

This is a section of [doi:10.7551/mitpress/14630.001.0001](https://doi.org/10.7551/mitpress/14630.001.0001)

# Context Changes Everything

## How Constraints Create Coherence

By: Alicia Juarrero

### Citation:

*Context Changes Everything: How Constraints Create Coherence*

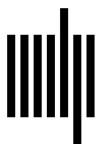
By: Alicia Juarrero

DOI: 10.7551/mitpress/14630.001.0001

ISBN (electronic): 9780262374774

Publisher: The MIT Press

Published: 2023



The MIT Press

---

## Empirical Research on Delayed Response

### Neuroscience Case Studies about Constraints

Empirical observations processed with certain analytic tools support the hypotheses presented in previous chapters. This chapter describes two sets of neuroscience experiments on delayed responses in laboratory animals. The first suggests that after training (context-dependent constraints), brains of macaque monkeys self-organize a coherent neural subspace with task-defined emergent properties. As revealed by a particular statistical technique, this neural subspace structures, initializes, and controls the monkeys' delayed response on each trial run such that it satisfies the task instructions, which vary each time. The second part of this chapter presents a related study that further explains the neural mechanisms that underpin delayed responses. Specifically, this second study concludes that self-organized and task-defined neural subspaces bring about behavior that is appropriate to the task at hand by organizing dynamic attractors in abstract neural state space. A combination of empirical research and simulations uncovers both the formation of such neural attractors and the constraints that control the monkeys' delayed actions and successfully satisfy the task's demands.

\* \* \*

Neural activity preparatory to voluntary behavior has been of interest to neuroscientists attempting to elucidate the contents of the mind's black box. The mind-brain complex, however, did not readily give up its secrets to studies that relied on summing and averaging single-neuron activity. In particular, when experimental conditions require the subjects to integrate and coordinate distinct streams of information, data sets of single-neuron recordings, no matter how large, are so highly heterogeneous (both across neurons and across experimental conditions) that no obvious moment-by-moment correlation with sensory input or motor output can be detected. When carrying out complex but contextually constrained tasks requires "coordination across neurons," single-neuron recordings have

proven particularly unsuitable, no matter how large the data sets or fast the processing runs (Cunningham and Yu 2014). When experiments require subjects to internally process a variety of stimuli before performing a task-defined behavior, data sets that sum and average individual neural responses provide no useful insight.

At first, the lack of clear-cut input/output patterns in single-neuron recordings was attributed to noise. Gradually, neuroscientists began to suspect that “this single-neuron complexity may be the realization of a coherent and testable neural mechanism *that exists only [as a dynamical system] at the level of the population*” (Cunningham and Yu 2014, 1503; emphasis added). Testing this hypothesis called for a different approach. As neuroscientists turned to population-level analyses, one class of statistical methods, *dimensionality reduction*, proved particularly useful in uncovering features of brain activity that could not be discerned with classical methods such as averaging responses across trials of single-neuron recordings. Dimensionality reduction, a statistical technique that transforms data from a high-dimensional representation, yielded insights into integrative neurological patterns that subjects organize in response to the constraints of training and the contextual cues of each trial run. Dimensionality reduction can extract those features that matter from noisy raw data.

Dimensionality reduction had previously been used to uncover neural mechanisms of monkeys performing tasks requiring attention. These studies, in contrast, focus on its application to *instructed delay tasks*, those that include an interval between stimulus presentation and the GO cue to initiate behavior. This delay period (after trial run instruction and stimulus presentation but prior to the initiation of motor activity) is called the *preparatory period*. Dimensionality reduction uncovered complex brain activity during that interval that was hidden in the chaos of single-recording data sets.

### Mark Churchland’s Team

The experimental conditions in both sets of experiments (Churchland, Cunningham, et al. 2012; Kaufman, Churchland et al. 2014; Mante, Sussillo, et al. 2013) are similar. The reaching experiments are described first. Monkeys fitted with electrodes on their motor and premotor cortexes were trained to reach for and touch (or trace) the correct target on a monitor in response to instructions. Some runs required a straight reach, others required a curved reach, some required the monkey to trace out a

maze on a screen in a clockwise direction, and on other runs it required a counterclockwise reach. Critically, there was a randomly varying delay interval between the preparatory period and the cue to initiate motion.

