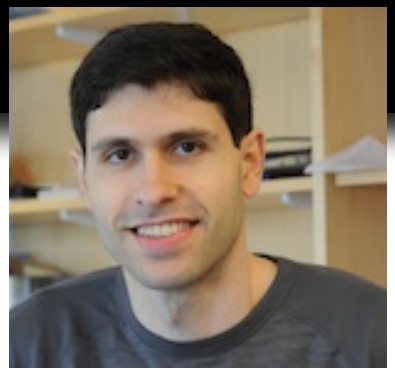
DiCarlo Lab's role on the MURI team:

Determine if and when the primate brain uses bidirectional computing (feedback) to successfully represent objects in images that challenge current feedforward computer vision systems (MURI Aim 1)



Elias Issa



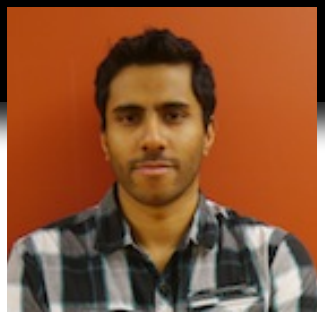
Kailyn Schmidt



Kohitij Kar



Arash Afraz



Rishi Rajalingham

James DiCarlo MD, PhD

Professor of Neuroscience and Head, Department of Brain and Cognitive Sciences
Investigator, The McGovern Institute for Brain Research
Massachusetts Institute of Technology, Cambridge MA, USA

brain+cognitive sciences

Logic of our approach and contribution

State-of-the-art computer vision is far from perfect

Humans perform better than state-of-the art CV on many tasks/images

=> human brain contains important to-be-discovered processing mechanisms

When monkeys perform similarly to humans on these tasks/images

=> monkey brain contains important to-be-discovered processing mechanisms (e.g. bidirectional computing?)

We can bring to bear more powerful tools to study these mechanisms

Overview of what we are contributing to the MURI team effort (Aim 1)

We can measure the evolution of the brain's solution in two ways:

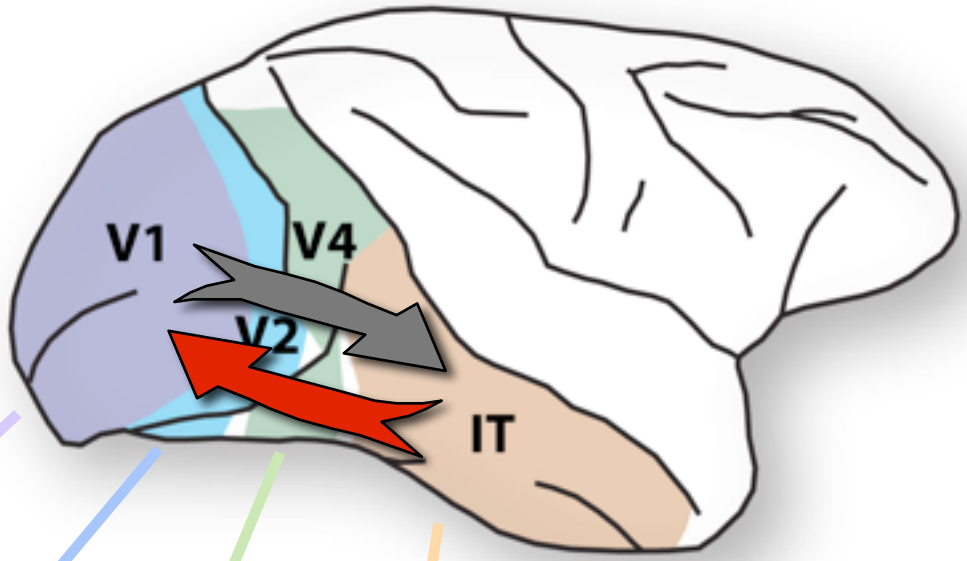
1. ***Subject's perceptual report***
2. ***IT decoded population activity***

Sought after deliverables:

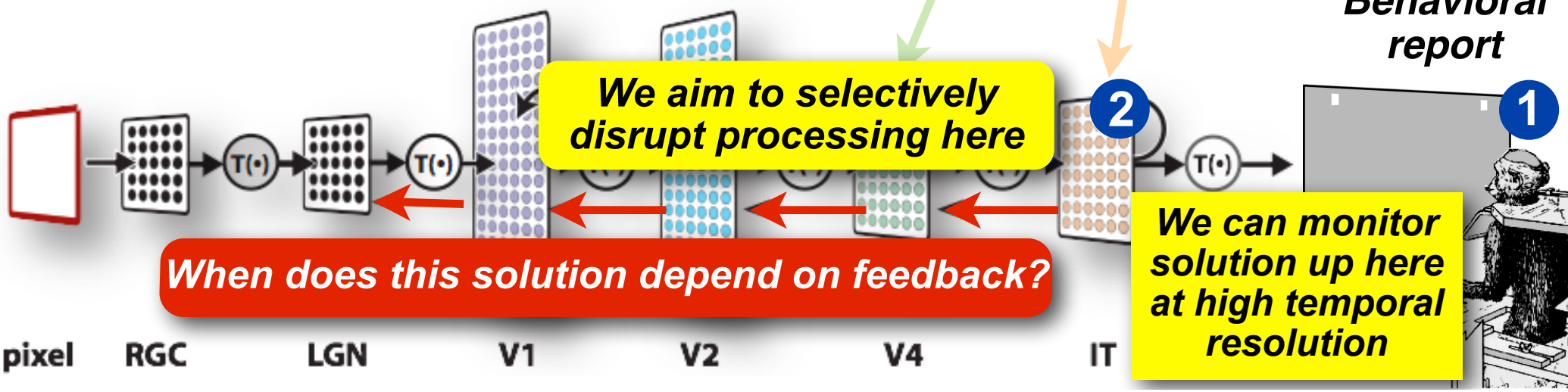
Which tasks and which images require ventral-stream cortical-cortical bidirectional computing?

↓ ↓

Guide CV models



Behavioral report

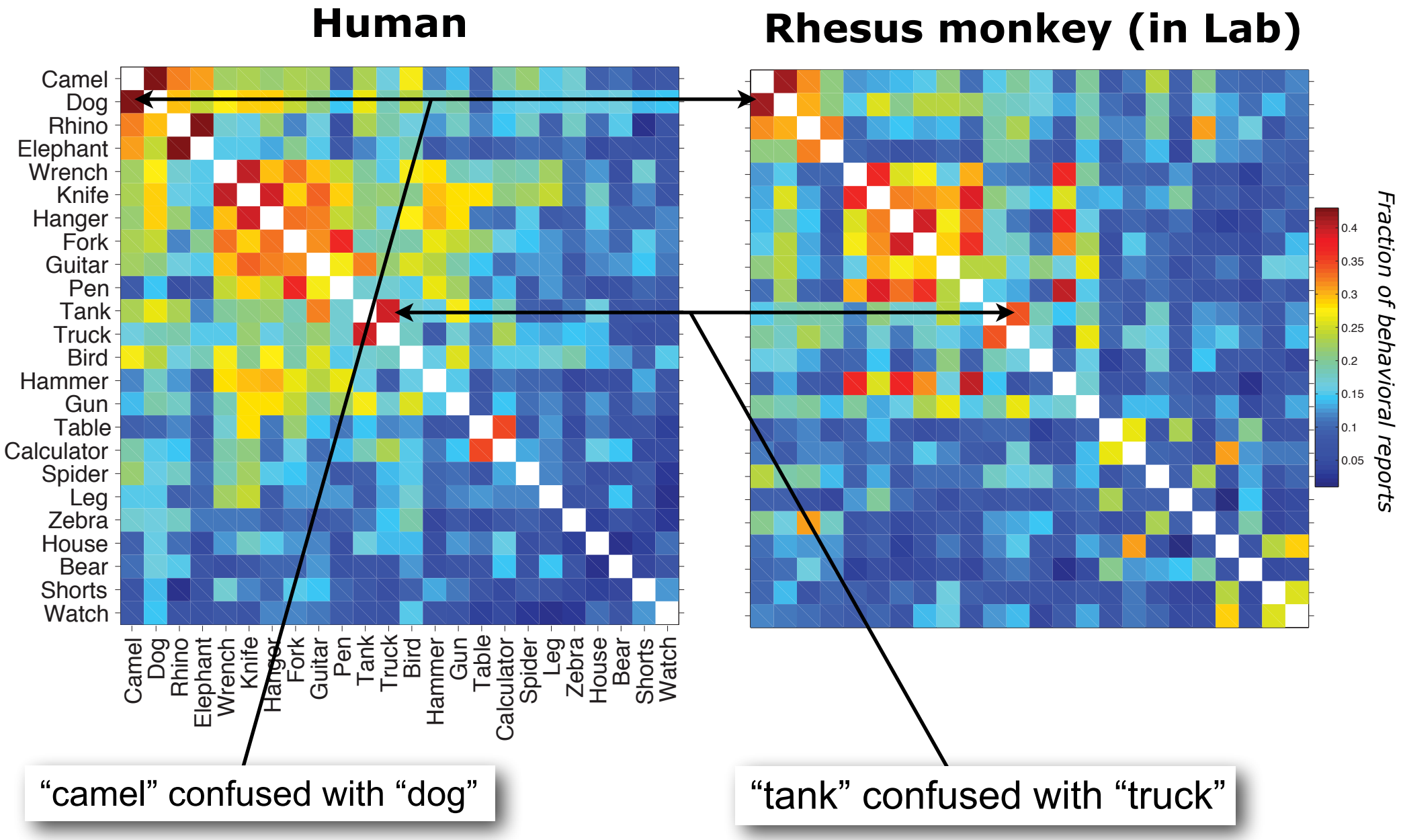


We aim to selectively disrupt processing here

When does this solution depend on feedback?

We can monitor solution up here at high temporal resolution

Background: **Monkey performance and pattern of object confusions on visual object recognition tasks is indistinguishable from human**



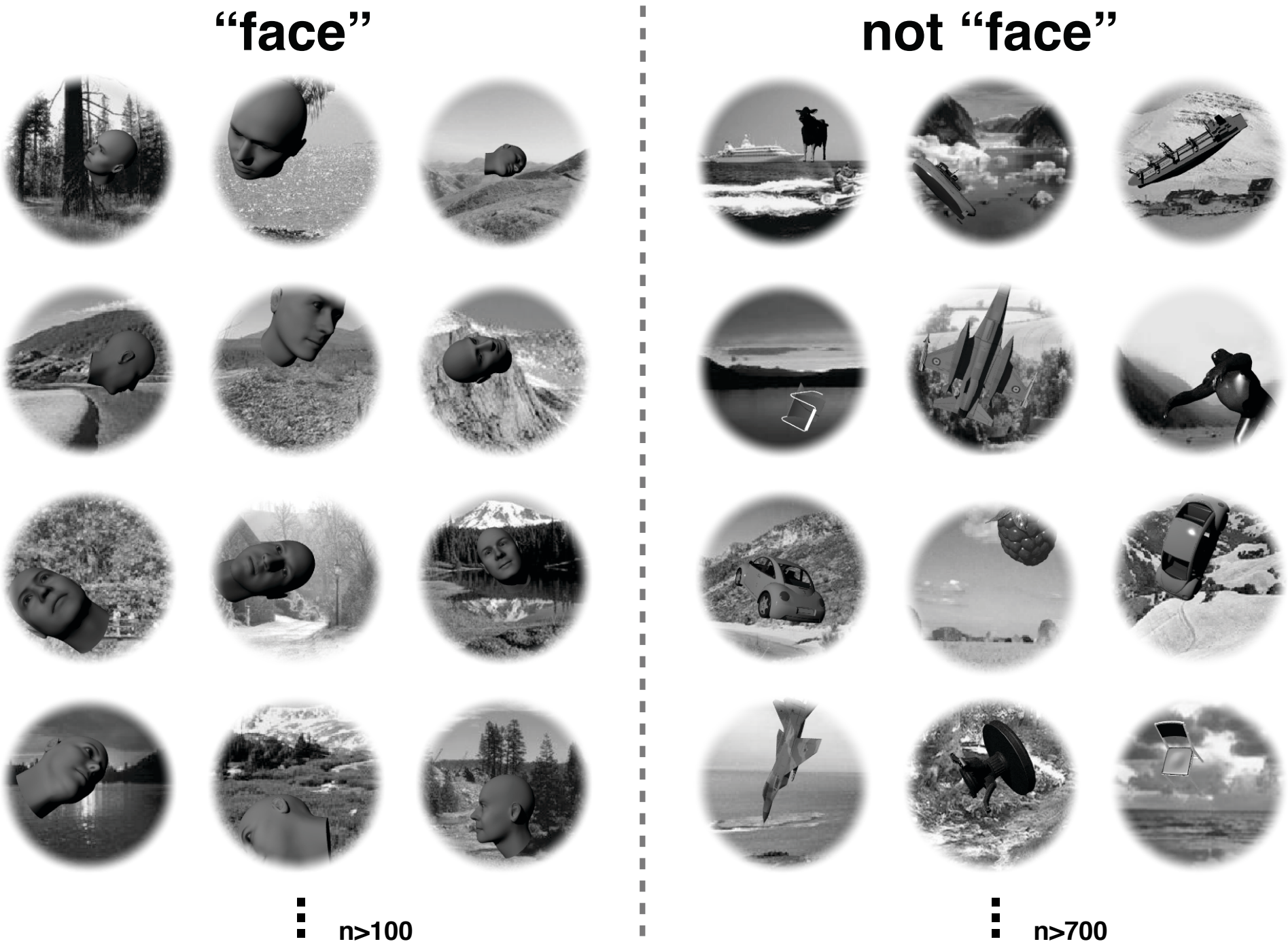
Comparison of Object Recognition Behavior in Human and Monkey
R. Rajalingham, K Schmidt, J.J. DiCarlo, ***J Neuroscience*** *(in press)*

Our prior recent work:

