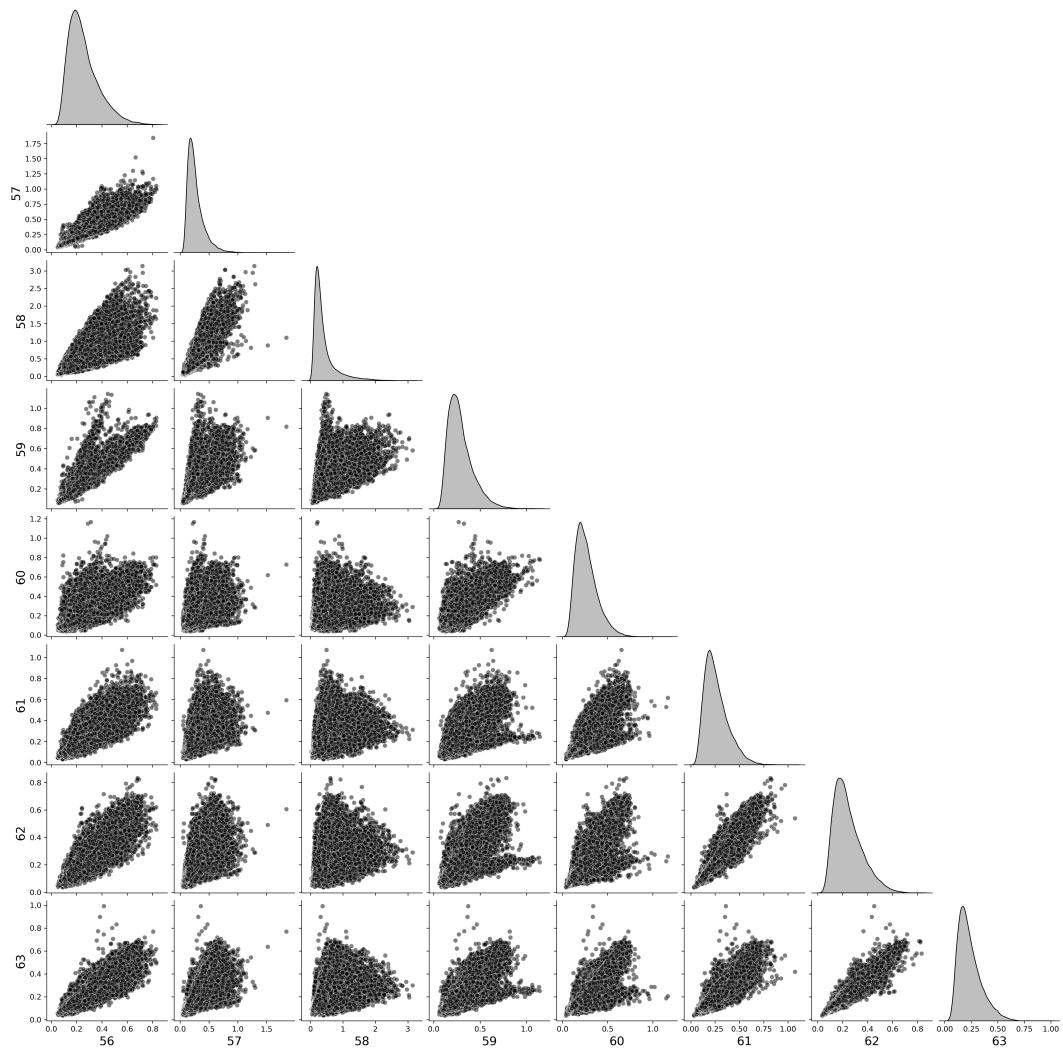


Figure 6.1: Pairplot of calibration data

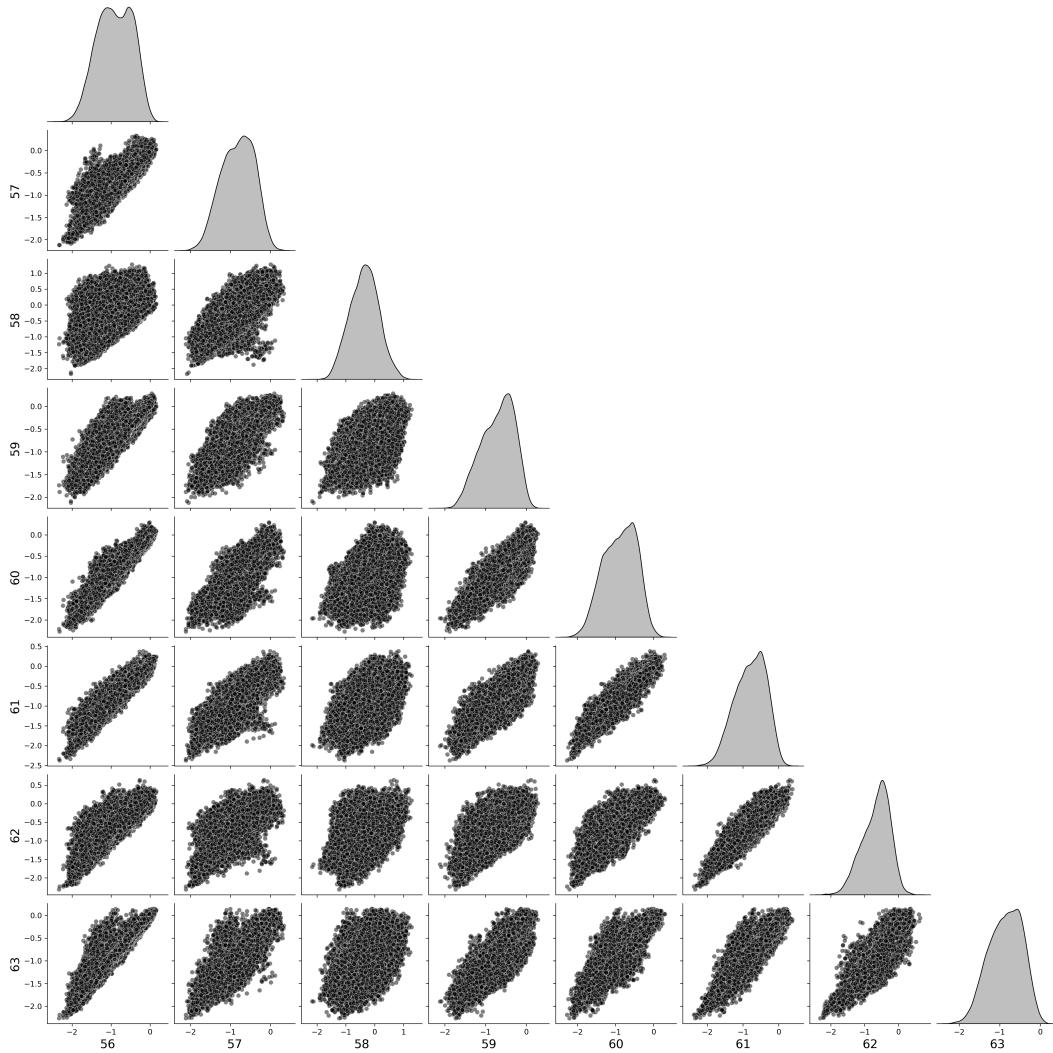


Figure 6.2: Pairplot of log-transformed calibration data

Structure of data manifold → multimodal, spiky, etc. means we have to be careful about statistics Motivate PCA vs mixture models Show PCA and mixture model fits Do we see shifts over learning under different models? Fit models across learning, use the stats to show shifts? Entropy of GMM model fits over trials? “Policy entropy”? (Bits of information contained in the policy? [0,1] Does a subject penalize changes to their controllers/action distribution? Do they follow a KL-divergence type of measurement when improving their policy? Do we see this in the dynamics of the statistics over trials?

#### “SUBSPACE CONFINEMENT”

The fits enable you to overcome the subspace confinement, give us a reason-

ably tractable model of our multimodal data. Weve proven that the data lies in distinct, low-dimensional subspaces of the activity space. The GMMs describe this fact, particularly the rank of the variances. Explain what these fits are, visualize comparisons to data, prove that theyre reasonable Motivate the GMM fitting with the overview of the data as a model to capture stats !!! you cant recover the subspace-confinement if you look at the combined covariance (if subspaces overlap) – this means analyzing bulk covariances will hide subspace confinement in the data. This could be happening even at the intra-trial level Explain how the data is very sparse! Each trajectory is very low-probability. The trajectories are very thin, subjects repeat movements which are common, this is whats captured by the GMM

How many discrete movements make up a trajectory? How do we define a discrete movement? Can we define a segmentation for movements, and use this as a measure of “smoothness” of control? Hypothesis: movement modularity or segmentation decreases over learning as subjects discover smoother solutions to the task. These movement segments are part of the subjects natural movement repertoire. Do subjects freeze degrees of freedom to simplify the learning task? (E.g. How many PCA components are active at once, how does this change over time?) Hypothesis: early in learning subjects will activate only a subset of available movement dimensions as defined by task-space movement repertoire, PCA or similar model of variability. If we compute PCA on each trial, does the dimensionality increase, decrease, or something else over trials? The answer points to exploration vs. exploitation, hierarchical or mixed-controller solutions. This may highlight how the structure of variability develops over trials. Hypothesis: Early in learning, subjects will explore more, and their dimensionality will be high (similar to the dimensionality of natural movement or the calibration task). Over learning, dimensionality will decrease as subjects “hone in” on solutions to the task. Dig into subjects movement policies– what structure is there? Are subjects choosing from a repertoire of policies (e.g. learning one policy for each target, a kind of hierarchical policy, or learning the decoder directly to flexibly produce movement solutions?) Hypothesis: subjects are doing everything all at once– theyre internalizing the decoder by

learning an internal model (a mapping from targets to movements) of their new environment, and they're using targets as cues to remember discrete movements. This is highlighted when subjects make incorrect movements— how they recover from these mistakes will highlight the difference between an internal model and model-free learning. Hypothesis: Early in learning, we'll see subjects repeat many of the same movements as they're repeating what works in a model-free manner. Later in learning, we'll see subjects responding more flexibly, having developed a better internal model of the movement-task contingency. NB: these hypotheses contradict other hypotheses about exploration. If subjects are optimally exploring the space, collecting data about the contingency, they wouldn't be making the same movements repeatedly, they would be making a variety of movements. This might come down to individual subjects' strategies for how they approach the task. We may see some subjects choosing to explore, rather than solve the task. Other subjects may be focused on solving the task and learning through their mistaken movements. Can we take the first movement segment (of each block?) as representing subjects' predictive forward models? Does this give us an indication of subjects' forward models? Or of their discrete policy decision memory? Do these get better over trials? Hypothesis: Early in learning, these "first corrections" will be poor, in the sense that they will often be incorrect and likely default back to some common movement policy (perhaps a movement from the natural movement dataset?). Late in learning, we'll begin to see more flexible, assumedly internal-model driven responses to these "mistakes". Are subjects using one controller or several? Are they composing them or switching between them? Are they learning which controller to switch between or are they developing a single controller to solve for all targets? Does this explain perturbation response? Subjects might be building  $p(\text{best controller} \mid \text{target, state})$ , that is choosing from distinct policies, rather than developing a de novo policy. Hypothesis: Subjects are doing everything at once: they are internalizing a model of the task dynamics and decoder to form a general controller/policy, they are forming or recalling individual controllers/policies for individual targets. We can investigate this by looking within trial trajectories, where we should be able to classify behavior

as policy switching/activation vs. generic controller/policy response.

Dig into the data, what does it look like from different angles? Means in EMG and task space Activity distributions in EMG and task space Covariance structure? Tangent space? Manifold structure – UMAP over time? Colorful plot showing shift over time... Discuss trajectory, data scaling On what scale (trials, timesteps) is the model altered? the policy? → What does the trial-to-trial variance look like? Subtract out the mean trajectory?