*Preparatory cortical activity* (or motor preparation) was defined as neural activity during the delay period between the instruction and display on the one hand, and the GO cue, on the other. The central question investigated in all these studies is, What exactly does preparatory activity do and how does it segue into actual and task-appropriate responses?

### Preparatory and Perimotor Neural Activity

Recordings showed that individual neurons are uncorrelated during the early preparatory period. For approximately the first 200–300 ms, not only does the temporal structure of recordings of individual neurons “var[y] widely across cells: some even increase their firing rate, some decrease, some arrive at an approximate plateau level, while others undulate” (Shenoy, Kaufman, et al. 2011, 38).

It was initially hypothesized that the combined activity of individual neurons during the preparatory period might represent a summing of all the physical dimensions that make up the full data set: distance to the target, direction of the target, or speed of arm motion to the target, the arm’s angle of rotation, and so on. However, this turned out not to be so. Few individual neurons appear tuned to these lower-level parameters (Cunningham and Yu 2014; Churchland, Santhanam, et al. 2006; Shenoy, Kaufman, et al. 2011).

Instead, dimensionality reduction revealed population-level neural activity that integrates a wide variety of specific factors to represent the task at hand. That is, during preparatory activity, the monkey constructs an abstract task space in neural space. This abstract space recodes noisy, high-dimensional data into a low-dimensional and complex neural space whose parameters capture salient aspects that covary with relevant elements of the task. The space also initializes (sets the initial conditions) for the subsequent *perimotor stage* of cortical activity, the interval between the GO cue and the actual initiation of movement. The task space organized during the preparatory period is therefore a “dynamical system [that] controls movement” (Churchland, Cunningham, et al. 2012, 51) in the sense that it is predictive of actual reaction times and movement.

Specifically, “Consider the space of all possible preparatory states (all possible Ps). For a given reach, there is presumably some small [contiguous] subregion of space containing those values of P that are adequate

to produce a successful [response]” (Shenoy, Kauffman, et al. 2011, 42). That is, it contains those values that would qualify as a successful performance of the required task. Applying dimensionality reduction to massively disparate data sets of individual neural activity during the preparatory period revealed this multidimensional but far smaller subregion of neural space (a task-defined and structured preparatory space) within which subsequent behavior must be constrained to successfully execute the [response] and fulfill the experimental task’s demands. As evidenced in the neural activity during the perimotor period, correct responses then flow from this organized preparatory subregion in neural state space into appropriate behavior.

Surprisingly, Churchland’s team found that the population-level activity of this organized space consists of rotating and continuous cyclic patterns of neural activity. That is, neural dynamics organized during preparatory activity are rhythmic. The monkey’s brain is recoding signals in the frequency domain. Discovering this rotation was surprising because “the [the task-appropriate responses required in the experiments] themselves are not rhythmic” (Churchland, Cunningham, et al. 2012, 51–52).

From the perspective presented in this book, the rhythmic rotations are significant because they remain invariant throughout the trial run: “for each data set the neural state rotates in the same direction across conditions (Churchland, Cunningham, et al. 2012, 53). The authors hypothesize that the continuous properties of the rhythmic activity (such as direction of rotation, amplitude, and phase) capture salient and relevant properties of the task that covary across different experimental conditions. They capture effective input, in our terminology. This conclusion is further supported by the fact that the properties of the rhythmic activity do not correspond to any physical features of the signals displayed. The continuous properties of the constructed task space appear to result “not from a multiphasic signal, but from how that signal is constructed” by the monkey (Churchland, Cunningham, et al. 2012, 55). These population-level patterns of neural activity, I submit, describe intensional content representing the task at hand (reaching for a target, looking at a target, determining if the relevant response is about dot color or dot motion, for example). It implies that the abstract space is eminently multiply realizable and indexically ordered. Indeed, the authors conclude that “many features of the observed rotations make sense in terms of how the actual behavior (EMG, kinematics) *might* be generated in the remainder of the trial run, rather than in terms of the behaviors themselves” (Churchland, Cunningham, et al. 2012, 55; emphasis added).