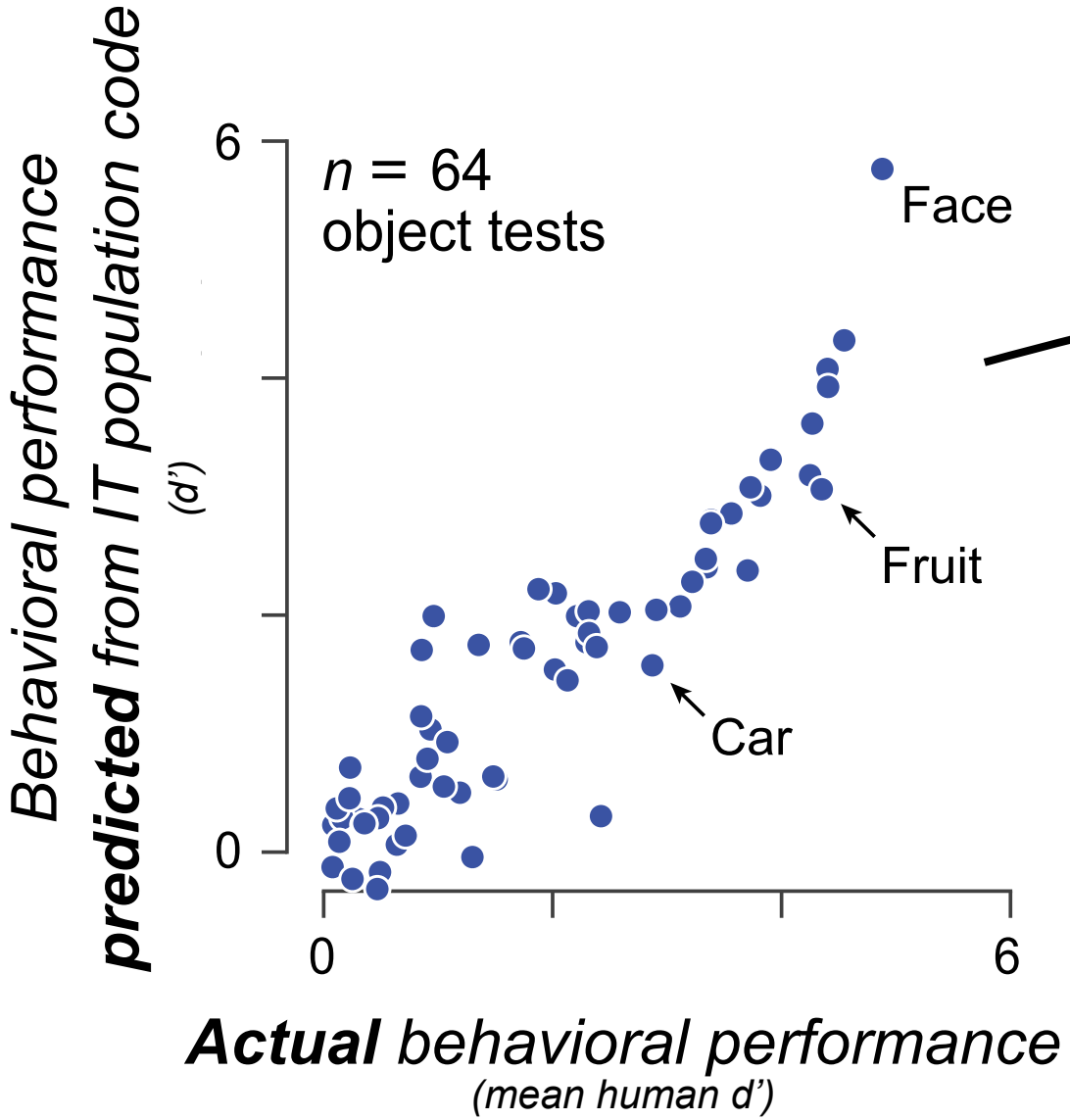


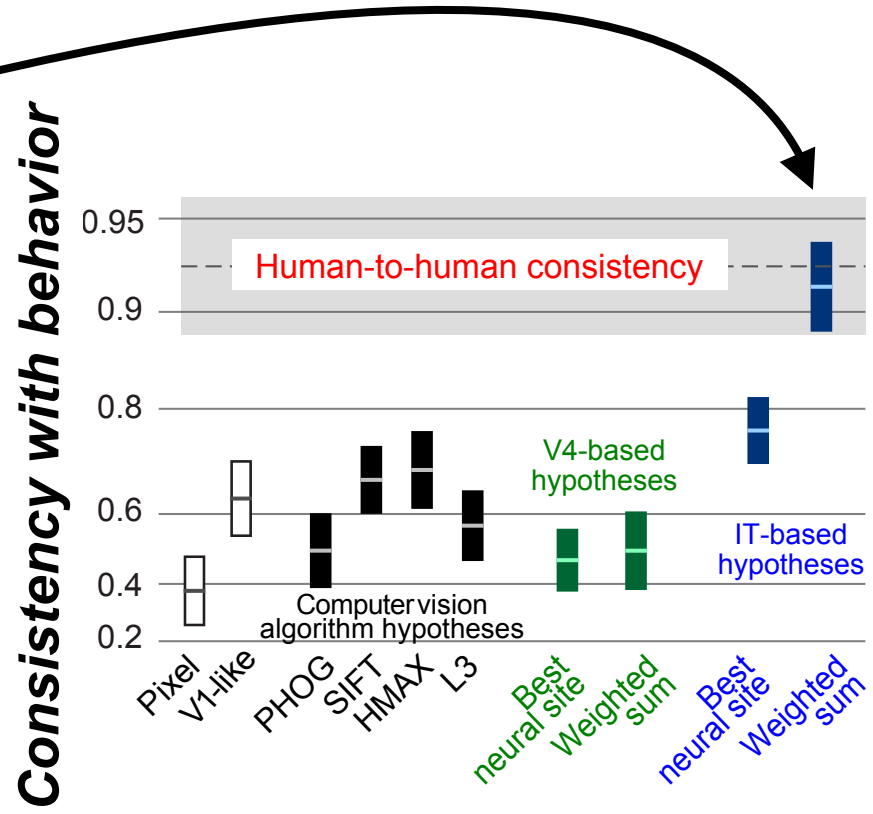
- generative space of images, each with a single experimenter-known foreground object and other latent parameters (e.g. position, pose).
- uncorrelated, new background every image

One example core object recognition test:



Background: **We can "read" IT population codes that are highly predictive of behavior**

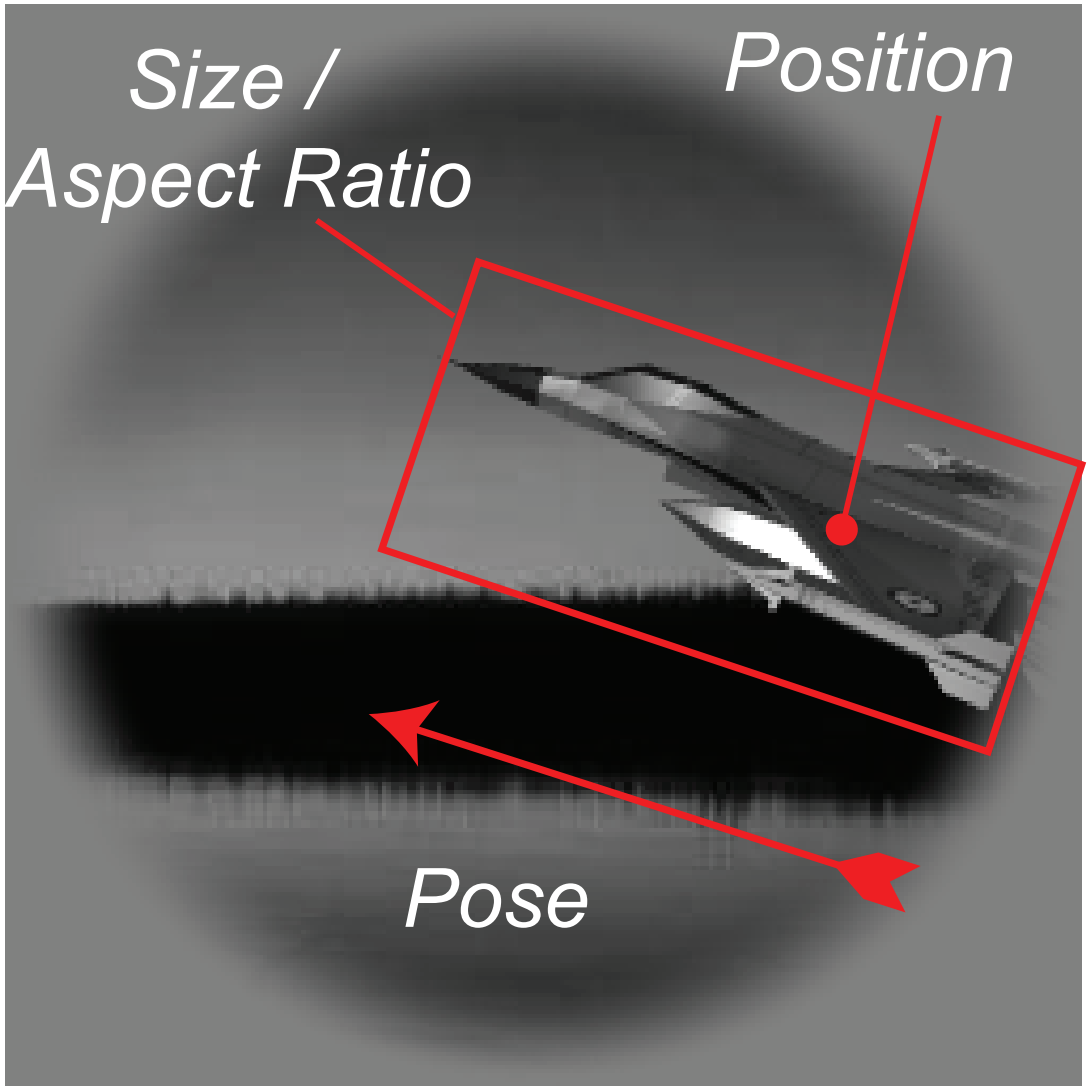


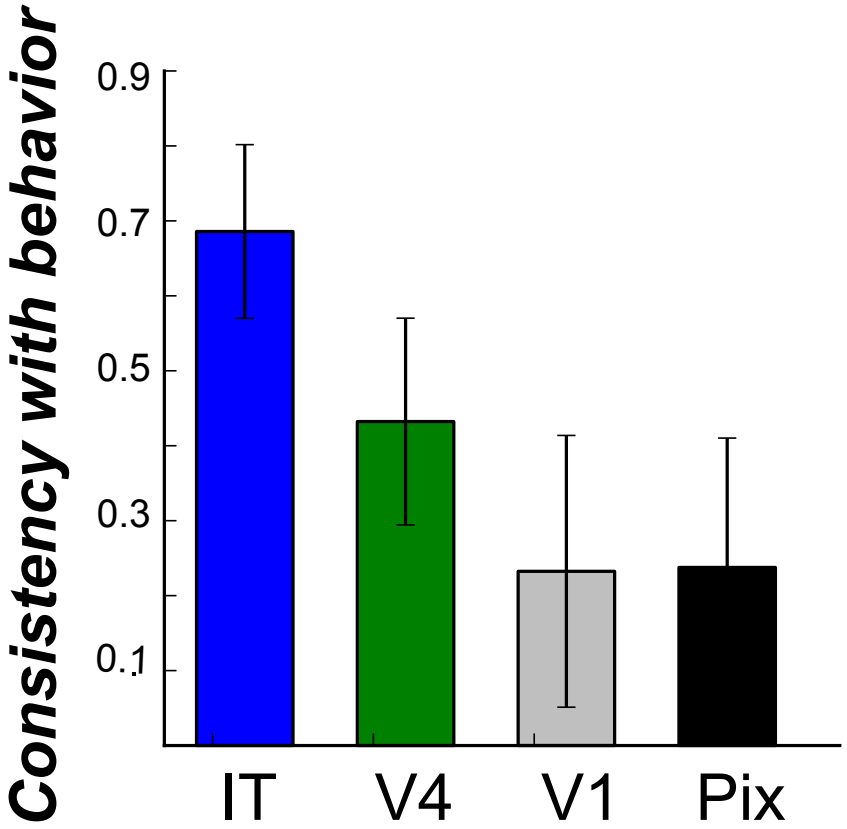


Background: **We can "read" IT population codes that are highly predictive of behavior**

IT codes also explain/predict behavior on other object latent variables

CV: "object detection"

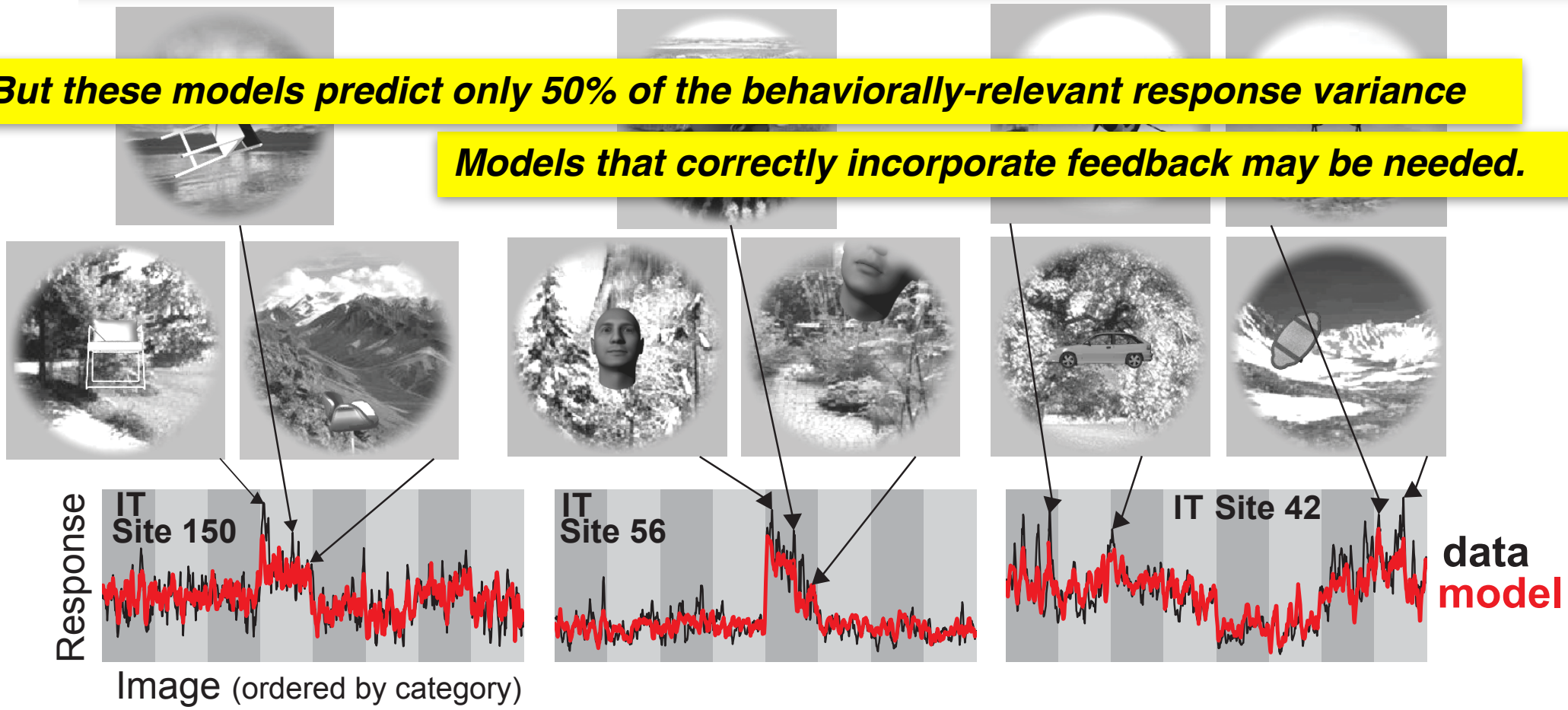




Background: We have shown that feedforward CNN models are very good predictors of ventral stream neuronal responses

But these models predict only 50% of the behaviorally-relevant response variance

Models that correctly incorporate feedback may be needed.