If nothing so far really explains subjects performance, including the decoder / task contingency itself, then the constraints imposed on subjects must come from the subjects themselves, from their data manifold. Discuss the geometry of the task; weve shown that aspects of the decoder arent predictive of subjects success—so something else (the data manifold itself!) is responsible for performance. Go deeper into the decoder— were inverting the mapping from latent “activations” to emg to get an emg-to-latent mapping which is norm-minimizing What is the structure of the data manifold? Show/explain the concept of subspace confinement The data isnt Gaussian, so PCA is going to suffer This is a really important point we can make – looking at the data manifold, PCA will give us something but its going to hide a lot of information about subspace confinement, covariance structure, etc. Its essentially going to align the axes towards high-variance directions, but not take into account the “spiky” nature of the data manifold. This will give us a false sense of the correlation structure between EMG channels. In other studies, theyve analyzed surface EMG data with PCA to then say that since there are high-variance components that capture gross hand postures / movements, and lower-variance, higher-order components that capture finer movements, this implies something about the neural basis for hand control. While it may, this result via PCA isnt surprising. What is more insightful is to understand the underlying structure of the data manifold, and follow its changes across learning to understand how the motor system copes with a new environment. We can clearly show how PCA misses an important aspect of the data manifold, its highly nonlinear structure, and how this plays a role in subjects performance and learning. We can highlight this by showing PCA pro-

jections of the data colored to show how they're mixed in terms of their target, etc.  
– not getting good separation of the data because of the non-gaussian manifold

## 6.2 How Does PCA Dimensionality Compare Across Tasks?

Show dimensionality of different tasks

Show two-dimensional projections of activity

Look at these PCA modes, compare them to natural movement...? Are these modes relevant to the task? Hmm....

“Movement” session Compare this to calibration, trial in terms of e.g. entropy to show that subjects are moving beyond their “Calibration” session

Hypothesis: subjects explore beyond their natural movements in the calibration task.

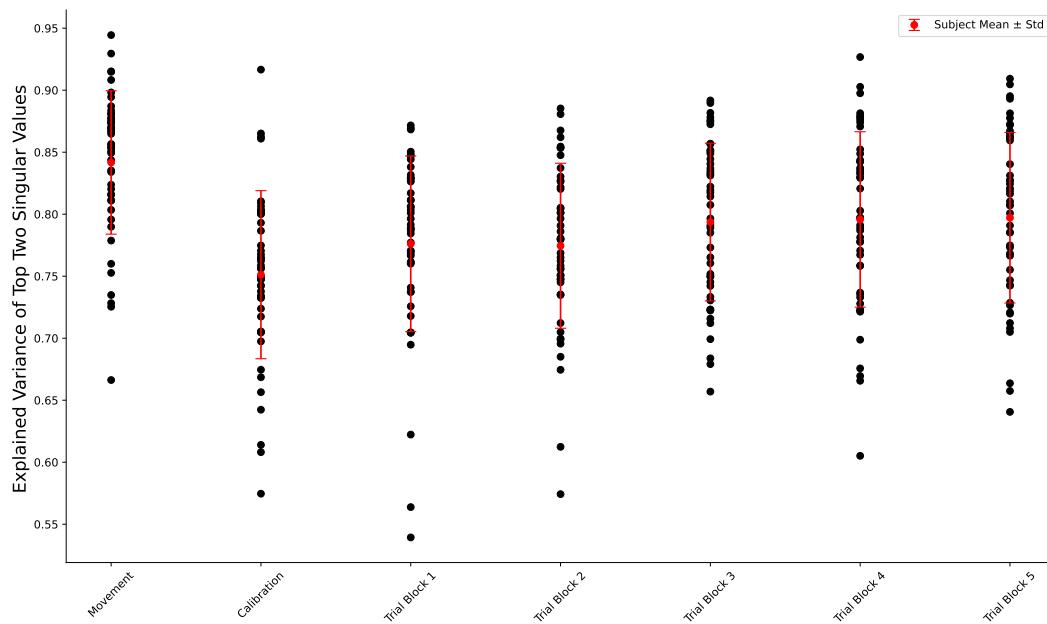


Figure 6.3: No correlation!

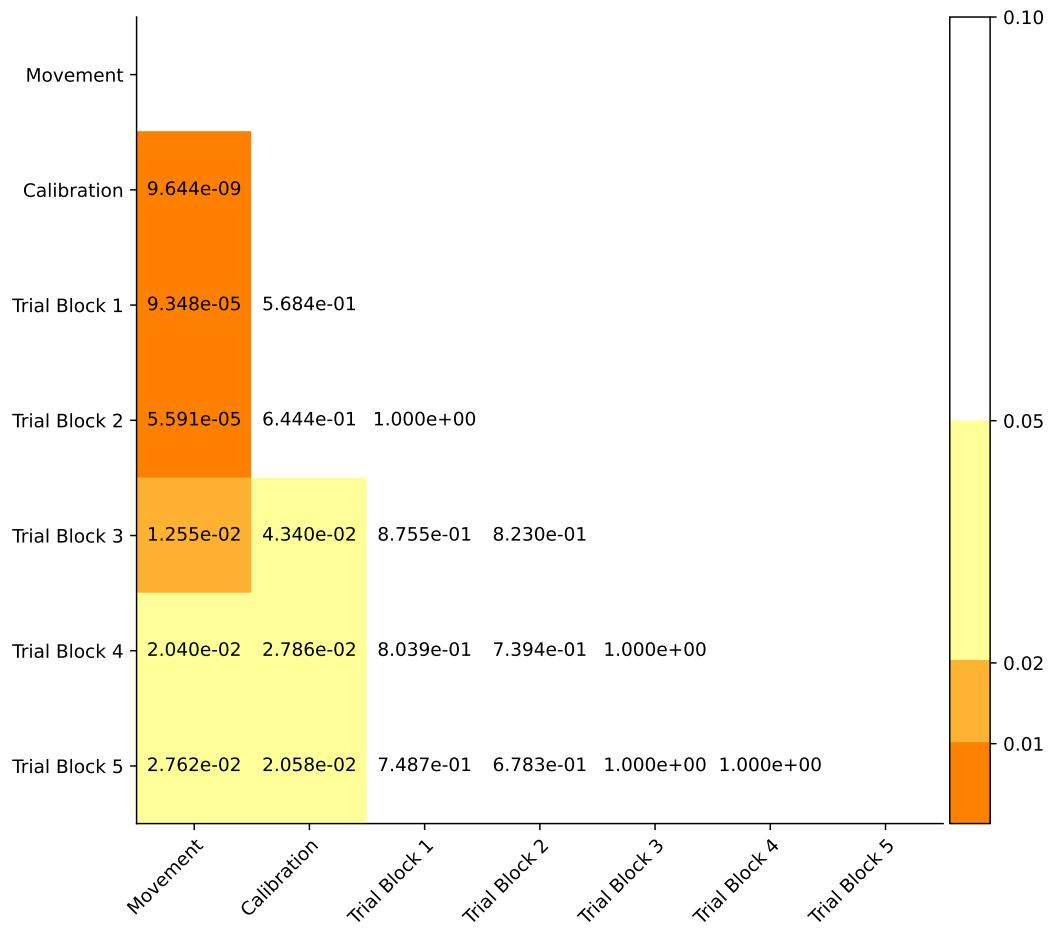


Figure 6.4:  $p$ -values for Tukey’s honestly significant difference test for pairwise means of a group of samples. Here we compare the explained variance of the top two singular values for the movement, calibration, and 5 blocks of task trials over subjects.

Figure 6.4

### 6.2.1 UMAP

UMAP the data to motivate the nonlinear manifold, the “fingers” of the subspaces, and the “honing” idea – this supports the weird data manifold Do a comparison on synthetic data to hammer the point home – compare to 64D Gaussian distribution with matching statistics..

### 6.2.2 Toy GMM model demonstration X-pattern of trajectories

Show how PCA fails to capture these statistics. Motivate the gaussian mixture model, even if these are overfit; we don’t need these models to generalize, we’re

using them for summary statistics, like a filter.

### 6.3 Subspace Confinement

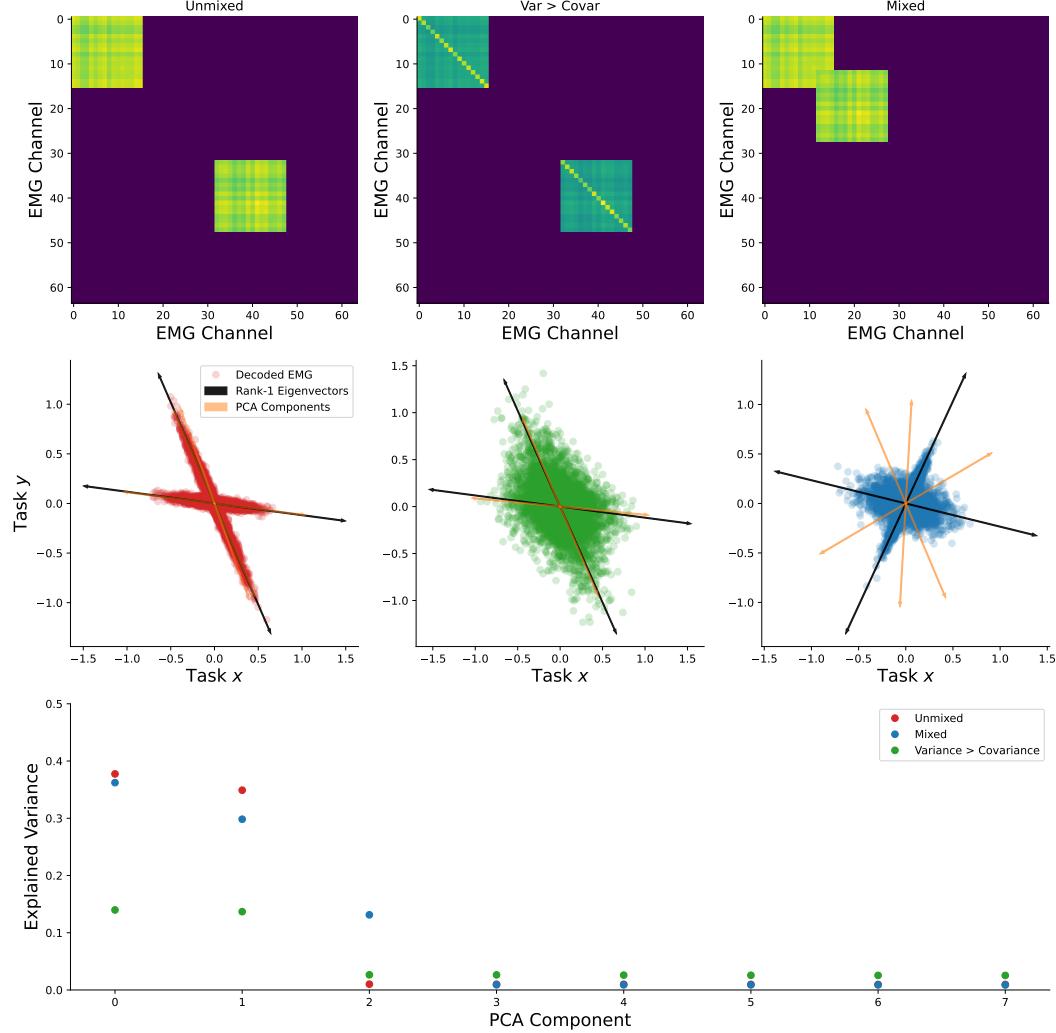


Figure 6.5: The variables here are: the rank (number of rank-1 components making up the mixture), the ratio of the variance to the covariance of those components, and the "mixing" of each rank-1 component which effectively increases the rank of the mixture, if we assume to be a single Gaussian. We can see how as the variance begins to dominate the mixture looks more like a multivariate Gaussian. When components mix (without changing the overall variance to covariance ratio), we have a similar result due, however, to inability for PCA to model such a mixture as a single Gaussian. The 2D trajectory plots are shown for reference, but this effect is happening in the full 64D space, the values of the decoder are not a factor in this effect. See how PCA mischaracterizes the distribution in the mixed case, but fares well in the unmixed and high variance cases. This motivates our use of mixture models, which explicitly attempts to fit multiple Gaussians and can thus deal with mixing. Note that the effect here is slight, but demonstrates the principle of interest.

How to transform gaussians into 2D

$$x \sim N(\mu, \Sigma) \quad y = Ax + b \sim N(A\mu + b, A\Sigma A^T) \quad (6.1)$$

Do subjects who are less “confined” (constrained) are expected to perform better?