Rotational invariance of population-level neural activity suggests, in short, that in response to training, neural activity during the preparatory period creates a persistent constraint regime that represents task-defined requirements and anticipates a range of possible movements from which a response will be selected depending on the details of the particular run. It represents the meaningful features of a contextually constrained and type-defined task. The end state of the preparatory period then becomes the initial condition of perimotor activity for that trial run. “If this initial condition is known, subsequent states [the actual behavior on each trial run] can be predicted” (Churchland, Cunningham, et al. 2012, 53). That is, once the GO signal is given, the characteristics of the preparatory abstract space transition naturally through the perimotor neural activity into actual and task-relevant behavior, be it a successful reach or, in a different experiment, a saccade (Shenoy, Kaufman, et al. 2011, 39). It is tempting to conclude from this experiment that, as Patten and Auble proposed, the monkey’s cortex generates an abstract model with emergent properties that subsequently constrains actual muscle activity top down such that “individual unit responses . . . reflect the underlying dynamical factors: the patterns present on each axis of [task-defined] space” (Churchland, Cunningham et al. 2012, 54).

### **A Further Study**

In a second set of studies of context dependence in delayed response, rhesus macaque monkeys were presented with a display of noisy signals—in this case, differently colored light dots moving toward the left or right of a monitor. The monkeys were trained to discriminate between the light dots’ prevalent direction of motion (called the motion context) and their prevalent color (color context) in light of a contextual cue (task instruction) given prior to the beginning of each trial run. Cued to a color context, for example, the monkeys were trained to select the prevalent color, red or green, of the dots displayed; instructed to a motion context, the monkey had to select their prevalent direction of motion, toward the left or right. The monkeys reported their choices by saccade eye motions to one of two targets placed on either side of the monitor.

At the start of each trial, the monkeys were given instructions about the context for the upcoming run: whether the run would be about color or about motion. Once the contextual cue was provided, the display showed the target options on either side of the screen: red or green for

a color context, right or left for motion. The placement of these targets varied randomly on each trial run (sometimes the correct prevalent color target was on the left; at other times, it was on the right).

The monkeys were then presented with a random but noisy display of light dots that were both moving and variously colored. After a few seconds' delay (which, again, varied randomly), the GO cue was given.

The experiment demonstrated that the monkeys were able to discriminate evidence relevant to the instructed context for the task at hand while largely ignoring irrelevant evidence. If the trial run's cued context was about color, the monkeys largely ignored the dots' direction of motion, regardless of their speed and uniformity. If the trial run was about motion, the monkeys largely ignored color, regardless of the display's brightness or color uniformity. Strength of sensory input (how bright the color, for example) did not override relevance, indicating that the behavior was constrained top down, according to the relevant, task-defined contextual cue.

While monkeys performed this task, population-level responses from neurons in a region of the prefrontal cortex (PFC) that selects and executes eye movements, commands visuospatial attention, and integrates information "toward visuo-motor decisions" were recorded. At the start of each run, population-level neural activity converged to a starting baseline located near the center of the plots (as if to be positioned to respond correctly in either context). In the interval after the contextual cue (choose color!) and the target locations (red on right of screen, green on left) were presented but before the dot display was turned on, the monkeys' neural activity quickly moved away from the earlier center baseline and lined up along a neuronal pattern the scientists labeled an *axis of choice* (Mante, Sussillo, et al. 2013, 79).

As in the various experiments conducted by Churchland's team, Mante's so-called axis of choice is a structured, low-dimensional subregion of the monkeys' neural space. Specifically, Mante's team established that this subregion is organized as a line attractor with the options for each choice context located on either end of the axis (green and red in runs with a color context, right and left in runs with a direction of motion context).

Mante, Sussillo, et al. hypothesized that the gradual repositioning of neural activity along an organized axis of choice is evidence that the preparatory period integrates sensory input and contextual cues into a single domain that captures the particular trial run's context; the monkeys' brain activity organized the variables of the task into a coherent neural subregion that defines the requirements of the particular trial run and

task-defined response options; the axis of choice spans the task variables, without the values.