PNAS

Performance-optimized hierarchical models predict neural responses in higher visual cortex

Daniel L. K. Yamins[a,1], Ha Hong[a,b,1], Charles F. Cadieu[a], Ethan A. Solomon[a], Darren Seibert[a], and James J. DiCarlo[a,2]

[a]Department of Brain and Cognitive Sciences and McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139; and [b]Harvard-MIT Division of Health Sciences and Technology, Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA 02139

Edited by Terrence J. Sejnowski, Salk Institute for Biological Studies, La Jolla, CA, and approved April 8, 2014 (received for review March 3, 2014)

Overview of our ongoing MURI team plan

MURI Team: Find images who's solution likely depends on bidirectional computing

Step 1) Find images/tasks where human behavior better than state-of-the-art feedforward CV model behavior (Beck)

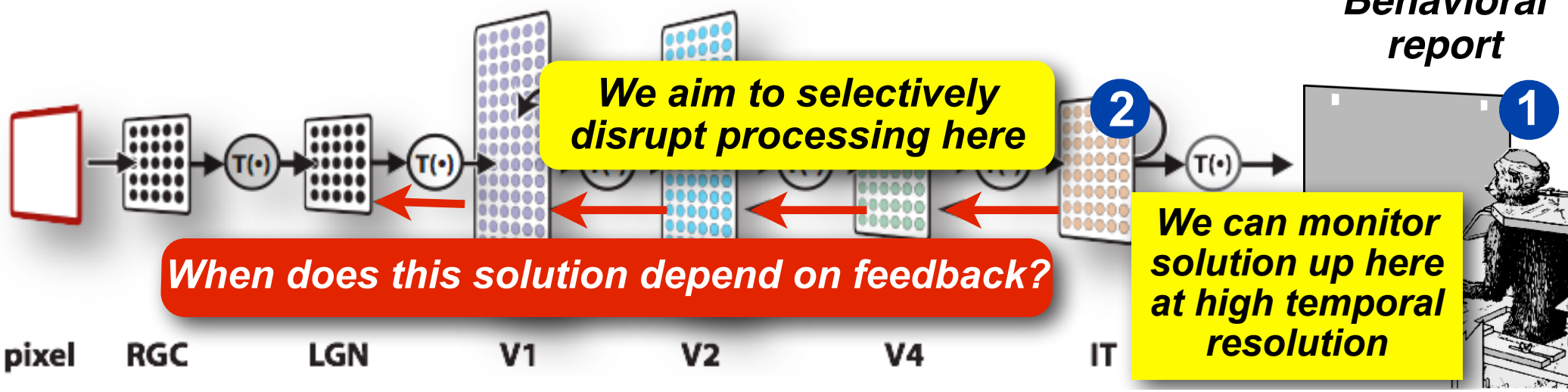
Check point: do those images show evidence of (slightly) longer than average processing time?

Step 2) Port same images and tasks to behaving monkey subjects so that we can study and manipulate neural circuits

DiCarlo Aim 0

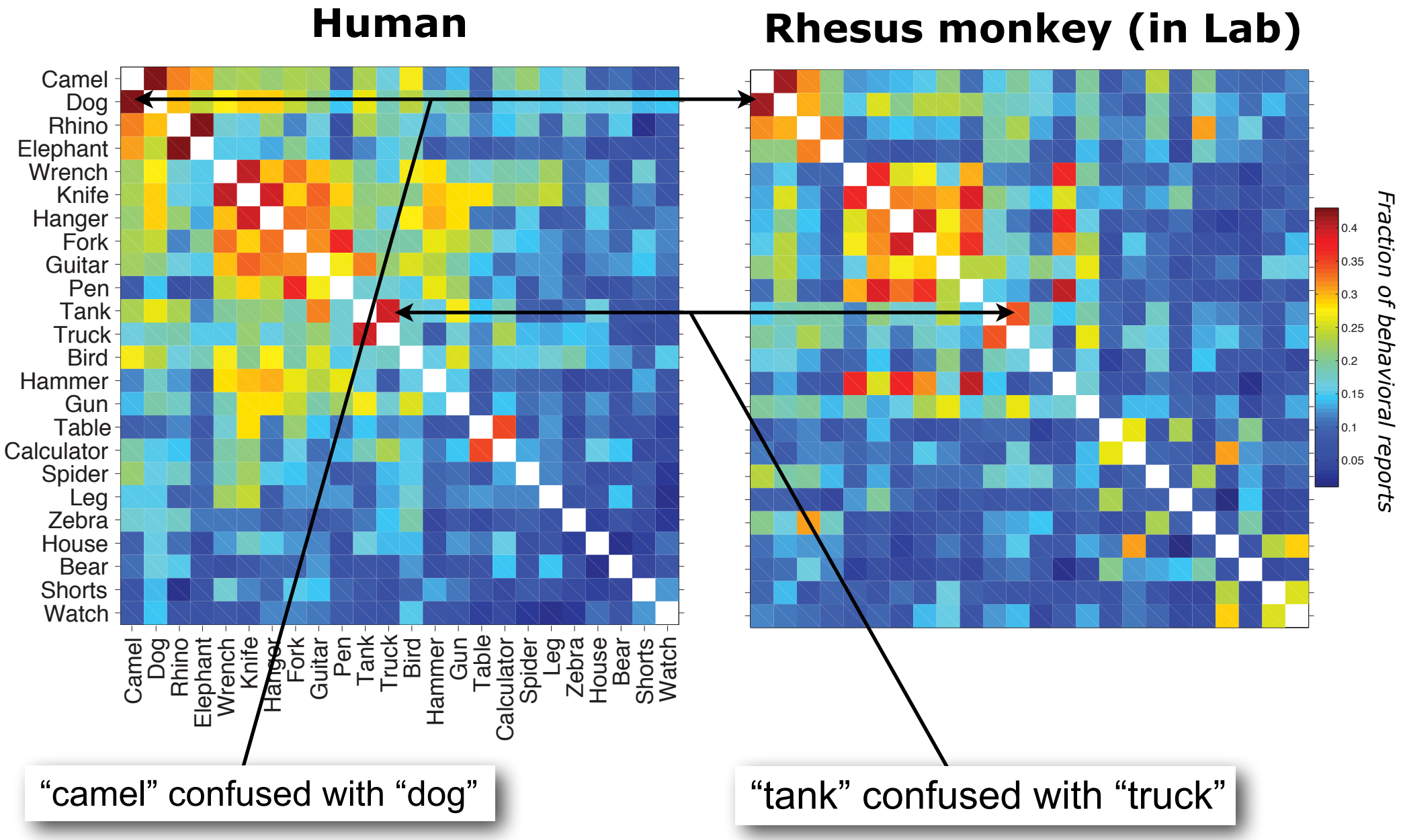
*Check point: do monkeys show same pattern of behavior as humans on each of these images? —> if not, **need** to further refine the images here.*

Step 3) Study and manipulate the relevant neural circuits in monkeys



DiCarlo Aim 0

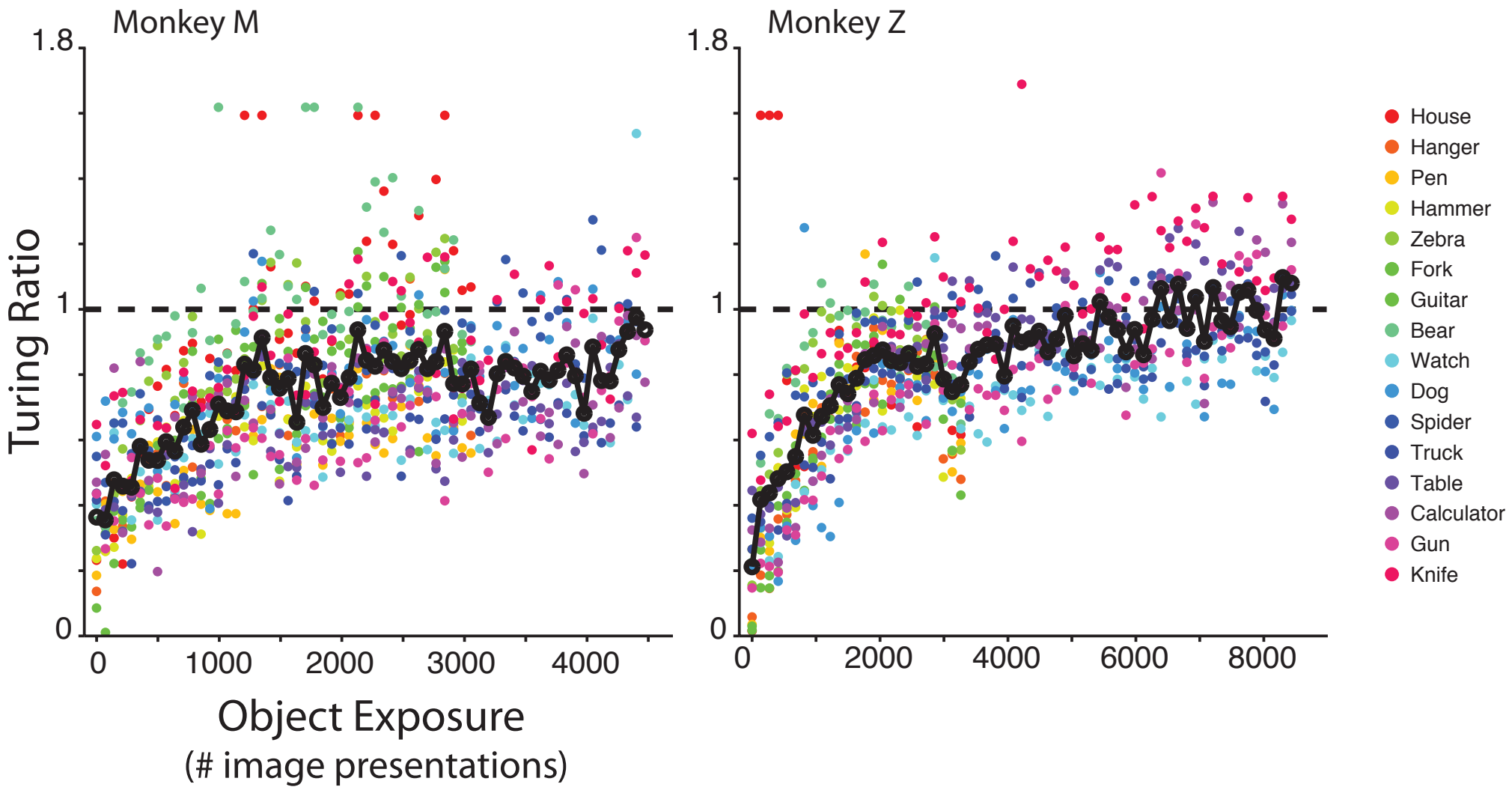
We have found that monkey performance and pattern of object confusions on visual object recognition tasks is indistinguishable from human



Comparison of Object Recognition Behavior in Human and Monkey
R. Rajalingham, K Schmidt, J.J. DiCarlo, ***J Neuroscience*** *(in press)*