## 6.4 Gaussian Mixtures

### 6.4.1 GMMs vs PCA

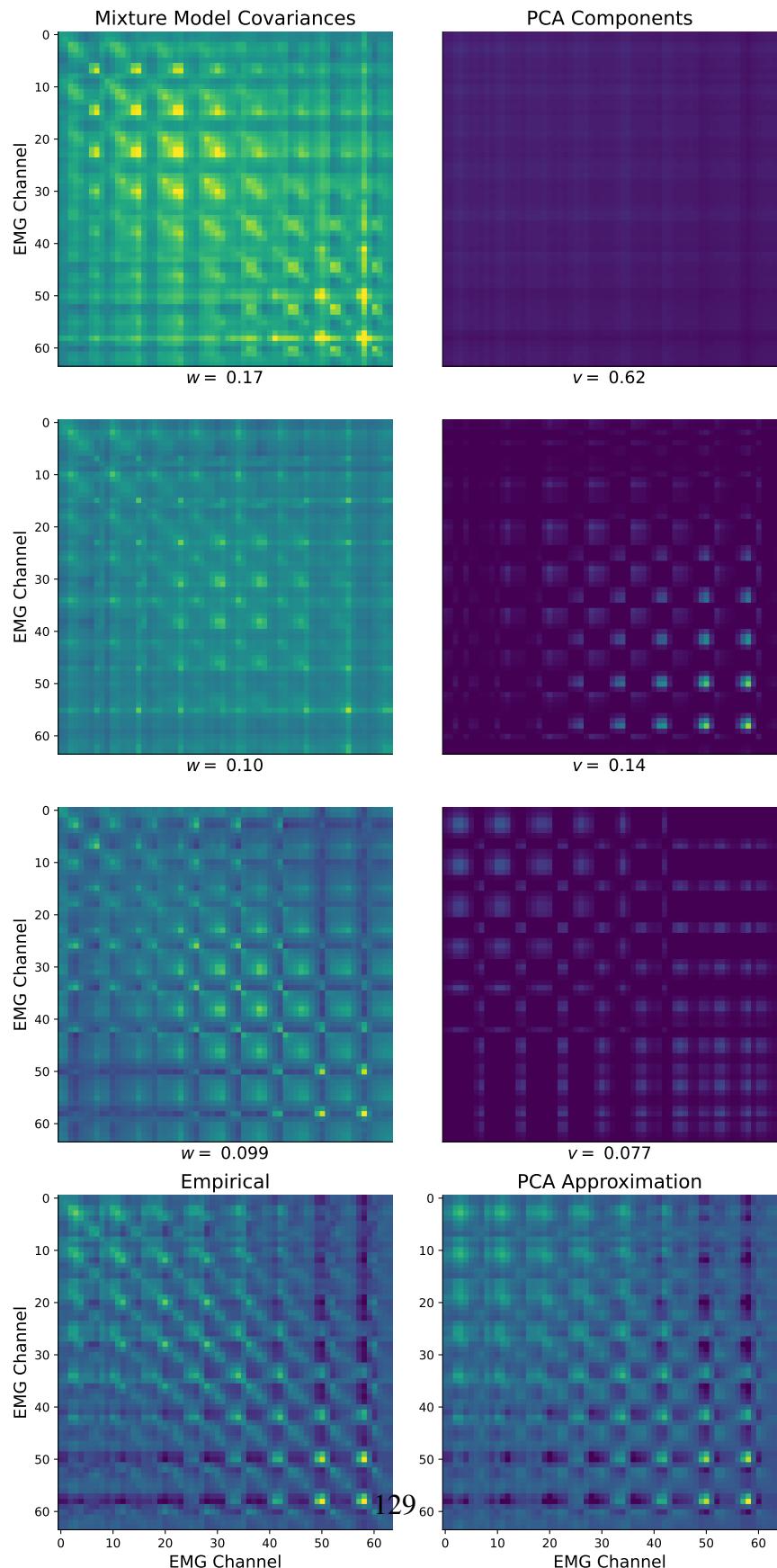
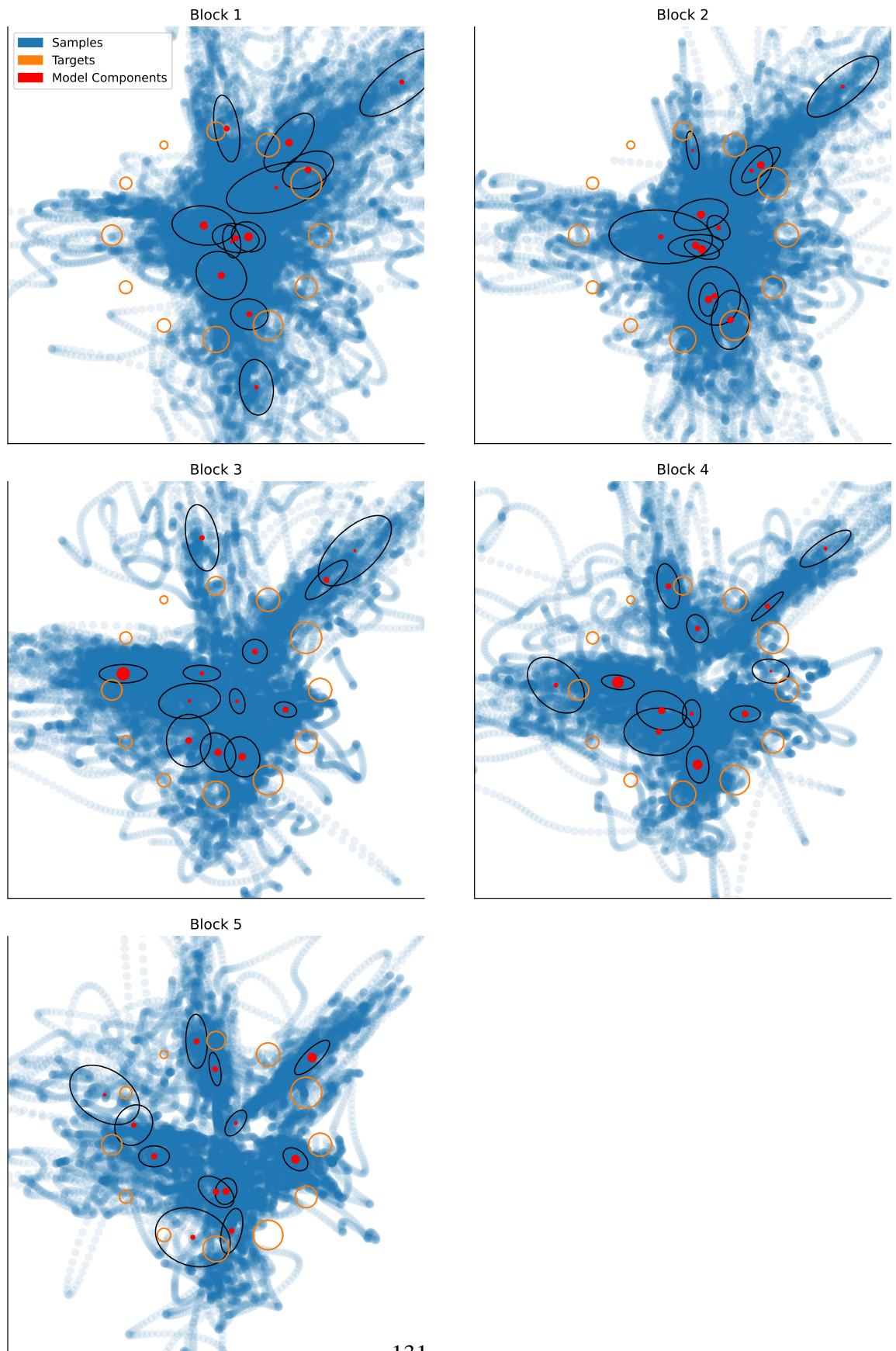


Figure 6.6: Compare the covariances of GMM fits to the covariances of individual PCA modes.



### 6.4.2 Example GMMs



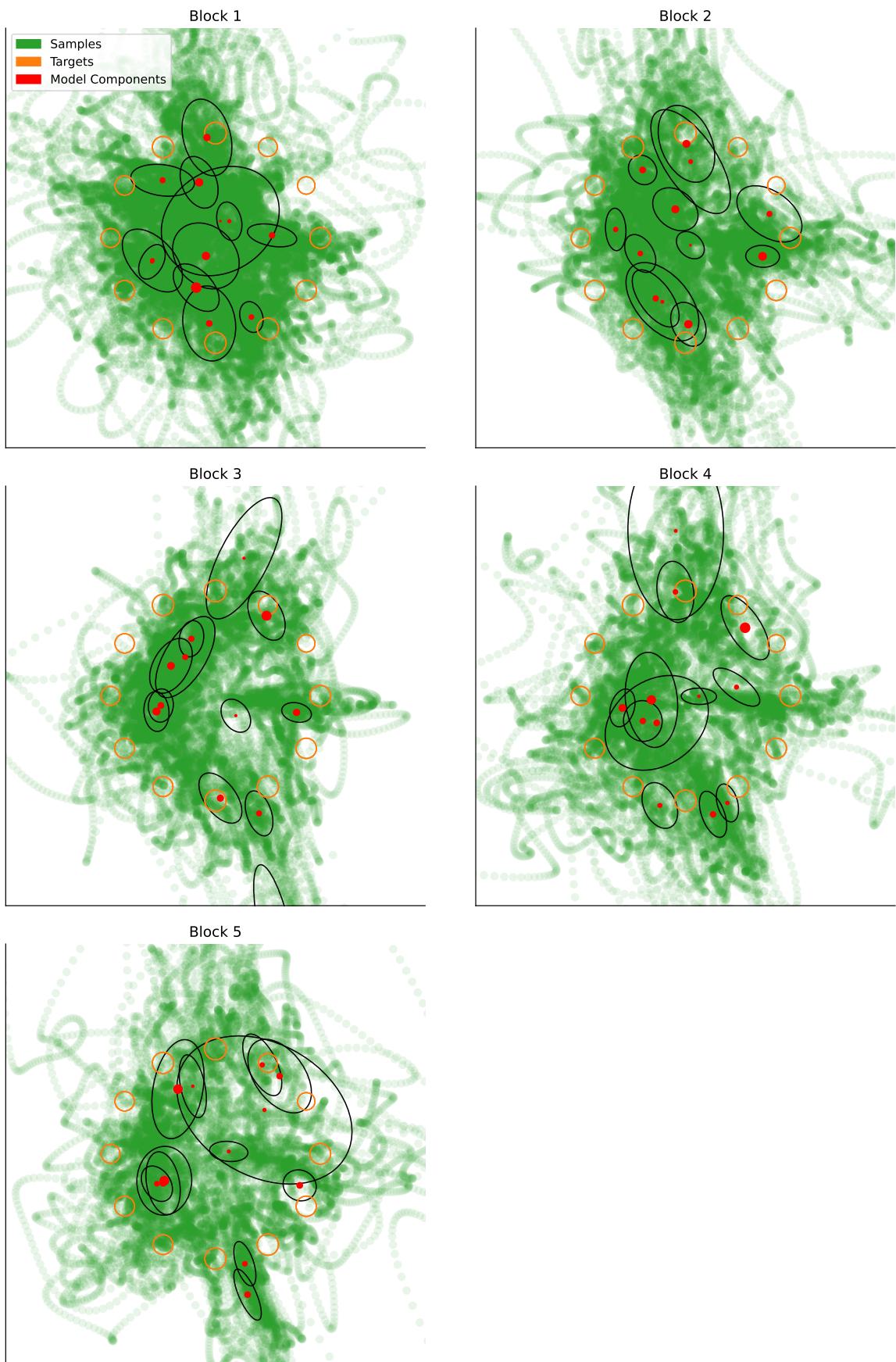
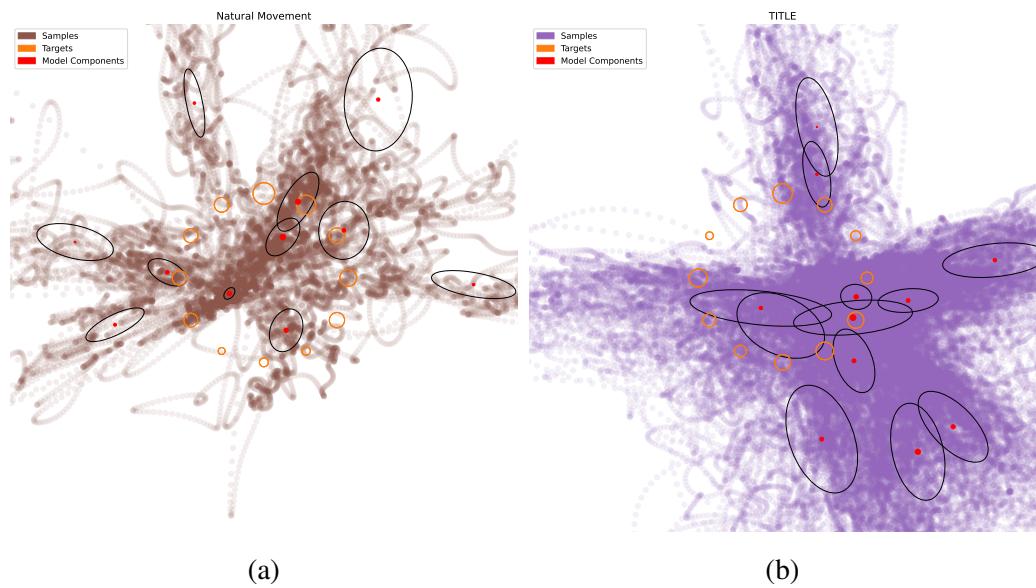