The study confirmed this hypothesis: comparing population responses in the PFC across contexts revealed that a single, stable neural pattern integrates sensory evidence in terms of the choices for each particular trial run. The integration combines in one domain the properties relevant to that trial run (dot color, red or green, or prevalent motion toward the right side of the monitor or left).

Choice thus begins by structuring a contextually generated type-defined subregion of possibility space in the PFC: “This will be a color run,” or “this will be a direction of motion run,” for example. This low-dimensional subspace is organized as the line attractor labeled the axis of choice; it represents the task requirements and the variables for the task: {color [red, green]} or {direction of motion [rightward, leftward]}. This possibility space aligns the options from which to subsequently select.

Significantly for our purposes, the authors of the study note that the results imply that integration and selection are two aspects of a common process separable only at the population level. The study by Mante, Sussillo, et al. thus agrees with experiments conducted by Churchland’s team: it proves the construction of an organized neural space that is contextually generated, task-defined, and multiply realizable, it spans possible behaviors that would satisfy that task. This neural space also defines the conditions of satisfaction that will constrain actual behavior top down such that the response is task-appropriate. In Lange’s terminology, the axis of choice describes a self-organized third dimension, a regime of constraints that structures a task-defined field, along with the possible range of values for that field. In the language of philosophical action theory, the attractor of the axis of choice might be considered the neural correlate of a *prior intention* to carry out the task of this trial run as instructed by the contextual cue.

Mante, Sussillo, et al. also and simultaneously recorded direct sensory responses to the motion and color displays. At the population level, PFC activity also showed that neural responses to sensory information alone oriented along two separate sensory attractors, one they called an axis of color, the other an axis of motion. Each of these *sensory axes* represented the order parameters of that run: they capture the significant properties (salience, cohesion) of the signals relevant to the contextual cue for that run. That is, they capture information about how strongly coherent the dot display was with respect to the context (color or motion), and how relevant to the contextual cue the features of the dot display were during



a given trial run. Neural activity along the axis of motion, for example, was less strong during trial runs cued to a color context—no matter how coherent the direction of motion or how bright the colors.

Interestingly, neural activity along these two sensory axes (motion and color) was momentary; neural activity returned to baseline when the display was turned off. Mante, Sussillo, et al. interpreted such transient sensory neural activity as providing only “momentary evidence . . . in favor of the two choices” (Mante, Sussillo, et al. 2013, 79), color or motion. In contrast, however, neural activity tuned along the axis of choice continued throughout the experiment until the actual response (saccade or reach) was completed. Unlike the sensory axes, the influence of the axis of choice appears to ensure that responses remain true to task by providing top-down governing constraints for the correct response to that trial run. The persistence of the axis of choice in constraining responses despite the intervening delay period can be interpreted as the standing cause of the response.

\* \* \*

Once the task at hand is identified and the corresponding subregion of neural state space is structured accordingly, how did the monkeys select one target rather than another? That is, how does neural activity (now constrained in a lower-dimensional space that is task-defined) select the correct response for that task on that run (red and not green) and then perform the correct action—looking toward the red target on the right rather than the green one on the left, say? In a simulation, Mante’s team found that a different vector participates in this final stage of the process.

When population responses across both axes are combined, the momentary information (red or green, right or left) registered along the sensory axes (for motion and color) and supporting a particular response for a given trial run arcs the neural population activity toward the correct answer—that is, toward the relevant sensory axis, motion or color. That is, the orientation of the combined neural activity that represents the integration of information from both the axis of choice and the relevant sensory axis projects strongly in the direction of the relevant target on the screen and strongly orthogonal to the irrelevant one. It represents the phase transition from many separate streams of information into one coherent set of interdependencies.

Mante, Sussillo, et al. call this coherent and arced aligning of neural activity a *selection vector*. By integrating the run’s instructed contextual alternatives with the momentary sensory display of that trial run (as given by the relevant sensory axis), the selection vector encodes relevance, a

highly context-dependent parameter. Unlike the orientation of the axis of choice, which is unaffected by the momentary recordings of the sensory stimuli along the sensory axes, the selection and performance of the actual response is therefore controlled by integrating both. Once the monkey must act, integrating all contexts simultaneously (the particular trial run context's axis of choice plus the current and momentary sensory information of the display) constructs and deforms a selection vector in neural state space—and orients it toward the relevant target and away from the irrelevant one. From our perspective, projecting strongly parallel to the relevant signals but orthogonal to the irrelevant ones is evidence that the selection vector's constraint regime integrates information on the display into those invariances previously established by the axis of choice (the constraints that govern the task at hand). In effect, a combination of context-dependent constraints, organized as a selection vector, recodes momentary sensory information in terms of its relevance (for this run, given the task context).