DiCarlo Aim 0

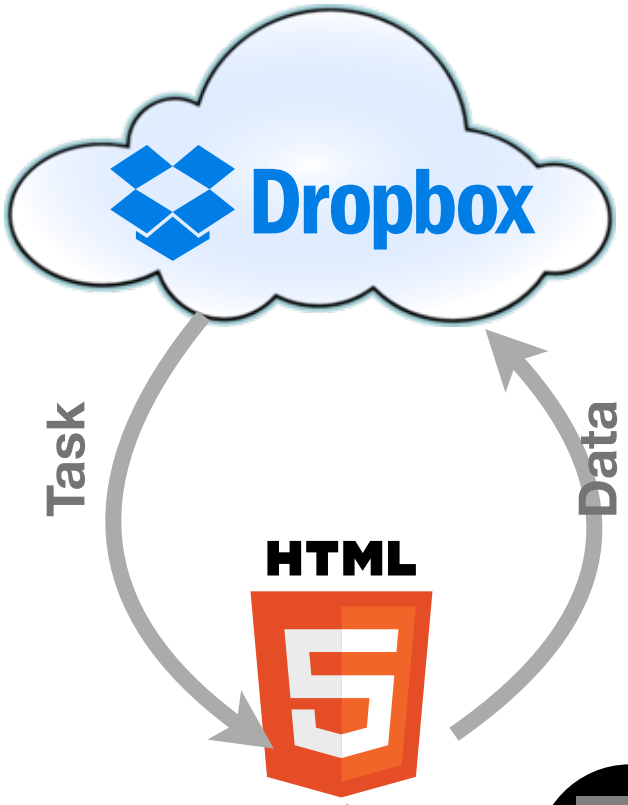
Monkey training to learn new objects



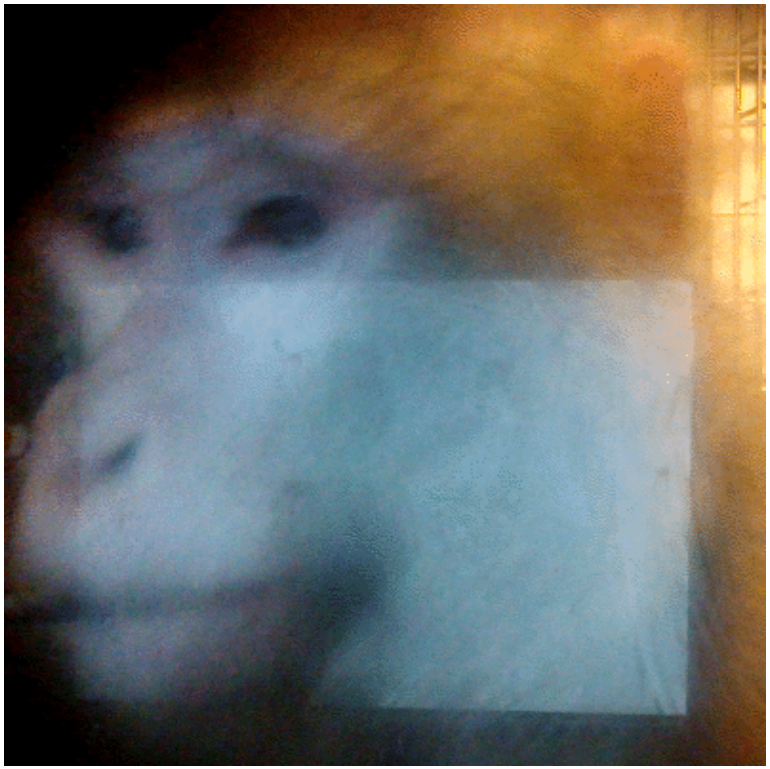
Comparison of Object Recognition Behavior in Human and Monkey
R. Rajalingham, K Schmidt, J.J. DiCarlo, ***J Neuroscience*** *(in press)*

DiCarlo
Aim 0

Higher-throughput monkey psychophysics



- Wireless
- Battery powered
- Automated
- Low cost
- Scalable



Match-to-Sample

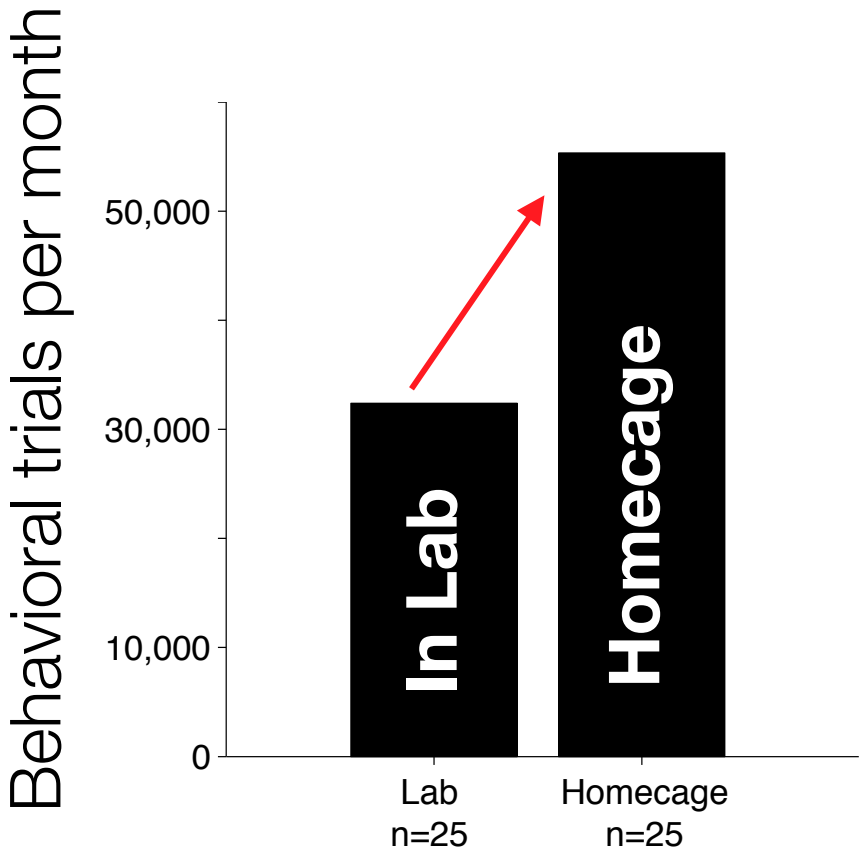


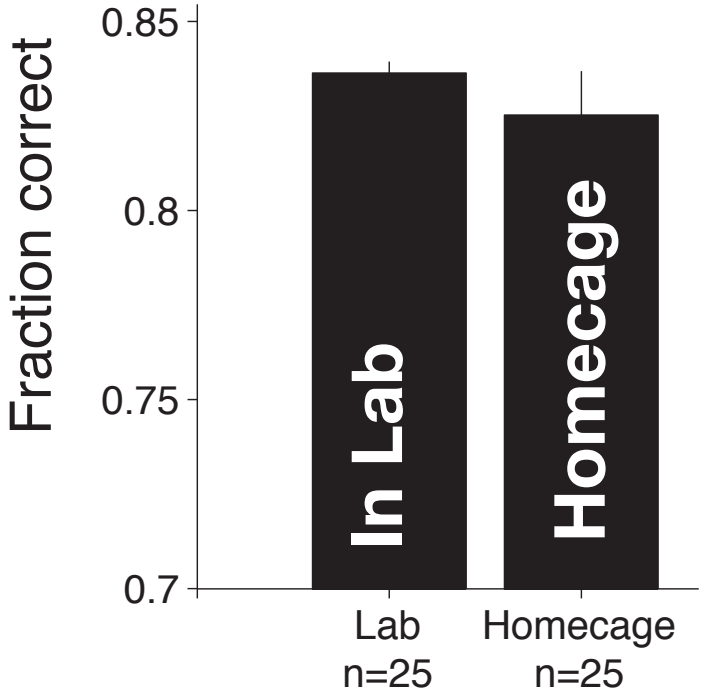




DiCarlo
Aim 0

Higher-throughput monkey psychophysics



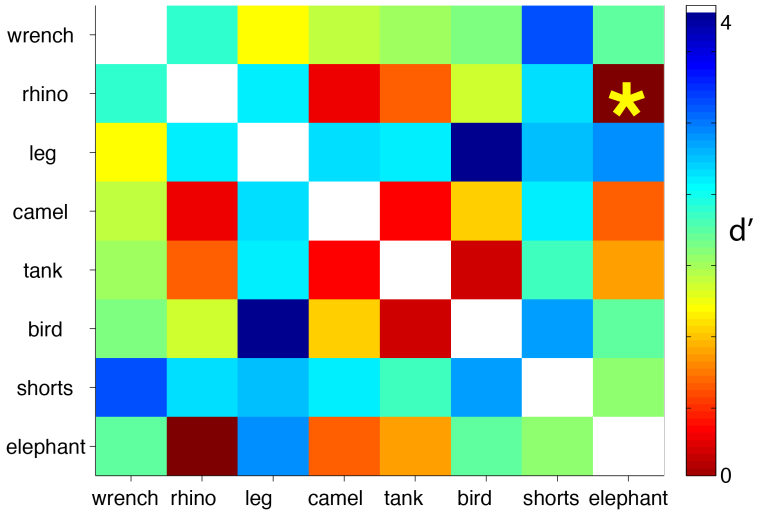


DiCarlo Aim 0

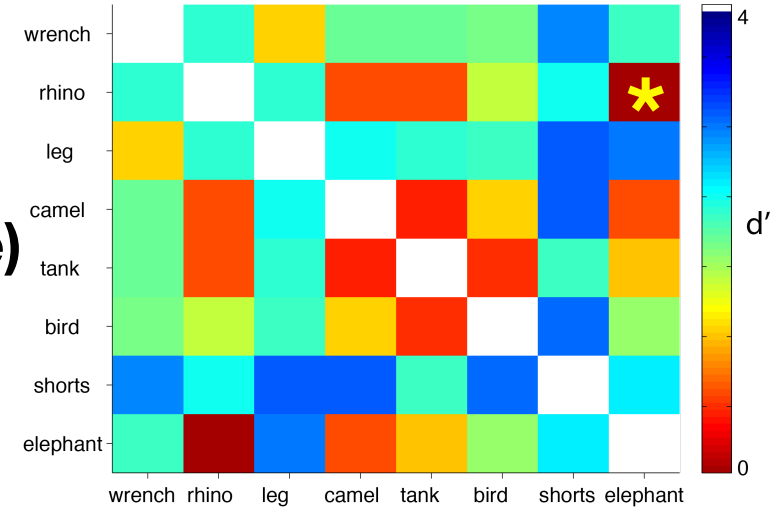
Higher-throughput monkey psychophysics

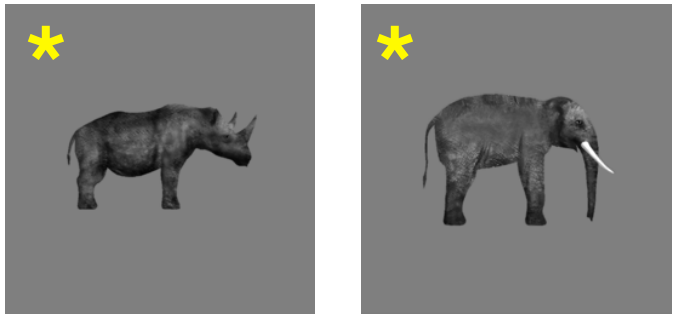
8x8 confusion matrices

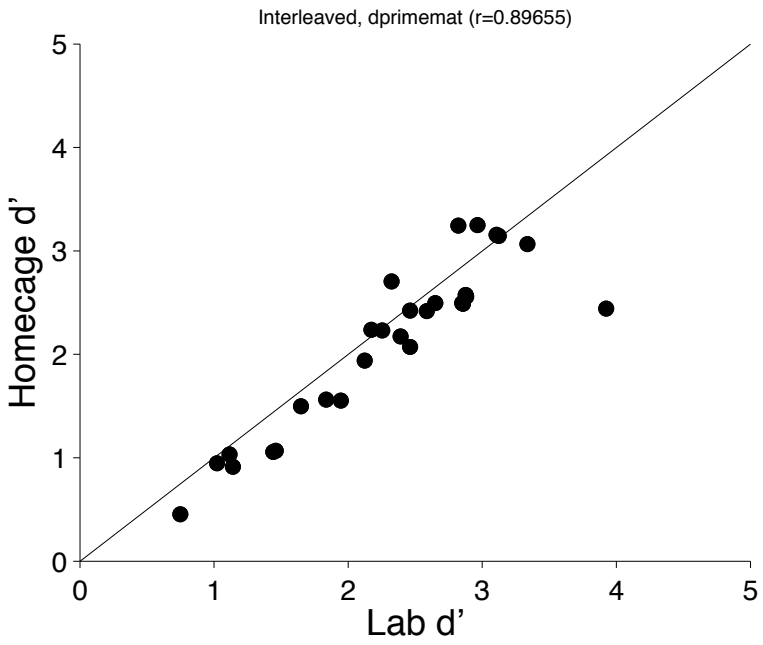
Monkey (In Lab)



Monkey (Homecage)