Figure 6.8: CAPTION  
132



(a)

(b)

Figure 6.9: CAPTION

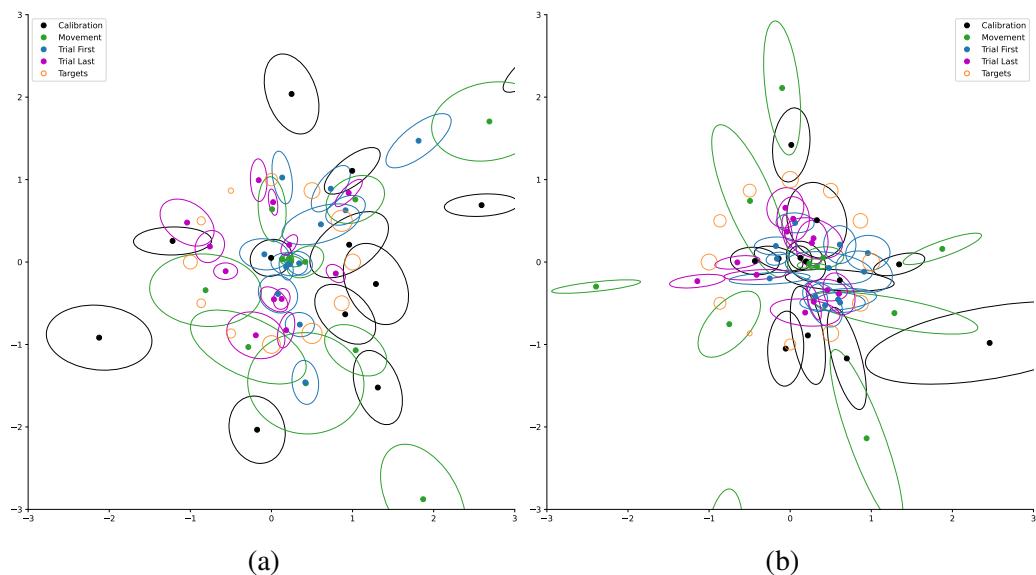


Figure 6.10: CAPTION

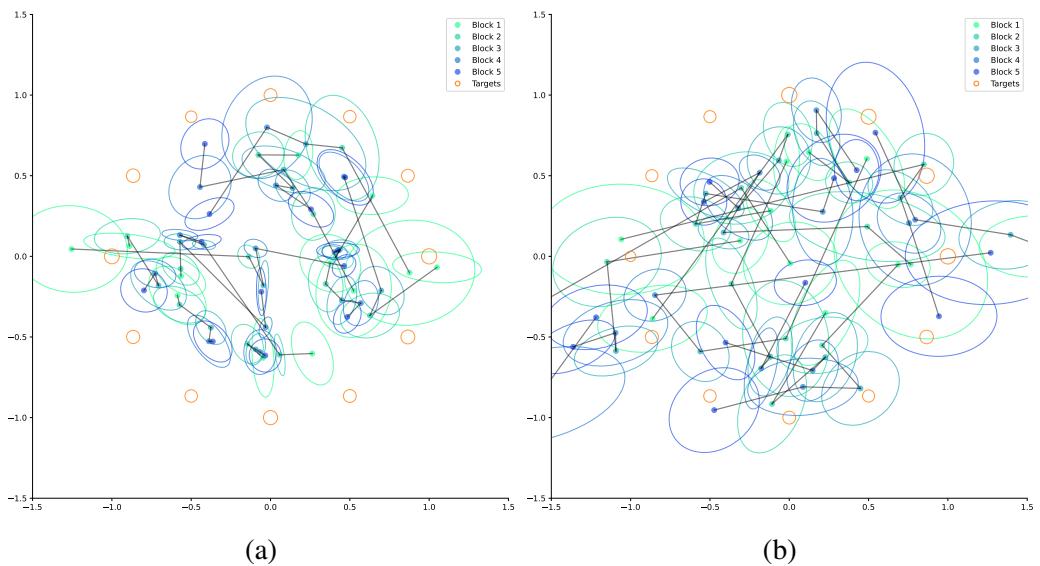


Figure 6.11: CAPTION

### 6.4.3 GMM Pseudo-derivative

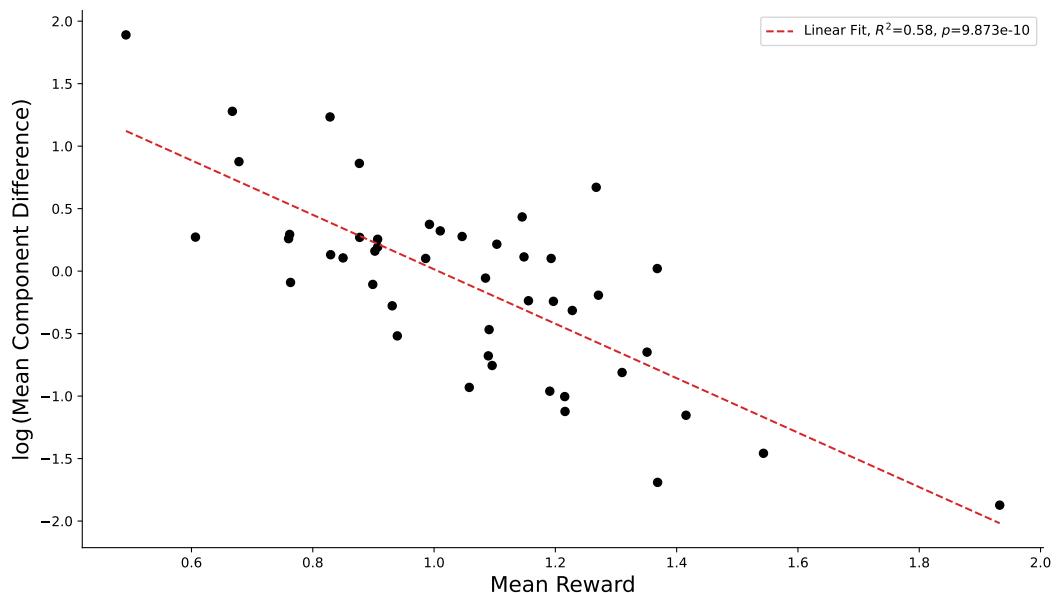


Figure 6.12: CAPTION

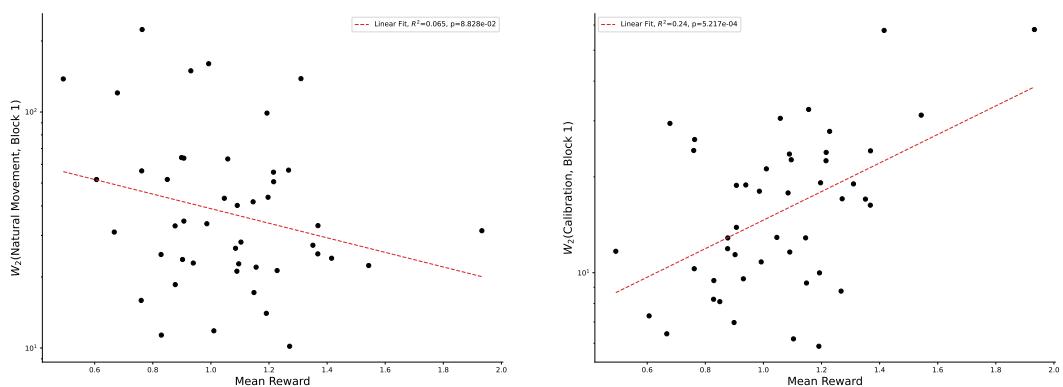


Figure 6.13: CAPTION

#### 6.4.4 GMM Effective Rank

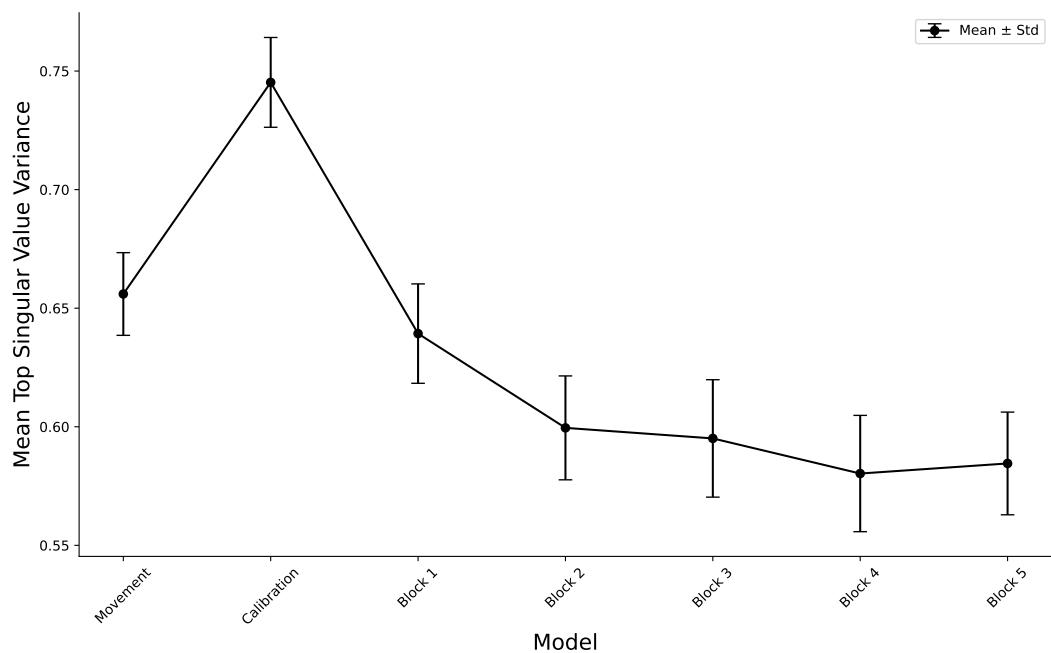


Figure 6.14: CAPTION

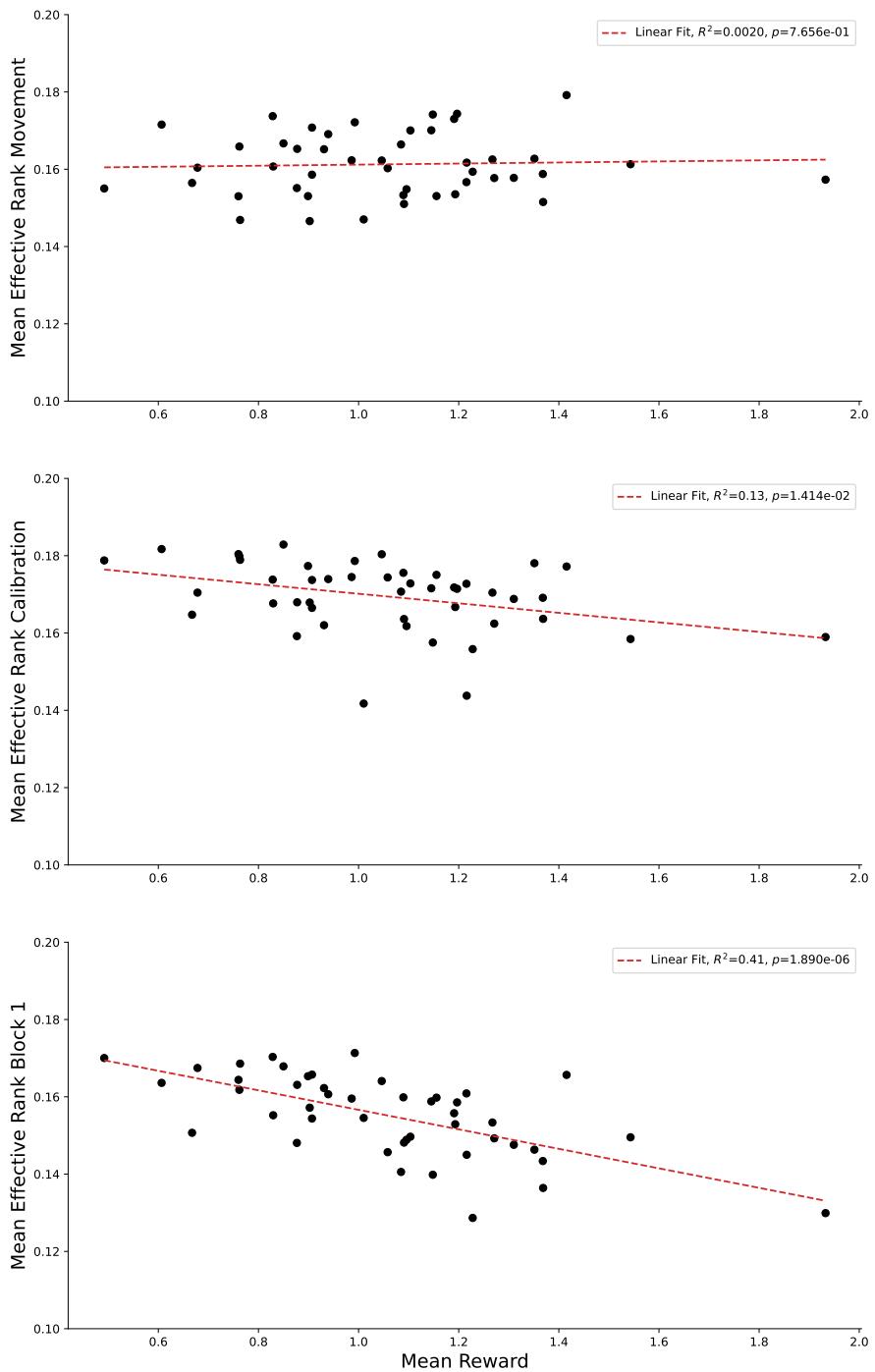


Figure 6.15: CAPTION

## 6.4.5 GMM Entropy

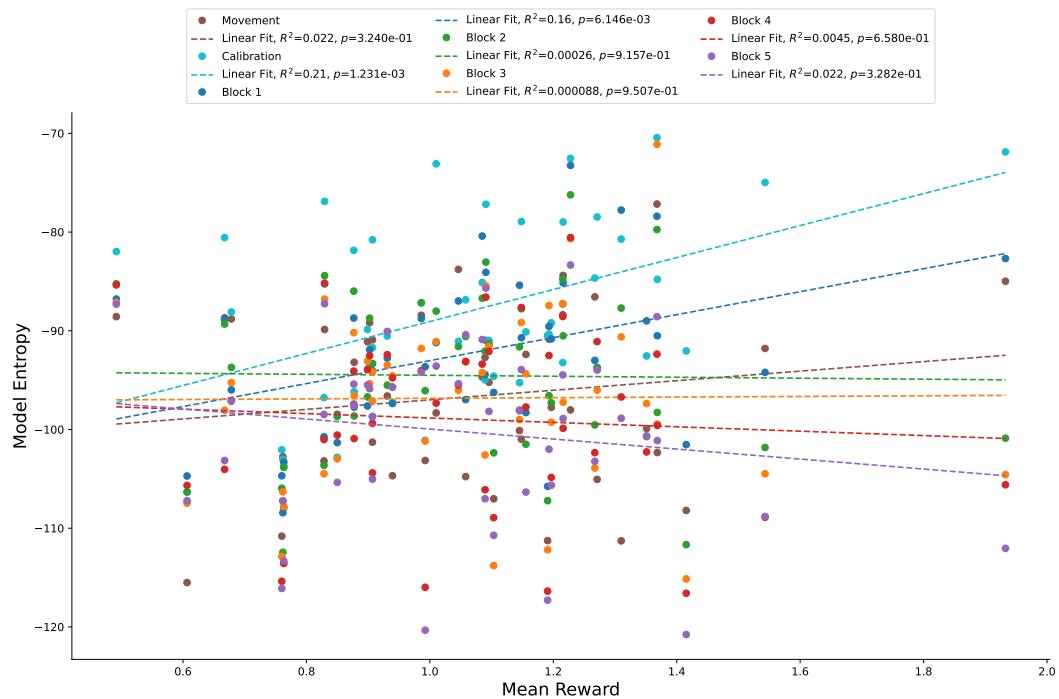


Figure 6.16: CAPTION

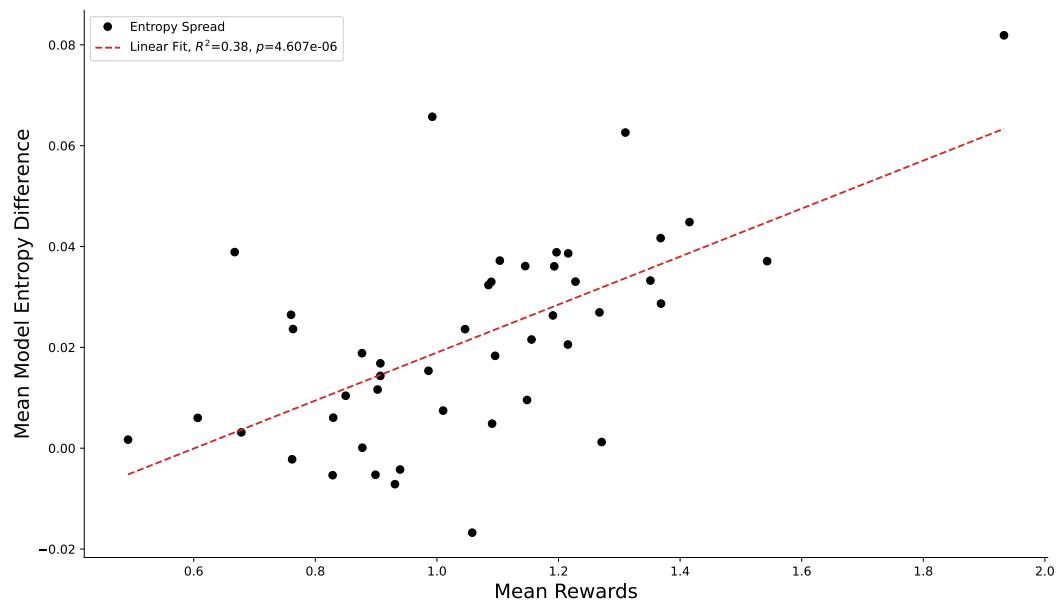


Figure 6.17: CAPTION

## 6.4.6 Log Transforming the Data

Show gaussian mixture fitting, show how they compare to data, show how stats of these change over trials

## 6.4.7 Influence of Natural Repertoire

Statistical distance between prior data and trial data Mahalanobis distance, but this breaks down for multi-model data Is there a metric we can use to measure similarity between these two datasets? Subjects are hard-constrained by their hands physiology, but we think they're also soft-constrained by their natural movement repertoire. We suggest casting this as a constrained policy optimization problem, where there is a regularization of the policy based on the prior movement data, which we suggest captures these natural constraints. The alternative we offer here is a norm-minimizing solution. This is connected to the task-space vs. null-space activity as the norm-minimizing solution effectively ceases to move in the nullspace of the task Fit GMMs to the trial and prior data, visualization these fits

Do we see a bias from natural statistics to trial statistics? Do subjects avoid distribution mismatch between their base/prior/default policy and their optimal policy? How do they efficiently explore and use this new data to update their internal model? What exploration strategy does a subject use to avoid mismatch? I.e. does variability in control/task space help or hurt subjects?