Mante, Sussillo, et al. describe the influence of the selection vector as *context-dependent relaxation*. “This mechanism explains how the same sensory input can result in movement along the line attractor in one context but not the other” (Mante, Sussillo, et al. 2013, 82). The process of context-dependent integration that induces the selection vector explains why a particular saccade is performed—toward the relevant target (displayed on one side of the monitor) and away from the other. And so, the correct (appropriate to this trial run) action is performed.

In contrast with the axis of choice, which determines effective input, the selection vector produces and regulates effective output. Applying dimensionality reduction to population-level responses thus confirmed the construction and presence of what we have called here effective input and output in the monkeys' motor cortex: population-level responses revealed that “identical sensory stimuli . . . can lead to very different behavioral responses depending on context” (Mante, Sussillo, et al. 2013, 78). This is possible only because of the previous Many-to-One self-organization of a type-defined task space.

By conditioning selection and behavioral activation on instructional context, the selection vector might also be considered as the neural correlate of a *proximate intention of action*. Proximate intention integrates the *intensional* content of the intention with motivation to the appropriate behavior such that the intensional content is satisfied.

From the perspective of this book, Mante's team shows how effective input and output are generated and acted on. First, and in response to the enabling constraints of training, neural activity induces a Many-to-One

transformation by integrating sensory signals into a coherent set of constrained interdependencies. These organize a type-defined abstract space that encodes continuous population-level properties. Dimensionality reduction reveals population-level neural activity that structures an indexically defined and abstract subregion of neural space. This neural space represents the interdependent type-defined task requirements as perceived by the subject. The neural space is multiply realizable; it spans all possible behaviors that might be required to satisfy those constraints. Once integrated into a coherent dynamic and in view of the constraints it embodies, a specific behavior is then selected and acted on in response to more timely and local constraints (provided by the display pattern of each run and as determined by the selection vector's orientation).

The studies by Mante et al. and Churchland et al. are congruent: the abstract subregion of neural state space organized during the perimotor period of Churchland et al.'s experiments corresponds to the line attractor of Mante, Sussillo, et al.'s axis of choice. In both cases, these are type-defined subregions that encode multiply realizable task-specified governing constraints that control for correct behavior while, top down, allowing it to vary depending on context.

From the perspective presented in this book, furthermore, combining the information represented by the axis of choice with sensory information into a selection vector interlaces semantics and syntax into a coherent set of interdependencies enabled and governed by the real-world constraints of the training and trial run instruction. As embodied in population-level neural patterns, the selection vector issues in meaningful behavior constrained top down by indexically defined interdependencies. The suite of delayed response experiments illustrates how enabling constraints (of training) generate coherent, multiply realizable neural dynamics with emergent (task-defined) properties that evolve into and control task-appropriate actions.

© 2023 Massachusetts Institute of Technology

This work is subject to a Creative Commons CC-BY-NC-ND license.

Subject to such license, all rights are reserved.



This book was set in Sabon by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Juarrero, Alicia, author.

Title: Context changes everything : how constraints create coherence /  
Alicia Juarrero.

Description: Cambridge, Massachusetts : The MIT Press, [2023] | Includes  
bibliographical references and index.

Identifiers: LCCN 2022030581 (print) | LCCN 2022030582 (ebook) |

ISBN 9780262545662 | ISBN 9780262374781 (epub) | ISBN 9780262374774 (pdf)

Subjects: LCSH: Complexity (Philosophy) | Causation.

Classification: LCC B105.C473 J83 2023 (print) | LCC B105.C473 (ebook) |

DDC 117—dc23/eng/20230124

LC record available at <https://lcn.loc.gov/2022030581>

LC ebook record available at <https://lcn.loc.gov/2022030582>