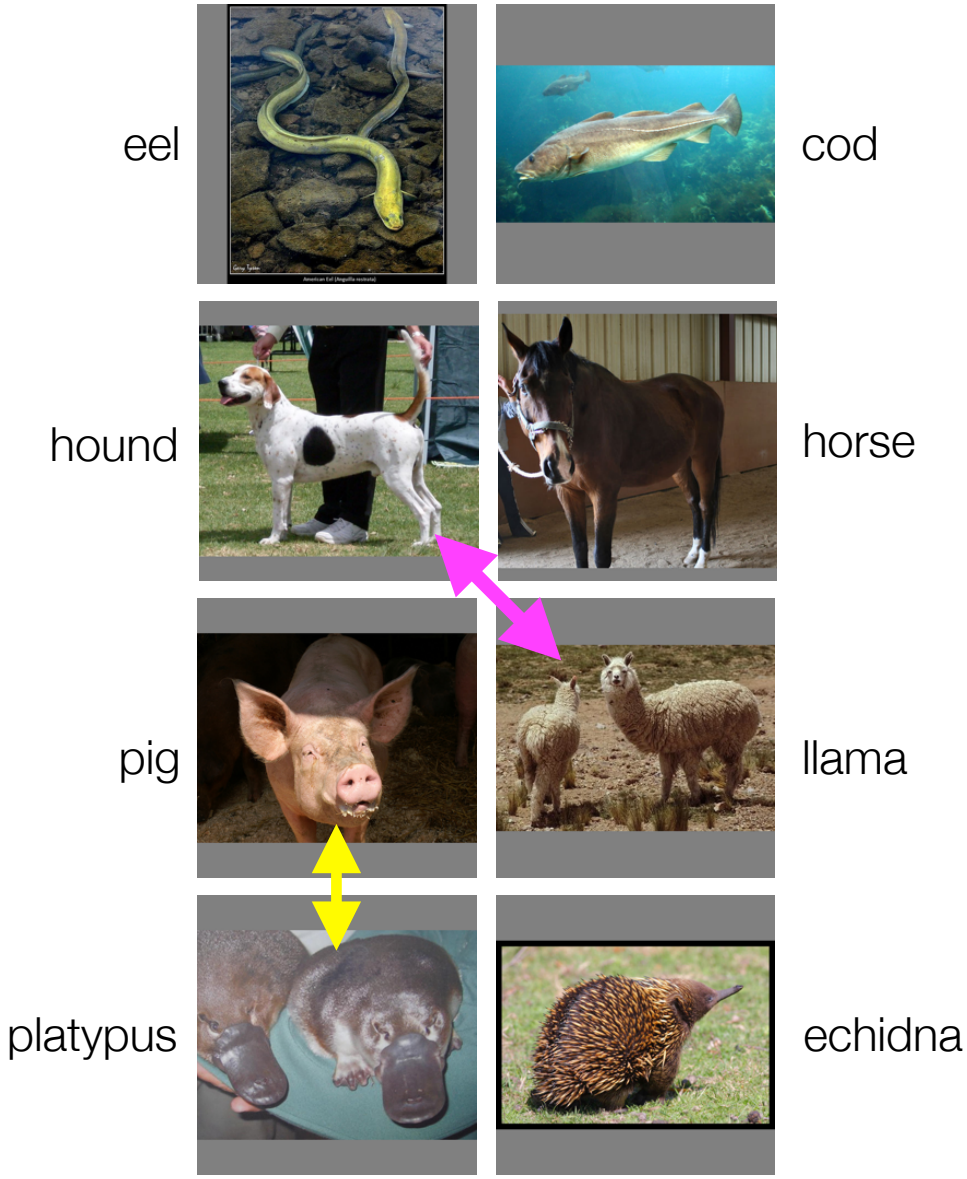




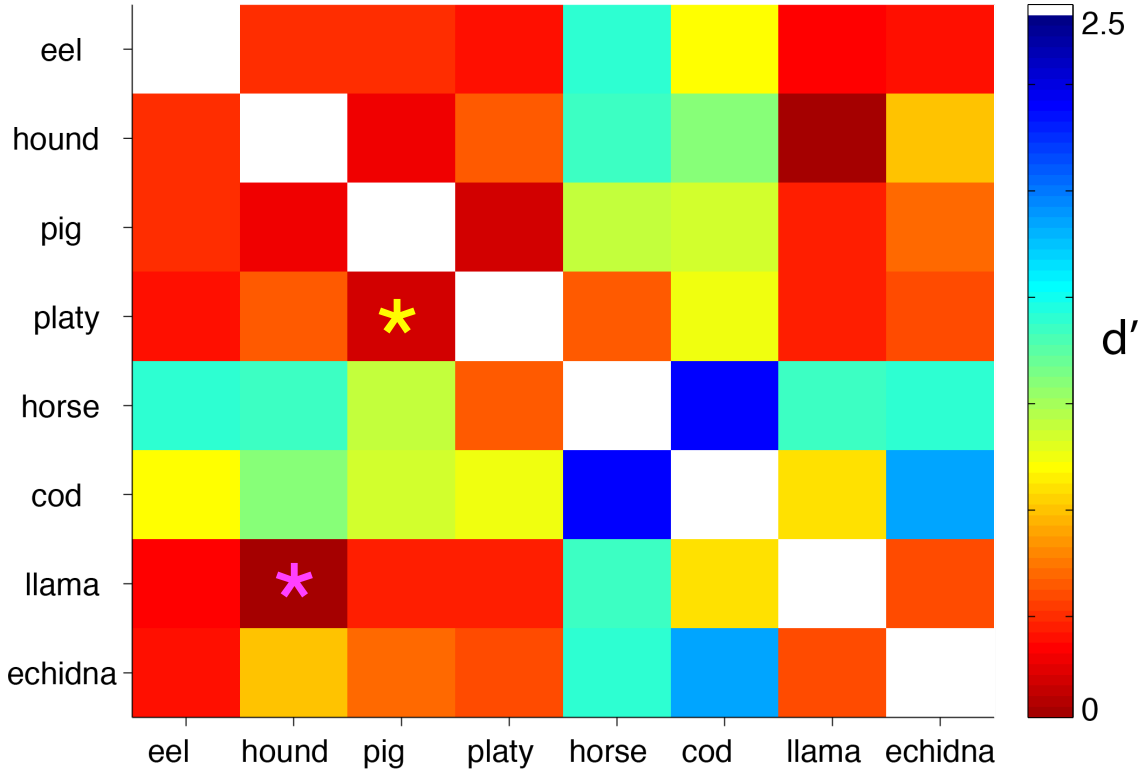


DiCarlo
Aim 0

Monkey behavior on ImageNet images (from Beck)



8 categories, 112 images
3 days, ~10,000 trials
200 ms presentation, 2-way choice



Overview of what we are contributing to the MURI team effort (Aim 1)

MURI Team: Find images who's solution likely depends on bidirectional computing

Step 1) Find images/tasks where human behavior better than state-of-the-art feedforward CV model behavior (Beck)

Check point: do those images show evidence of longer than average processing time?

Step 2) Port same images and tasks to behaving monkey subject n study and manipulate neural circuits

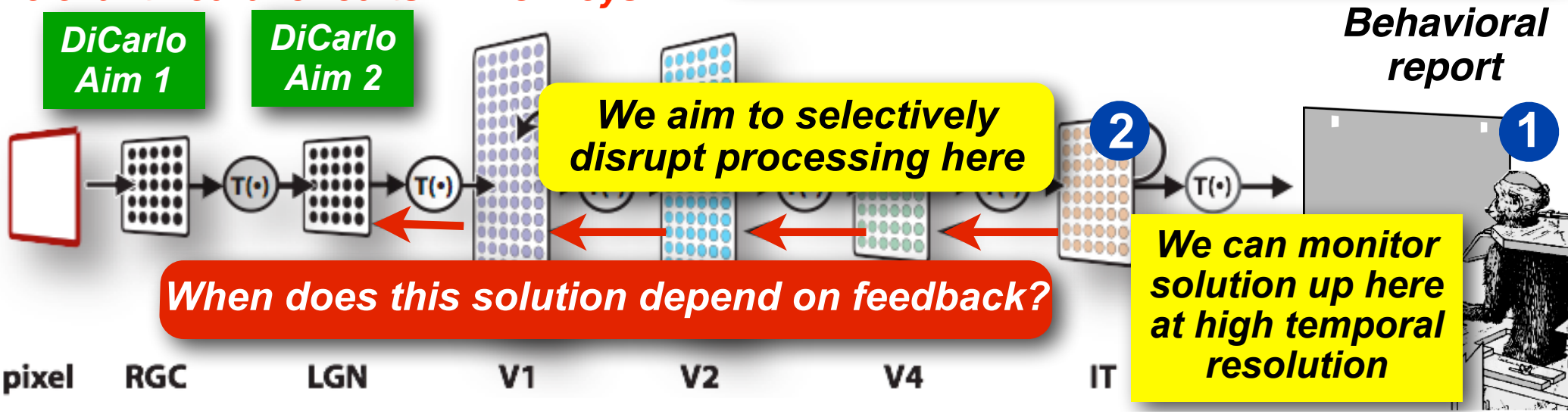
DiCarlo Aim 0

Done/ready

Check point: do monkeys show same pattern of behavior as humans on each of these images?
*—> if not, **need** to further refine the images here.*

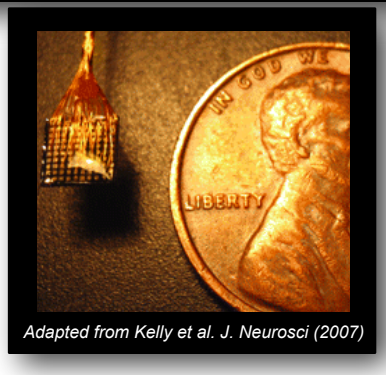
We need more data from humans and monkeys (our MURI team is working on this)
We plan to also test on "detection" tasks

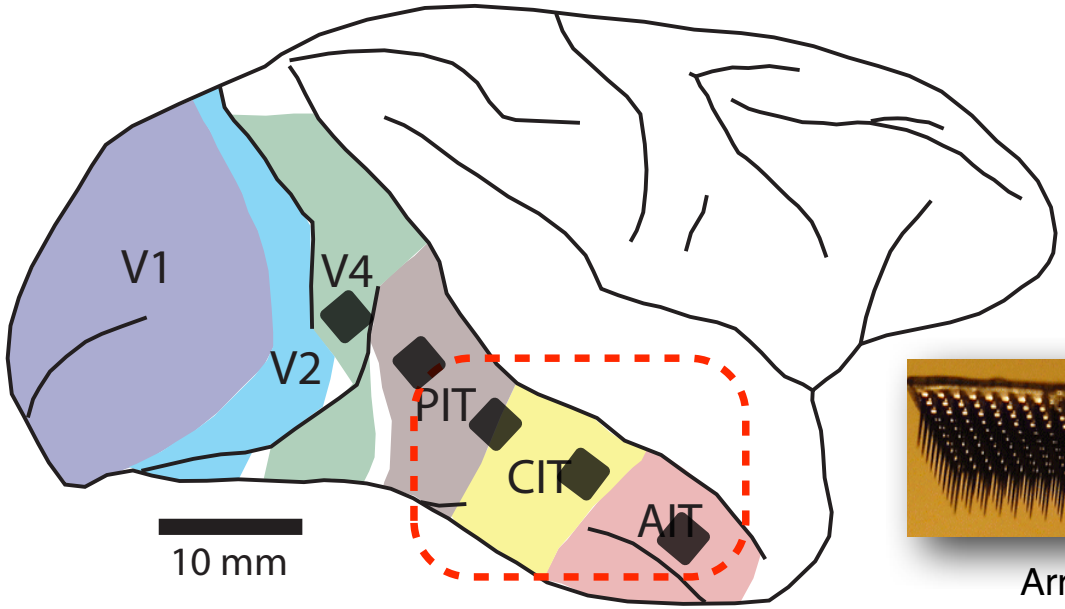
Step 3) Study and manipulate the relevant neural circuits in monkeys



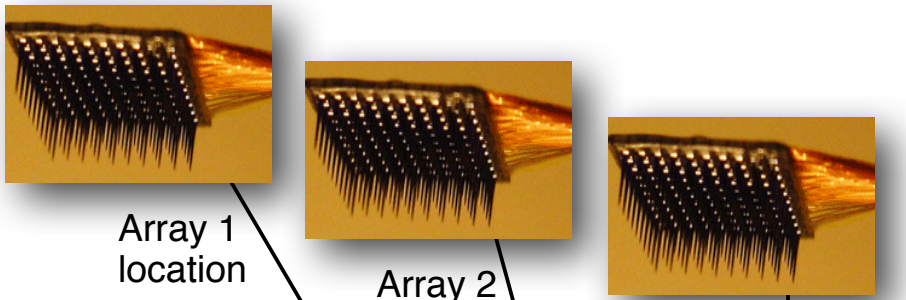
2.4.2 Expose a predicted signature of bidirectional computing by measuring the temporal evolution of the IT population representation of objects (DiCarlo, Beck)

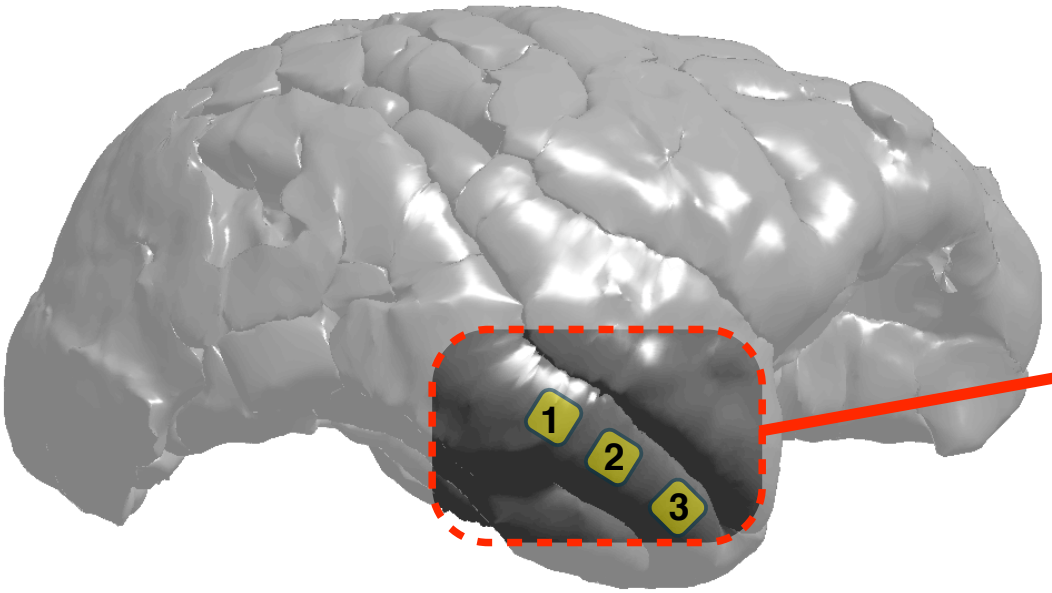
Plan: simultaneous recording of hundreds of neural sites in IT