# **Notes**



## **Chapter 7**

# **CHAPTER NAME**

*QUOTE QUOTE QUOTE QUOTE QUOTE QUOTE QUOTE QUOTE QUOTE*

— AUTHOR, *SOURCE*



## 7.1 GMM Null Space

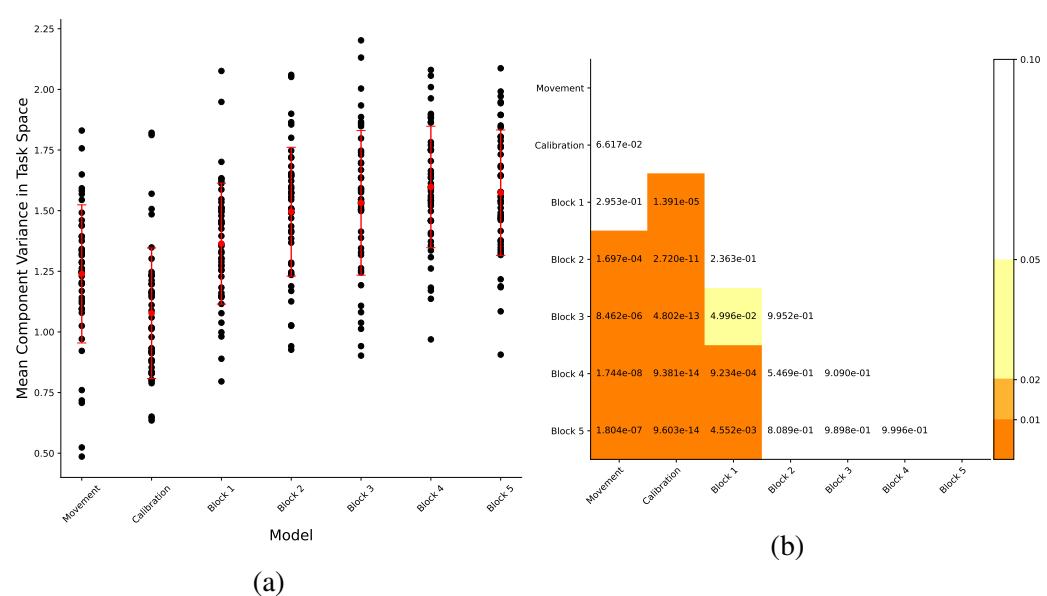


Figure 7.1: CAPTION



## 7.2 Nullspace Ratio of Task Error

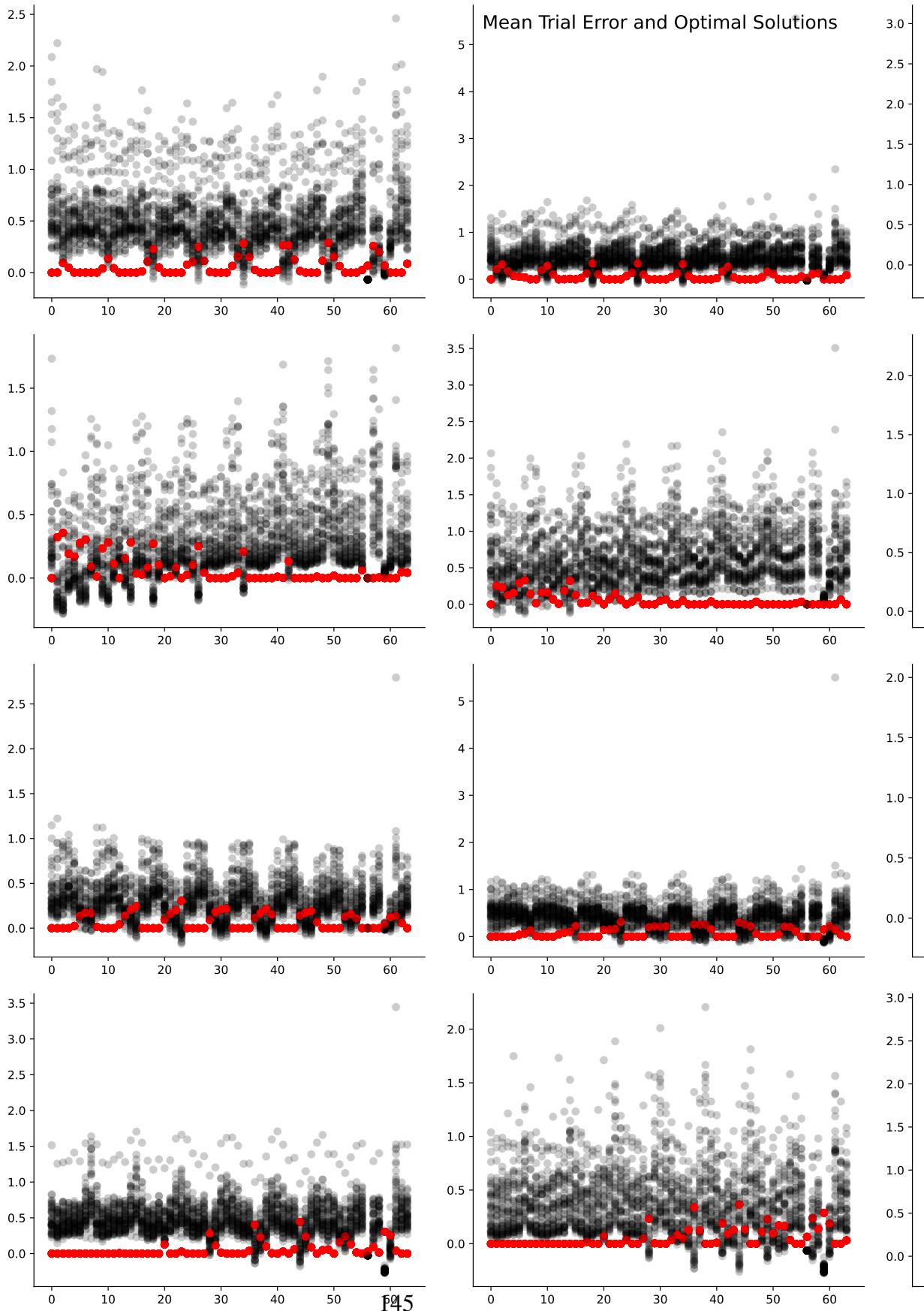


Figure 7.2: CAPTION

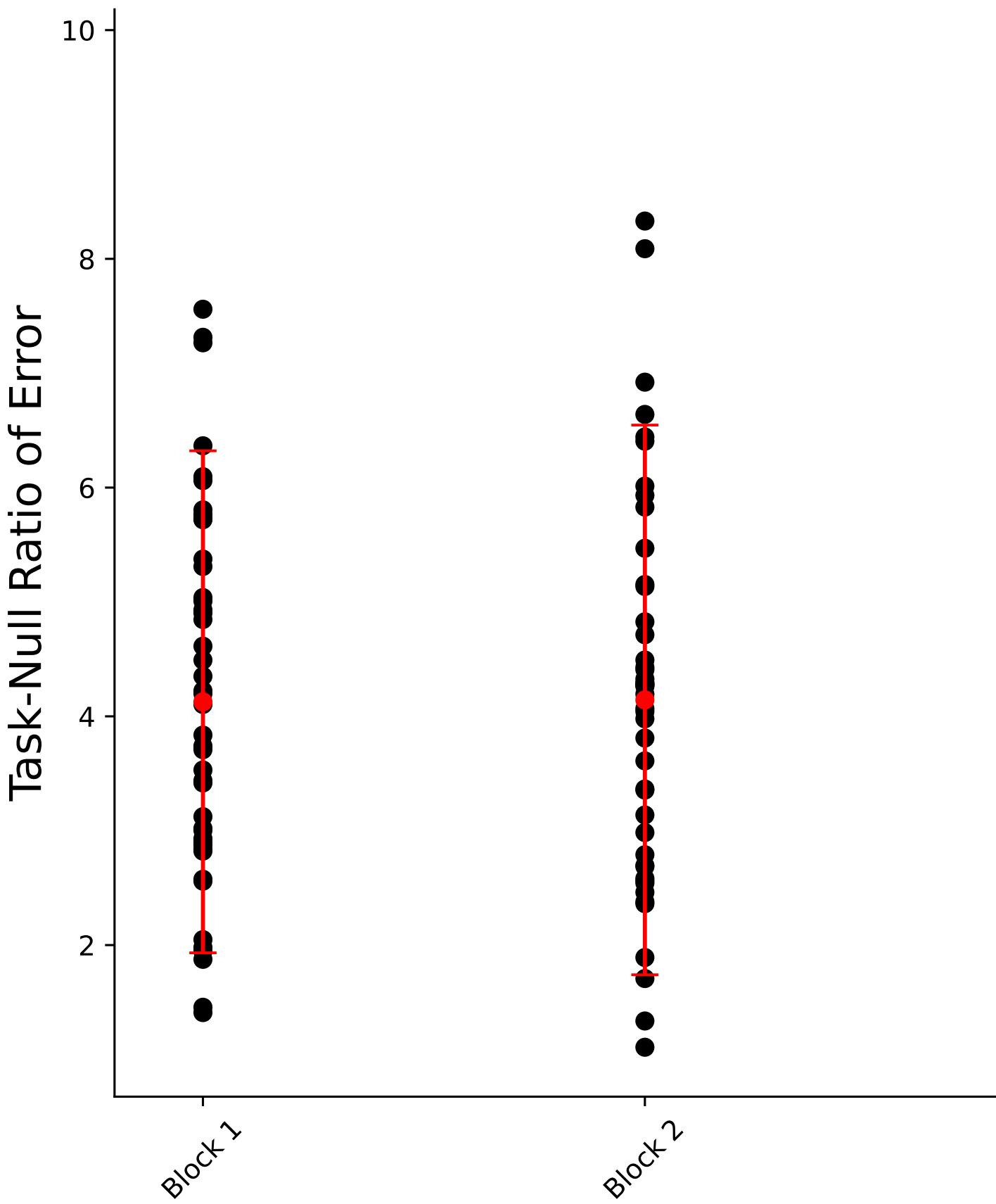


Figure 7.3; CAPTION  
146

## 7.3 Nullspace Optimization Solutions

Compare subject EMG solutions to optimization solutions NB <https://math.stackexchange.com/questions>

Hypothesis: subject error ratio of err-pinv / err-weighted will be  $\approx 1$  as subjects tend to minimize their error to the prior-weighted solution rather than the pinv solution.  
Todo – use a metric Check if fitting different subsets of the prior data with GMMs makes the distance thing work out! Try calibration only Try movement only There is a result here regardless of the outcome!!! This is about choosing a cost function and comparing to the data Can I fit a cost function to the data? This is like fitting a model... We know subject solutions hit the target, but what do they minimize over time? Can we regress these solutions against potential regressors?

How do subjects EMG solutions compare to optimization-derived solutions with varying levels of prior influence? I.e. Are subjects biased towards their prior data?

Do the variances of subject errors (subject solution - optimization solution) lie in the null space directions on average or not? I.e. Where do subjects expend most of their variance beyond achieving the requirements of the task

Nonnegative min-norm solution - activity = mostly in the null space or task space? This says something about the error! Does the error live in the null space or task space, does this change over time?