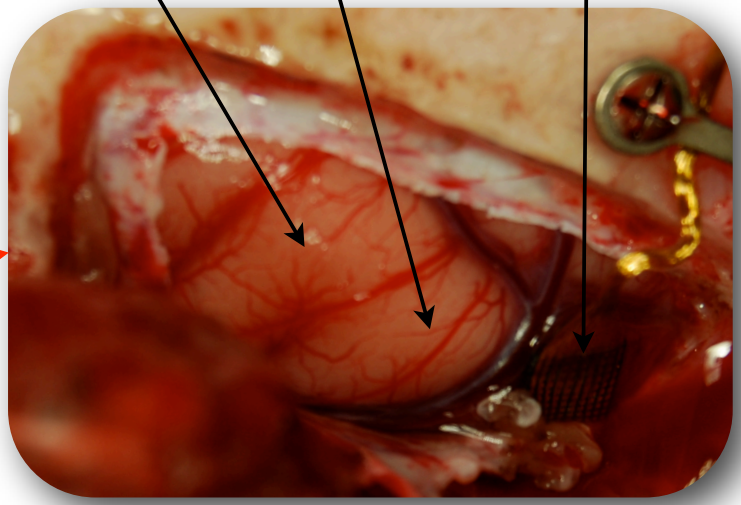




Three, 96-electrode arrays





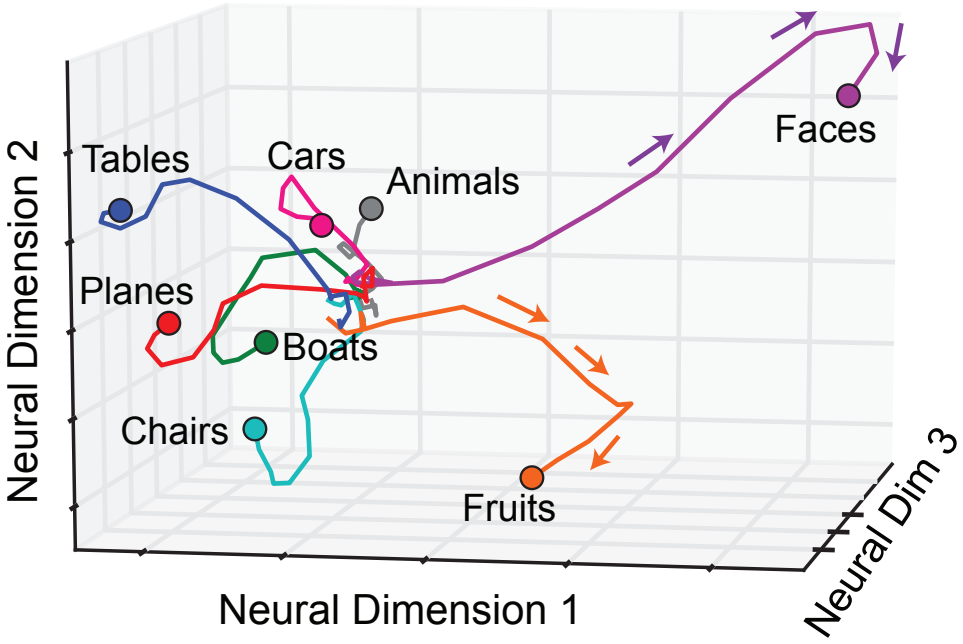


DiCarlo Aim 1

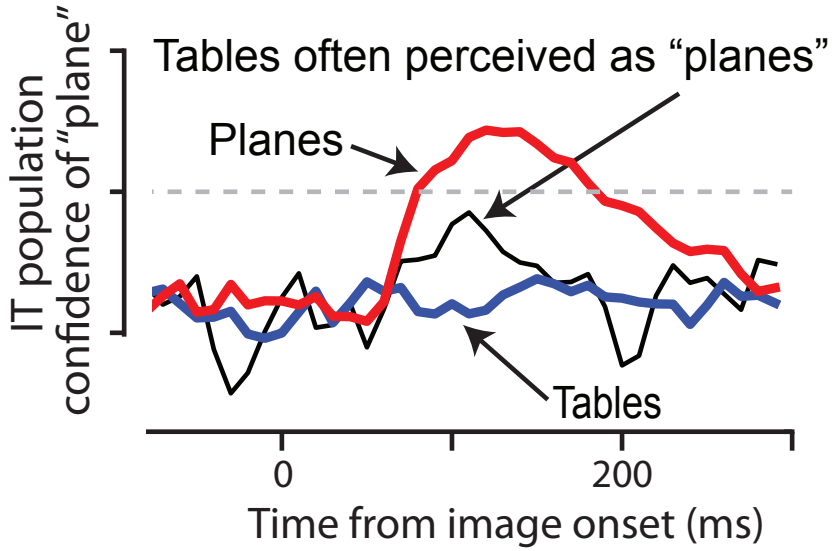
2.4.2 Expose a predicted signature of bidirectional computing by measuring the temporal evolution of the IT population representation of objects (DiCarlo, Beck)

Analyze the temporal evolution of the IT population object code.

Dynamics of IT population in low-dimensional projection



Appropriately projected IT dynamics reveals trajectory deviations consistent with confusions



MURI logic: This approach gives us moment-to-moment access to the neural signals underlying these CV-challenging tasks. If some or all of these tasks require bidirectional computing, then the population signal should emerge later in time.

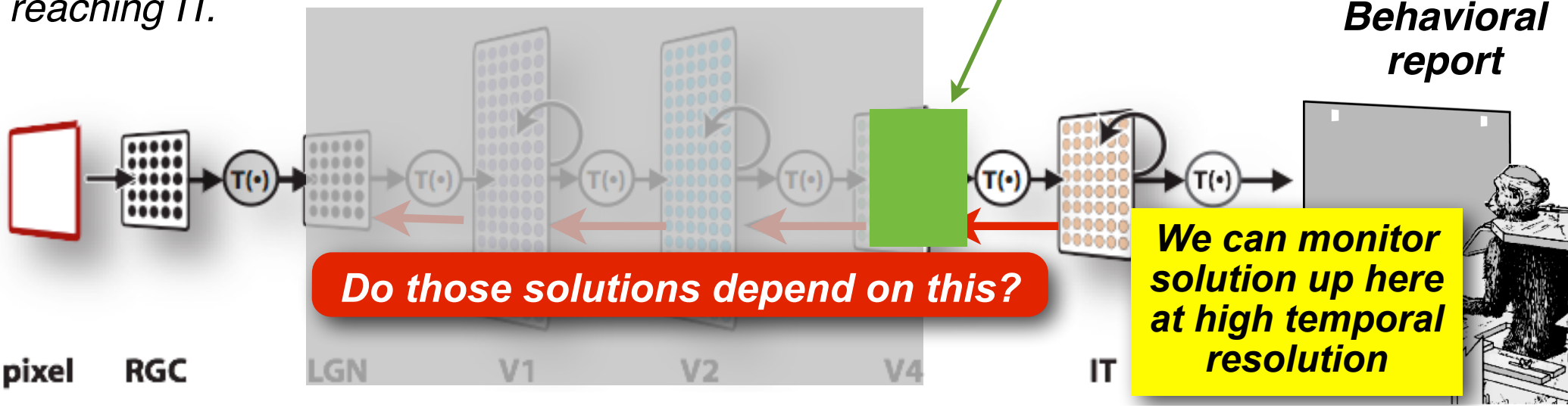
Progess/Status: We have trained two monkeys in object categorization tasks. We are refining tasks and images with the MURI team. Will start neural recording after that.

(MURI 2.4.4) Test causality: When do IT object representations depend on bidirectional computing? (DiCarlo) .

Approach 1: Temporally-delayed blockade of dominant cortical input to IT cortex

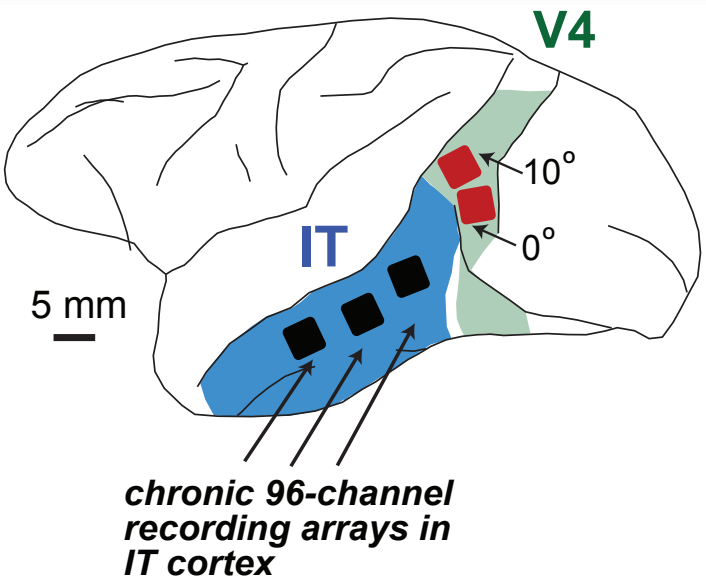
Apply optogentic blockade of central V4 or PIT at time lags after image onset on randomly-interleaved trials.

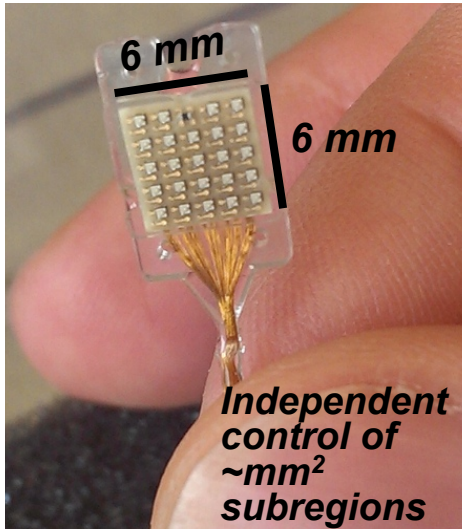
Logic: information built from cortical-cortical bidirectional computing is prevented from reaching IT.



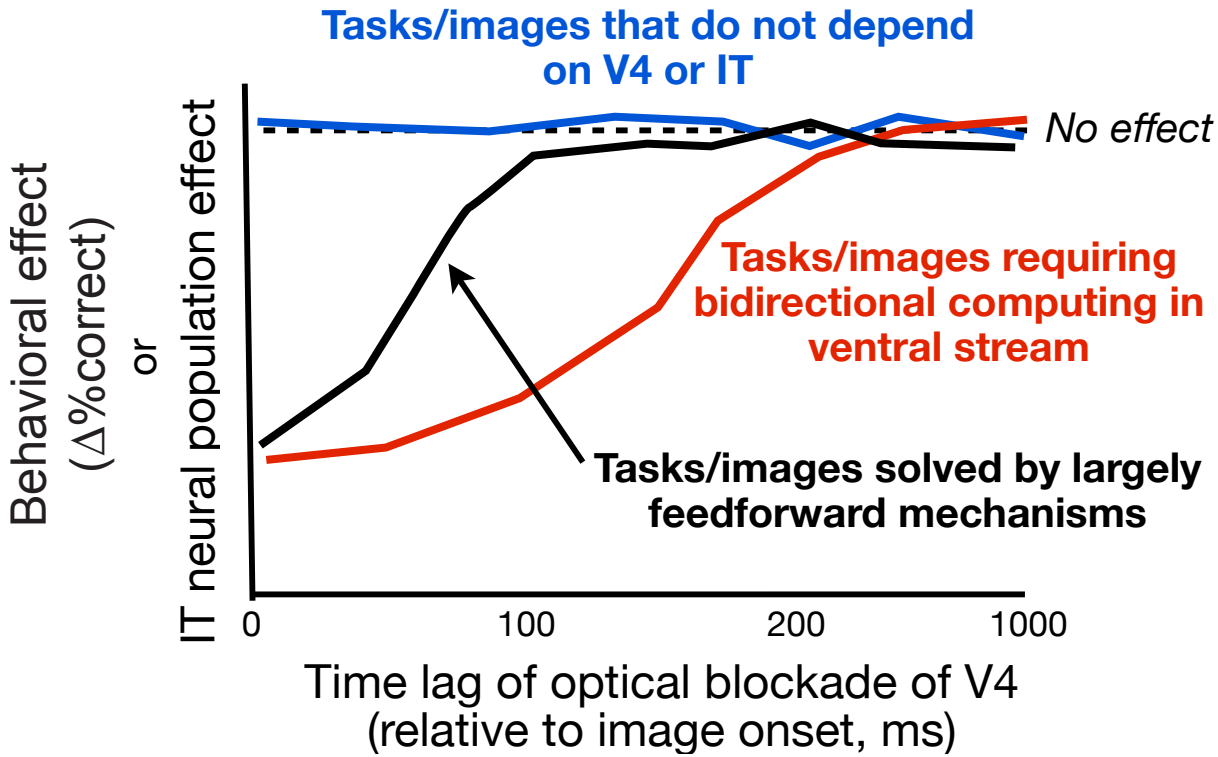
DiCarlo Aim 2

(MURI 2.4.4) Test causality: When do IT object representations depend on bidirectional computing? (DiCarlo) .





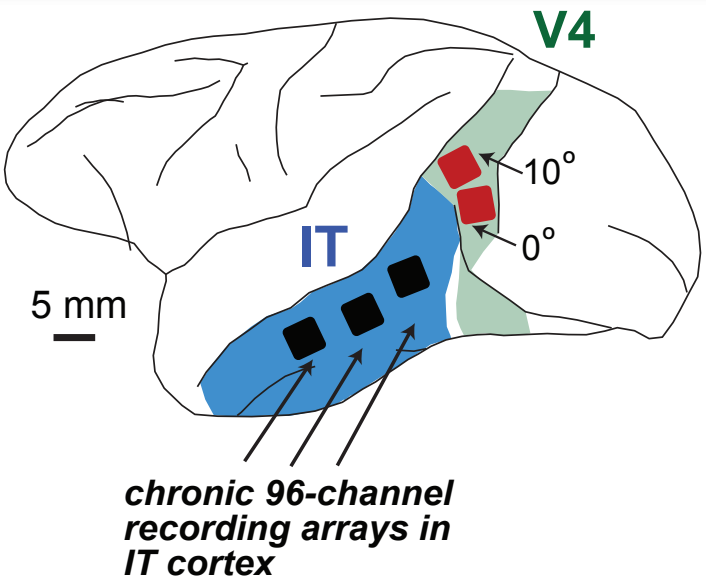
Expected form of our results:



MURI logic: For tasks and images that require bidirectional computing, we should disrupt IT and behavioral performance *(relative to reference tasks and images).*

DiCarlo Aim 2

(MURI 2.4.4) Test causality: When do IT object representations depend on bidirectional computing? (DiCarlo) .



Expected form of our results:

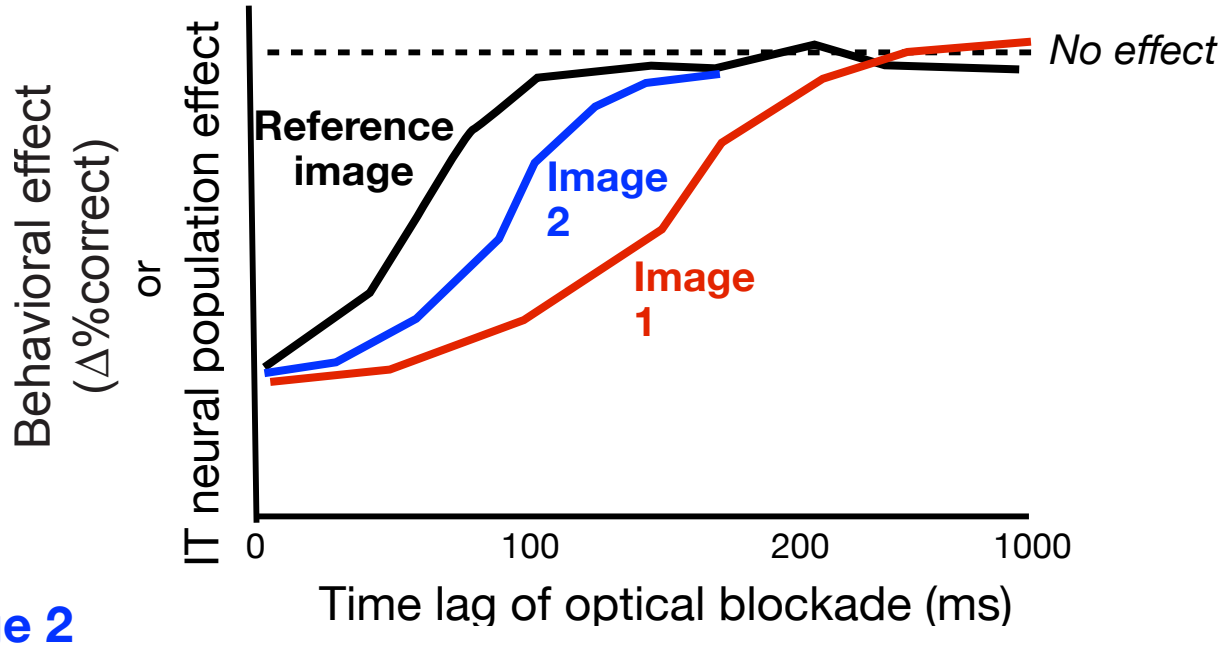


Image 1



Image 2



. . .