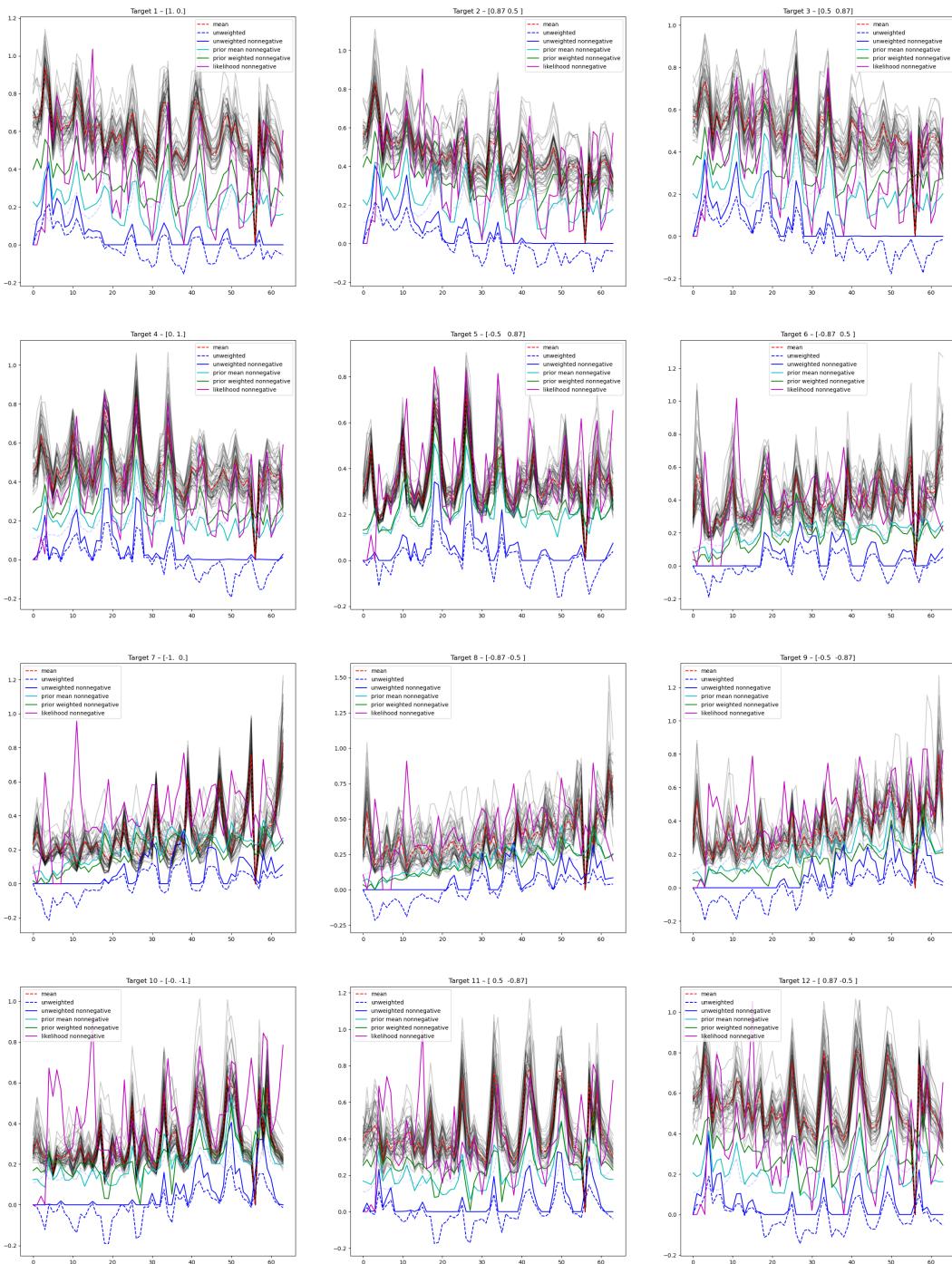


Figure 7.4: CAPTION

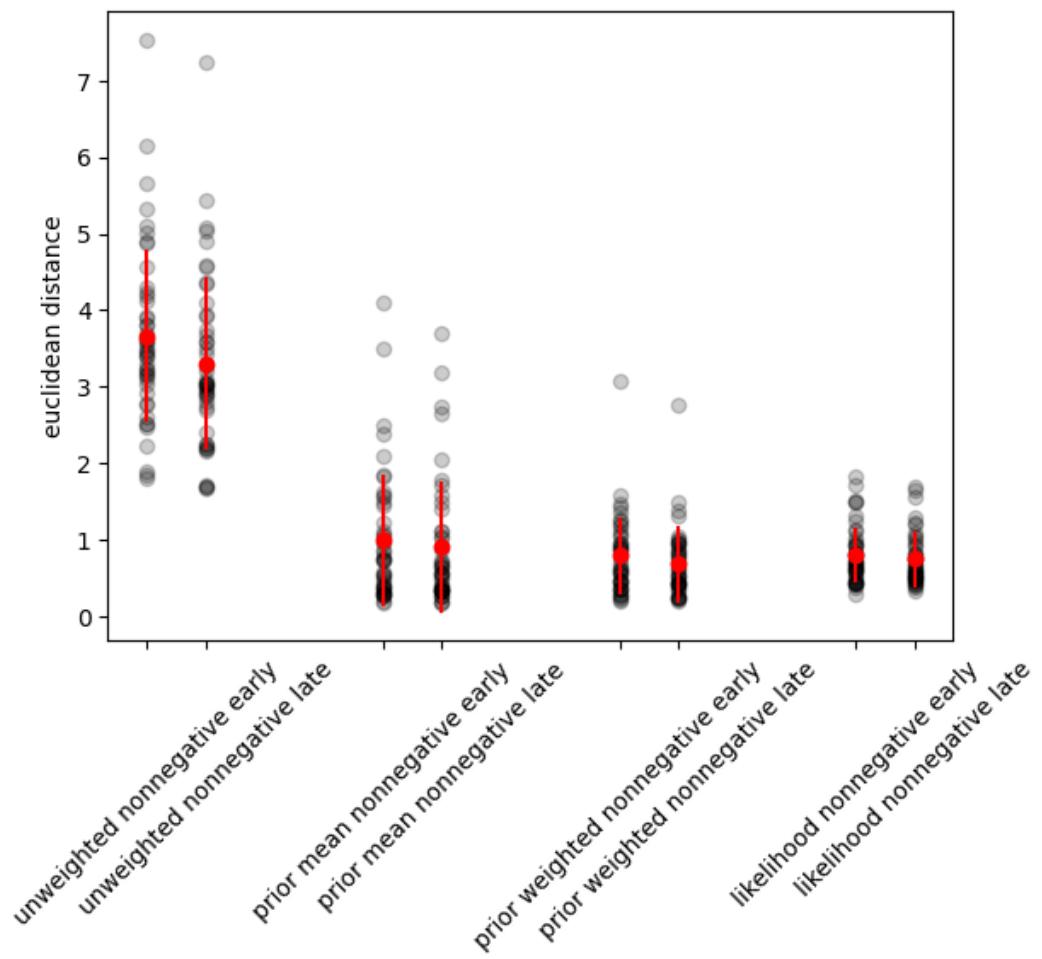


Figure 7.5: CAPTION

## 7.4 Classifying Nullspace Noise

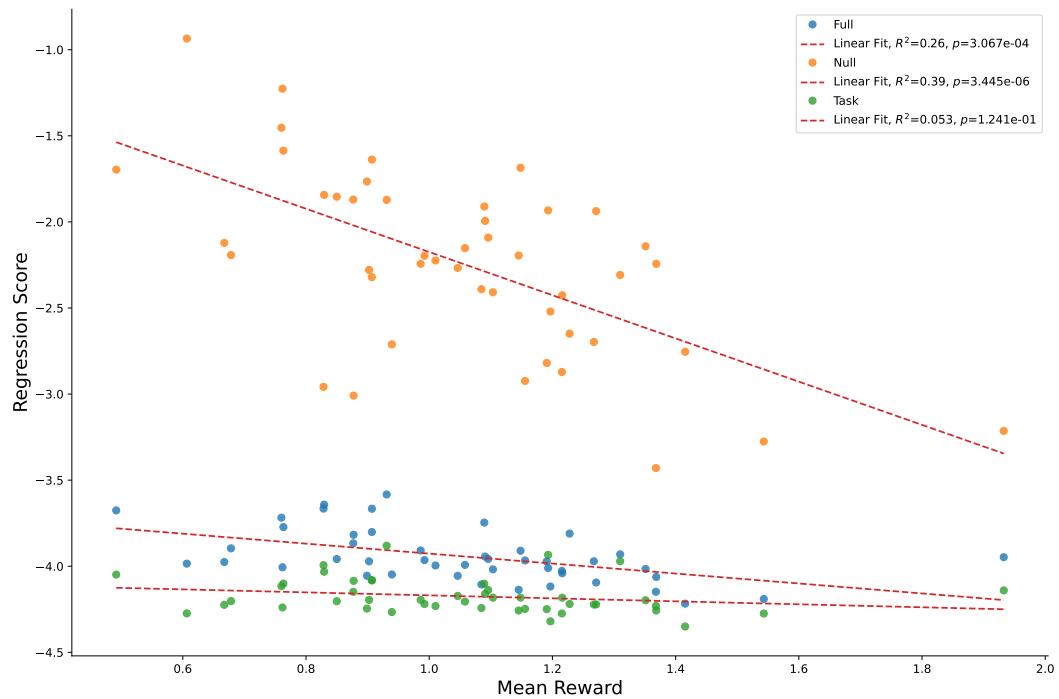


Figure 7.6: CAPTION

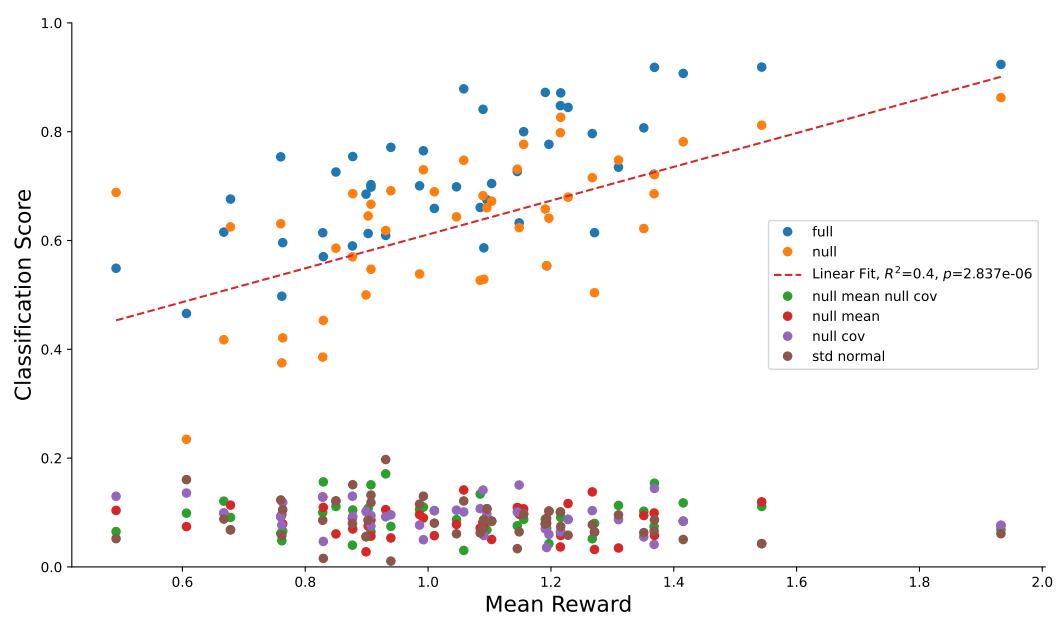


Figure 7.7: CAPTION

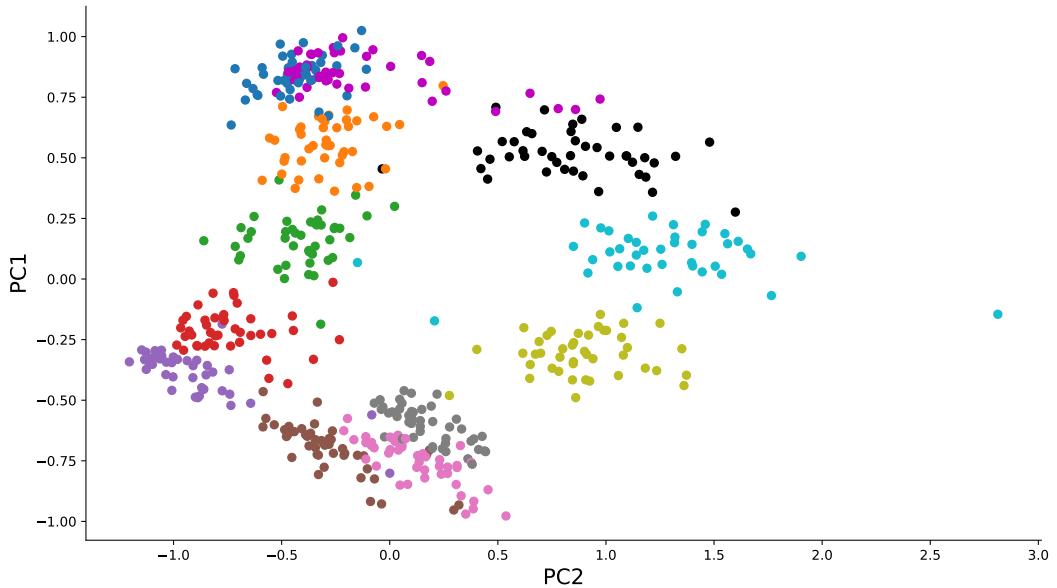


Figure 7.8: CAPTION

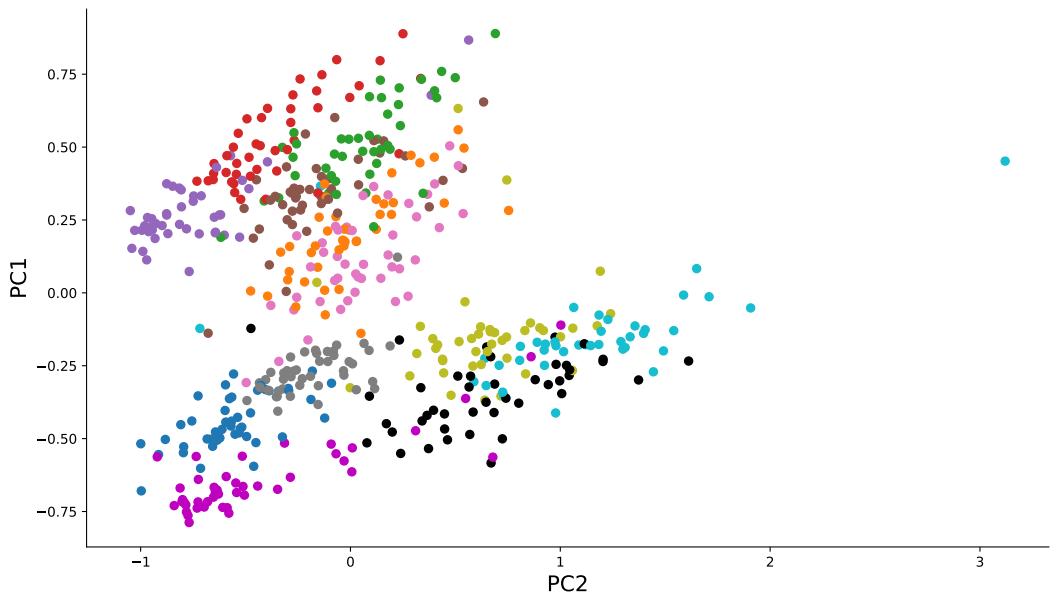


Figure 7.9: CAPTION

What do subjects learn? Do they learn specific rote movements, or are they flexible in their solutions?

Look at the null space for task relevant information – think about this, have we removed all the task relevant information before looking at this noise?



# **Notes**



## Chapter 8

# Conclusions & Next Steps

*When it comes to the problem of skilled movement, the algorithm is  
simply not known.*

— Wolpert & Ghahramani, 2000



## 8.1 Conclusions

## 8.2 Next Steps

### 8.2.1 EMG Hardware

Our preliminary data confirms the working principle of the setup and highlights the next steps for producing quality datasets. This is in accordance with the literature, where more advanced use of EMG is emerging as an important tool in understanding the complexities of motor computation[@Hug2011]. Our next steps are to build a new version of the EMG hardware doubling the number of channels to 64 electrodes placed across the forearm and to provide better hand constraints to ensure completely isometric contractions. The next hardware version will also include investments in shielding to provide proper noise mitigation. Most EMG in the literature is smoothed and trial-averaged due to noise, but we are confident that our records can be analyzed at the level of single trials, much like recent developments in neural data analyses [@churchlandNeuralVariabilityPremotor2006; @churchlandNeuralPopulationDynamics2012a].