Goal to inform computer vision:

Given a task, measure the importance of bidirectional computing for each image

DiCarlo Aim 2

(MURI 2.4.4) Test causality: When do IT object representations depend on bidirectional computing? (DiCarlo) .

Technical steps and progress thus far:

Demonstrate that optogenetic control can:

- ***suppress ventral stream neurons with rapid temporal onset***
- ***produce predictable behavioral disruption***

Develop, test, and deploy a method for delivering light to large regions of cortex

- ***Chronic LED arrays: Design and construct***
  - ***Heat and power testing***
  - ***Test optogenetic control performance in simpler cortical area (V1)***
    - ***Surgical implantation techniques***
    - ***Demonstrate predictable behavioral suppression***
  - ***Test optogenetic control performance in ventral stream (V4, IT)***
    - ***Confirm neural suppression***
    - ***Confirm behavioral disruption***
    - ***Test behavioral disruption at different time lags*** *(relative to image onset)*
    - ***Test behavioral disruption at different time lags for MURI images***

Partial MURI goal using Approach 1.

Implant for both IT population recording and V4 optogenetic suppression

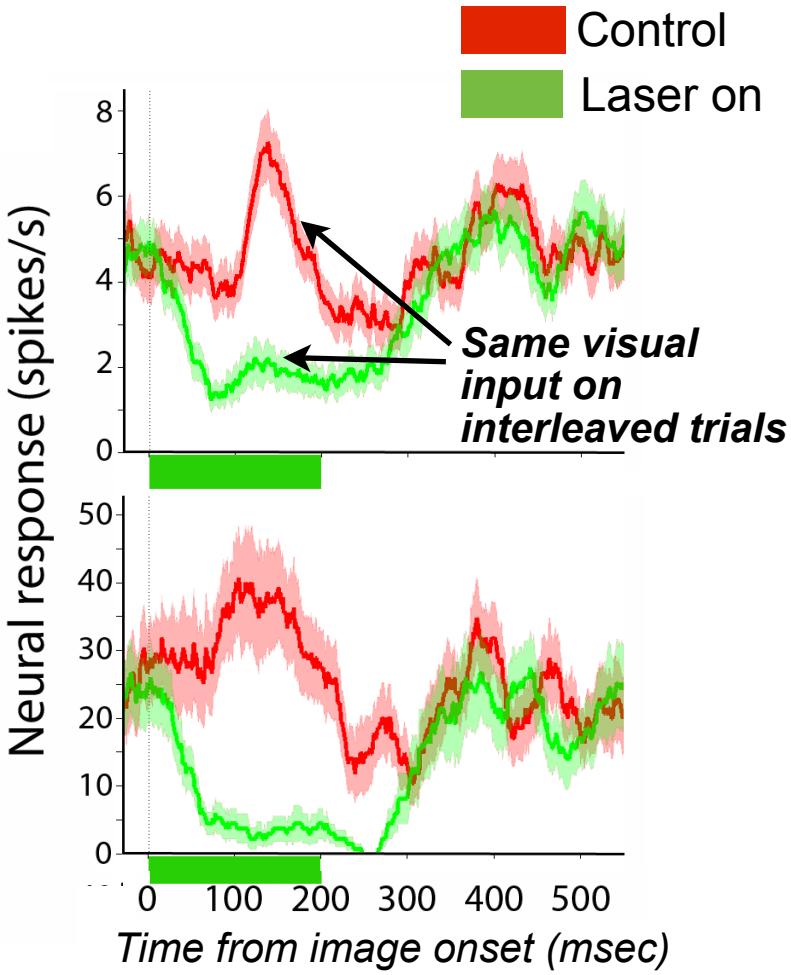
- ***Test behavioral and neural disruption at different time lags for MURI images***

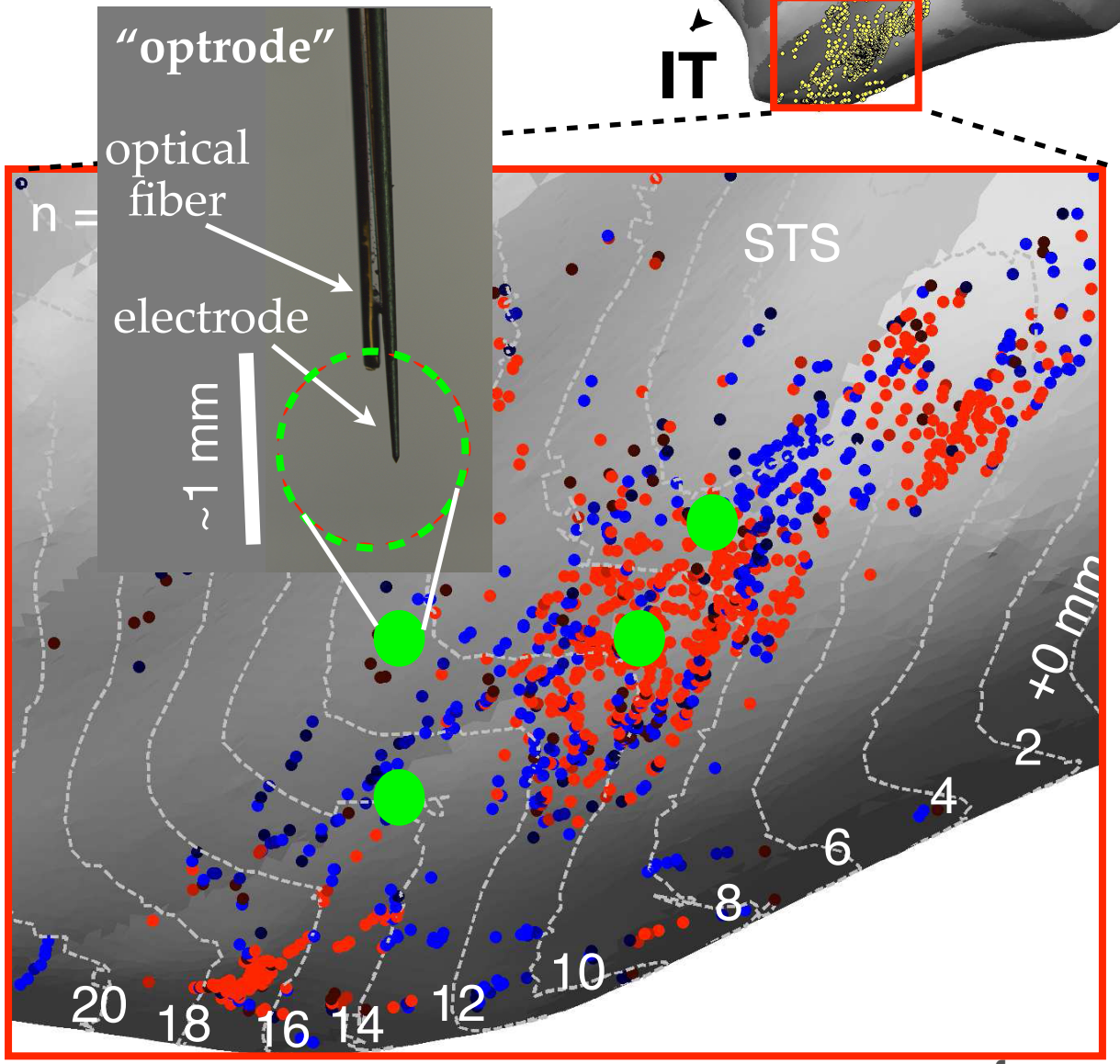
Full MURI goal using Approach 1.

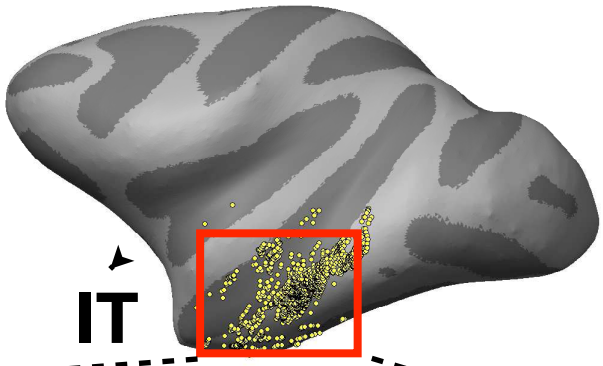
Demonstrate that optogenetic control can suppress ventral stream neurons with rapid temporal onset

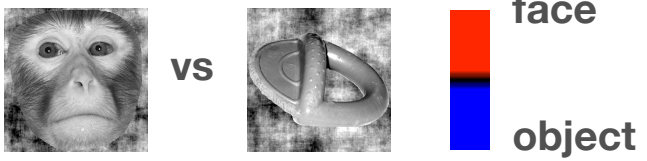
Example IT sites

Viral construct: AAV, CAG, ArchT



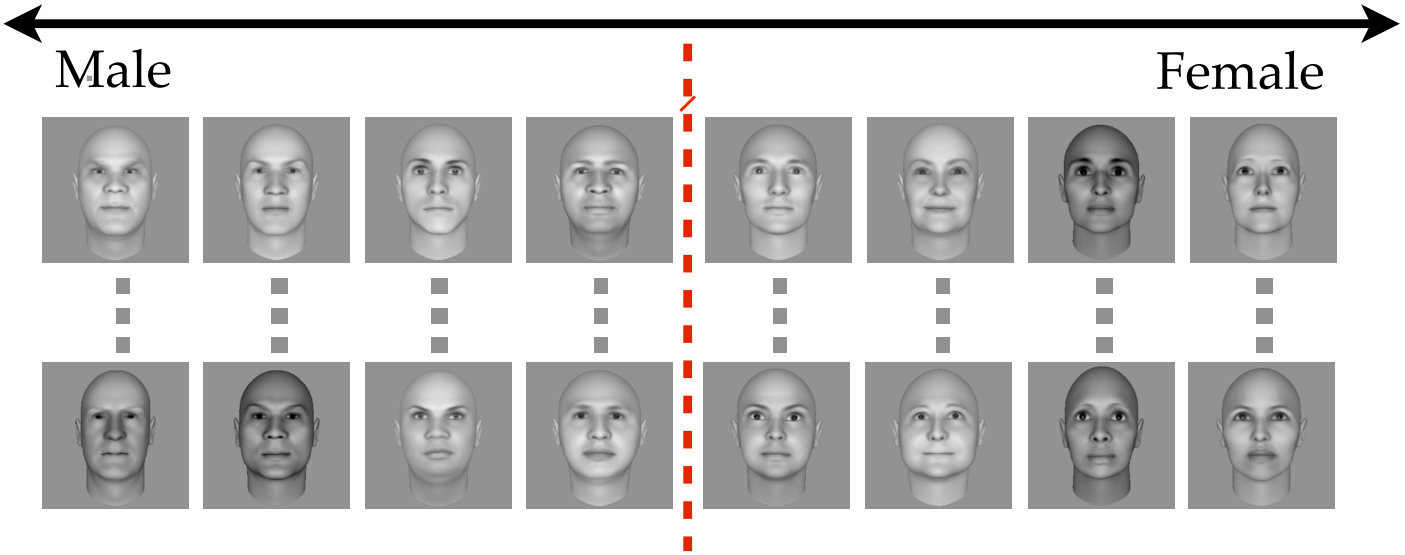


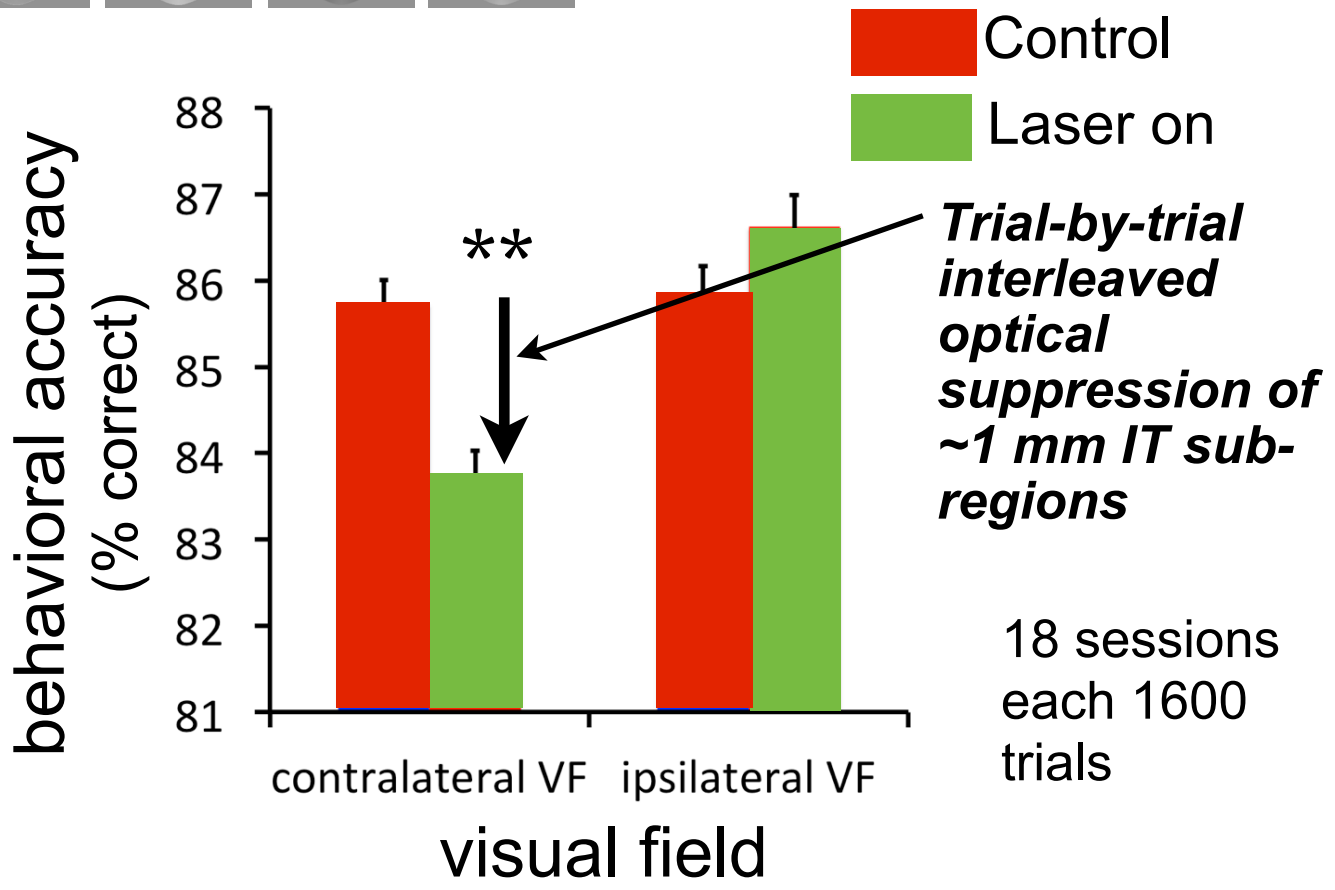




*Issa and DiCarlo, **J Neurosci** (2012)*

Demonstrate that optogenetic control can produce (predicted) behavioral disruption in object discrimination tasks