### 8.2.2 Eye Tracking

To completely close the loop in our experiments, we are working to integrate pupil and gaze tracking to more closely follow the perceptual aspects of our task. We hope to find correlations in line with the literature dealing with active learning[@yangTheoreticalPerspectivesActive2016; @huangActiveLearningLearning2008].

*i* EMG activity was recorded using hook-wire electrodes (44 gauge with a 27 gauge cannula; Nicolet Biomedical, Madison, WI) placed in the muscle for the duration of single recording sessions. [...] Electrode voltages were amplified, band-pass filtered (150–500 Hz, four pole, 24 db/octave), sampled at 1000 Hz, and digitized. Off-line, raw traces were differentiated (to remove any remaining baseline), rectified, smoothed with a Gaussian (SD of 15 ms) and averaged. [@churchlandNeuralVariabilityPremotor2006]

### 8.2.3 Preprocessing

Preprocessing of the EMG signal per-channel is another key area for improvement. EMG signal is the convolution of motor unit action potentials terminating near the electrode sites. Ideally we could filter each channel of raw EMG to infer the signal envelope. There is precedent in the literature for Bayesian filtering of EMG signal using a Laplacian distribution. Sanger used a Laplacian distribution to filter a one-dimensional EMG signal[@sangerBayesianFilteringMyoelectric2007]. In accordance with this choice, Nazarpour found that as more motor units are recruited, the EMG distribution shifts from super-gaussian to gaussian following the central limit theorem[@nazarpourNoteProbabilityDistribution2013]. That work suggests methods for estimating high-order statistics for better filtering at low contraction levels relevant to our experiments. Such methods may prove to be more rigorous for inferring motor unit activations from our raw signal.

A window of EMG of length  $T$  samples can be modeled as a convolution

$$\mathbf{z} = \sum_t^T \mathbf{h} * \mathbf{s}$$

where  $\mathbf{h}$  is a motor unit activation template, which itself is a particular neural spike waveform, and  $\mathbf{s}$  is the incidence of a spike, which might be modeled as a point process.

### 8.2.4 Calibration

With a filtered raw signal per channel, our goal is to devise mapping from EMG space to task space which are biophysically achievable by our subjects but require a degree of learning over many trials. In our preliminary task, we hardcoded mappings. Our next step would be to design a calibration task which asks subjects to actively explore the biophysical limits of EMG space that can be captured by our electrodes. This is akin to extracting features of spontaneous activity and passive viewing used in cortical BMI experiments to generate "intuitive" mappings [@Clancy2014; @sadtlerNeuralConstraintsLearning2014].

With high-dimensional EMG, we would ideally devise a principled method

of extracting modes from raw EMG that accurately reflect modes of neural drive, demixing neural modes across channels. Here we used PCA as a first step. One starting point would be to align with the cortical BMI literature and use factor analysis and Kalman filtering[@sadtlerNeuralConstraintsLearning2014]. The issue, however, is how to design a task which evokes the available modes of possible EMG activity before using dimensionality reduction to generate learnable yet non-intuitive mappings.

autoencoders (farina paper) [@vujaklijaOnlineMappingEMG2018]

Another direction for analysis is to study long-range correlations in EMG data within and across trials through empirical correlation functions[@crevecoeurGoldstandardApproachA]. This work may inform features of models which attempt to recover aspects of trial-to-trial learning.

### 8.2.5 Task Design and Data Collection

With a working calibration task, our next goal is to use this calibration data to generate mappings to track learning. Our center-hold, reach-out task combined these two steps into one task for validation of the setup using hardcoded mappings. We can maintain the center-hold, reach-out style of the task but with data-driven mappings, as well as construct novel tasks to test predictions of our models. Scaling this task up to multiple subjects across days will provide a dataset with which we can test hypotheses concerning the evolution of complexity and correlations across learning. This will also give us an opportunity to study inter-subject variability, for which there is precedent in the literature which suggests individual strategies[@crouzierIndividualDifferencesDistribution2019].

### 8.2.6 Modeling Control

One relevant aspect of the basic optimal feedback control model is that the optimal controller that arises from specifying a quadratic state and control cost is invariant to the target state. In spite of this, we can use the aforementioned task to test predictions of the basic LQR model with respect to state and control noise and imperfect dynamics.

We expect to validate the basic optimal control models for our setup as we've designed the learning environment specifically for the EMG signal provided by the subject acts as the input to chosen virtual dynamics, which can be chosen in accordance with our model. We can then test perturbations in our task with respect to noise, goal, and dynamics and compare subjects' responses to our models.

The question then becomes: when might subjects need to internalize a new control policy? When might they need to internalize multiple control policies? We hope to work towards answers of these questions alongside models of compositional control. Such control models could deal with, for example, target uncertainty as well as multiple competing targets[@gallivanActionPlanCooptimization2015;@gallivanParallelSpecificationCompeting2016].

- stochastic optimal control model comparisons - cost models - perturbations in goal - go-before you know / goal uncertainty - noise perturbations — do responses match the models?

- dynamics model fitting - internal model uncertainty - modeled with robust optimal control?

(task reads out from D muscles, find modes of that data; do PCA to get K  $\downarrow$  D dimensions, controller only responds to motion in those K directions)—does behavior + motor activity follow LQR? this question has already been asked, but it hasn't been asked for this kind of high-to-low dim mapping. Its been asked in tasks where muscles haven't been directly in control (Valero-Cuevas 2009).

Todorov: do a task, look at muscle signal. Muscles that aren't necessary for task have higher variability b/c they're not being optimized for task (but doesn't introduce perturbations). Also see Loeb (2012) for a negative result saying that muscle coordination is habitual rather than optimal, but it has issues (low number of muscles). Can we replicate previous reaching optimality results in our set-up? What's unique about our set-up is the PCA/dimensionality reduction in muscle activity space. This is important because you can create arbitrary muscle-cursor mappings, so you have to learn a new skill/mapping. This is different than perturbing a fundamental movement and forcing adaptation, which is what has been previously done. For our task,

the participants actually have to learn a new task/mapping, rather than just do what they already know and be robust to perturbations. We test the LQR hypothesis once theyve learned the task, because LQR isnt a learning theory, its a theory about optimal control. We can see if, once people learn a new skill, their behavior is optimal wrt LQR theory. If we establish this, then we can think about how this LQR model is actually learned (enter RL).

## 8.3 Modeling Learning

Ultimately, our goal is to adapt optimal control models which begin as coarse approximations and are updated both within and across trials. Adaptation typically refers to online alterations to control policies while learning might refer to across-trial policy changes. Our theoretical aim is to devise models of learning and movement construction which extend the optimal feedback control framework through additions of composition and error-based updates.

gradient descent stuff

Stemming from our work using simple gradient descent to update internal dynamics models, we would like to gain a better understanding of the loss landscape. It may be possible to compute the optimum analytically and to corrupt the dynamics matrix in a more principled way. We will also explore the action of the resulting gradient, and compute second-order derivatives, and compute derivatives with respect to the control law  $K$  as a comparison. These results can then be compared with results from the reaching adaptation literature. This work can be guided by analyzing our empirical data to understand what aspects of our trajectories in EMG and task space are changing over trial.

- explore connections between spectral analysis perspective of optimal control and empirical correlation functions from data
- make the connection between control and dynamics in model and experiment more tightly integrated.
- learning control via reward (RL) [@vanderkooijLearningReachTrajectory2021]

## 8.4 Open Questions

- What is the best calibration task to find the boundaries of the available EMG space? - What are the defining features of learning in our EMG-based task? - Are our experiments well-modeled by the optimal control framework? - How do subjects efficiently use error information from each trial and feedback from each time step to update their forward model and control policy/policies? How do subjects balance policy updates with model updates? - How does a subject sample the state space to efficiently learn? Do subjects sample optimally? - How can you empirically disentangle system identification (model estimation) and policy learning? If subjects are suboptimal, is it due to model mismatch or a suboptimal policy?

- On what scale (trials, timesteps) is the model altered? the policy? - Replanning at every timestep is a model predictive control algorithm - What prediction can we make for ID/learning every trial? - how does a subject avoid "distribution mismatch" between their base policy and their optimal policy? How do they efficiently explore and use this new data to update their internal model? - what exploration strategy does a subject use to avoid mismatch? - What is a subject's baseline/prior model?  $y_t = \hat{f}_0(x_t, u_t)$  or  $y_t \propto p_0(y_t | x_t, u_t)$  - What is the base policy / prior policy?  $u_t = \pi_0(\hat{x}_t)$  - How do we think about learning a distribution over trajectories in control law space, or perhaps equivalently, in covariance/precision space? - We might hypothesize that a subject will act as randomly as possible while minimizing cost, a maximum entropy solution that converges to an optimal controller?  $\mathcal{H}(p(u_t | x_t))$  - How does a subject penalize changes to their controllers? Do they follow a KL-divergence type of measurement when improving their policy? - How do we best modeling learning LQR controllers trial-to-trial? - how do we use existing controllers to construct movements? - how do we construct controllers under dynamical and goal uncertainty? - How does the observation mapping relate to the latent state covariance? The task state covariance? - How do we formalize this into a probabilistic graphical model? Why would we? - Would this make it easier to reason about what the goals are? - Would learning  $M$  become an inference problem? - Would solving the control problem become an inference problem...? - What

noise assumptions can we make? Can we not make? - How can we incorporate signal-dependent noise?

## 8.5 Model-based optimal control & LQR

comparisons to LQR – discussion of theory and that model Null and task variability for LQR model and data – discussion – motivate the idea of treating the data as iid samples (“dart throws”) at the target Compare the trajectory stats to the LQR stats Discussion about learning this LQR? Learning a linear policy? Do we expect our experiments well-modeled by the optimal control framework? What would give us an indication that subjects are using an internal model? Are subjects predictive? Are they responding to disturbances in an expectedly optimal way? What are the defining features of the model to compare? Trajectory-based? Trajectory statistics Null/task space variability? But note that our task isn't redundant in the task space, its redundant in the control→task mapping Can we pass our EMG through LQR dynamics and recover trajectories? Is the LQR a reasonable model for EMG control? As the task dynamic is designed to be linear-gaussian, does the control solution developed by subjects match the “decision” made by an LQR controller? We can investigate this by providing the LQR controller with actual subject data on each timestep, and comparing the difference at each timestep between the subjects control response and the LQR controller. (This is the opposite of what the inverse control problem is asking: at each timestep, what cost function, under a given model, best explains the response of the subject? Here we're asking for an assumed cost function, what would the optimal controller do, and how does this differ from decisions we see subjects making?) Hypothesis: Early in learning, no, subjects are not akin to an LQR. Later in learning, they become closer to LQR controllers as they internalize a model of the task contingency. Subjects begin responding in a “least action” manner in the course of learning.

## 8.6 Inverse Optimal Control

What are subjects' cost functions? Is it straight line movements in the task space? Something else? Hypothesis: subjects optimize for straight-line movements in the

task space, as well as for minimum-variance movements. We can test this hypothesis by tracking “straightness” of trajectories over the course of learning for subjects who improve over trials. Does this change along with performance? Hypothesis: If we take subjects movements as a priori optimal and fit cost functions to their movements, over trials we will see their cost functions converge to straight-line, minimum variance movement (the prediction of OFC). (NB: this is a stab in the dark, the work here is interpreting the cost functions we get out, and defending our choice of dynamics and other assumptions)

# Bibliography

- [1] Kober, J., Bagnell, J. A. & Peters, J. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research* **32**, 1238–1274 (2013).
- [2] Sauerbrei, B. A. *et al.* Cortical pattern generation during dexterous movement is input-driven. *Nature* (2019).
- [3] Bernstein, N. *The Coordination and Regulation of Movements* (Pergamon, Oxford, 1967).
- [4] Kitano, H. Biological robustness. *Nature Reviews Genetics* **5**, 826–837 (2004).