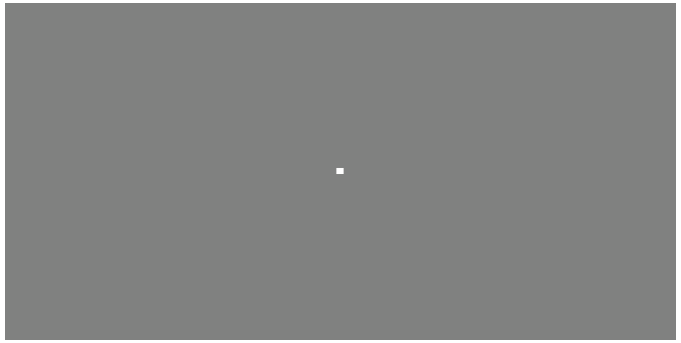


Trial-by-trial interleaved optical suppression of ~1 mm IT sub-regions

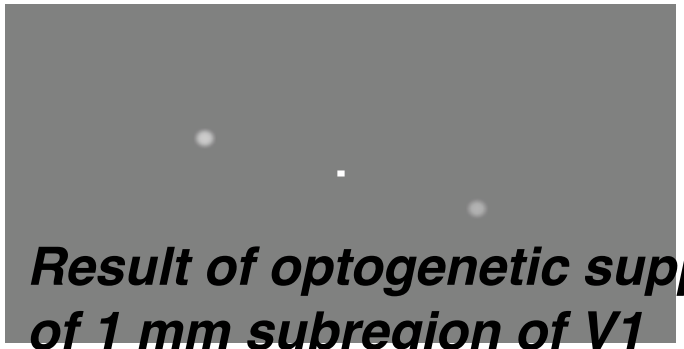
18 sessions each 1600 trials

Afraz, Boyden and DiCarlo, ***VSS*** *(2014)*
Afraz, Boyden and DiCarlo, ***PNAS*** *(2015)*

Demonstrate that optogenetic control can produce (predicted) behavioral disruption in object discrimination tasks



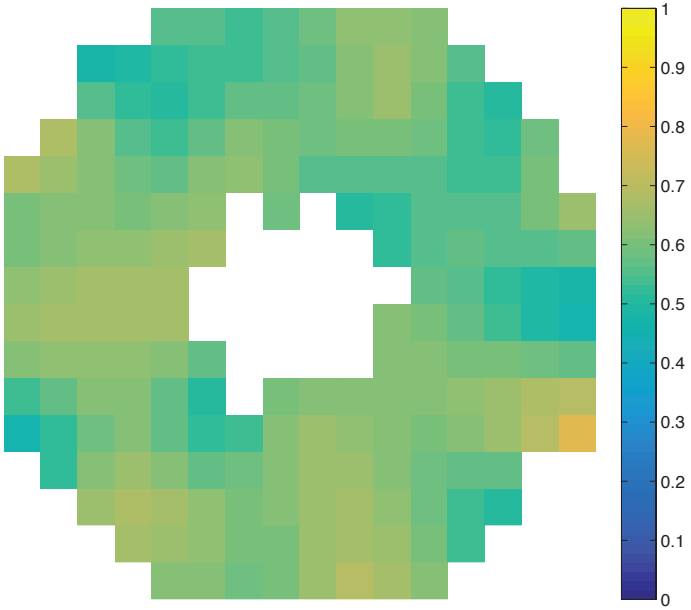
Fixation (200ms)

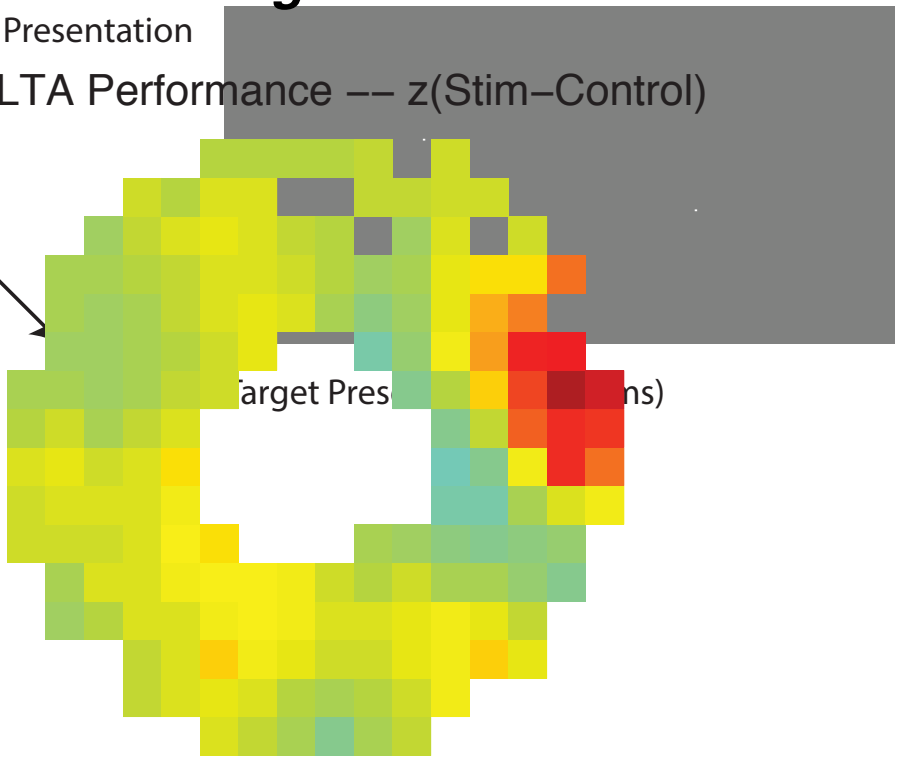


Stimulus Presentation (50ms)

Result of optogenetic suppression of 1 mm subregion of V1

Baseline behavioral performance is similar at all visual field locations





DiCarlo Aim 2 *(MURI 2.4.4) Test causality: When do IT object representations depend on bidirectional computing? (DiCarlo) .*

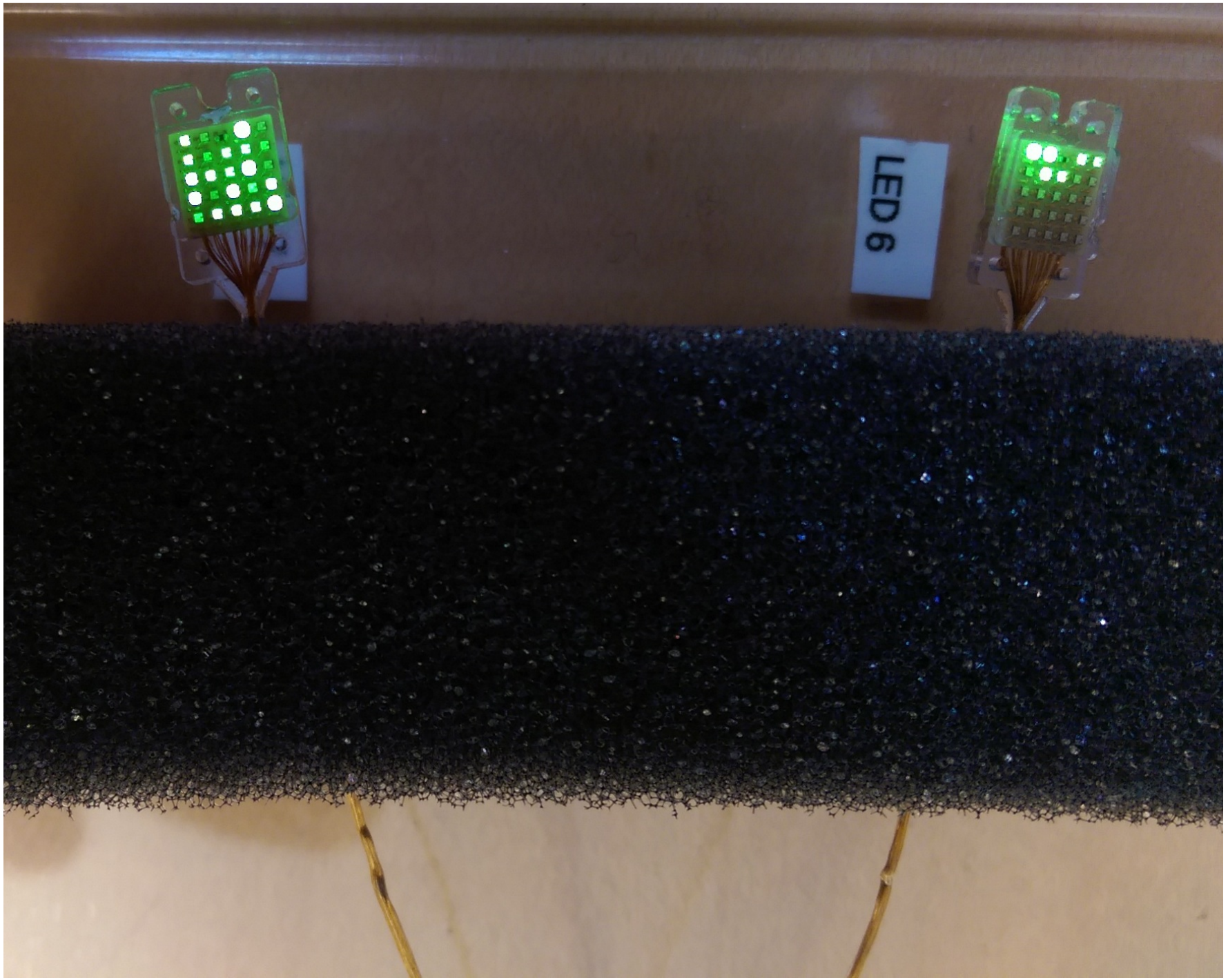
Technical steps and progress thus far:

Demonstrate that optogenetic control can: ✓
- ***suppress ventral stream neurons with rapid temporal onset*** ✓
- ***produce predictable behavioral disruption*** ✓

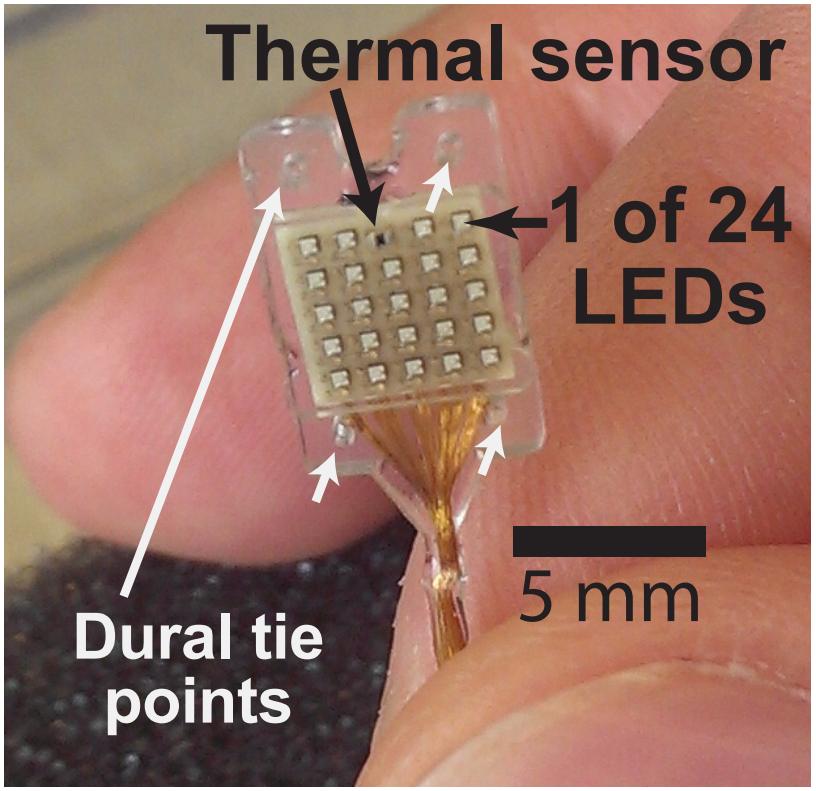
Develop, test, and deploy a method for delivering light to large regions of cortex
- ***Chronic LED arrays: Design and construct***
  - ***Heat and power testing***
  - ***Test optogenetic control performance in simpler cortical area (V1)***
    - ***Surgical implantation techniques***
    - ***Demonstrate predictable behavioral suppression***
  - ***Test optogenetic control performance in ventral stream (V4, IT)***
    - ***Confirm neural suppression***
    - ***Confirm behavioral disruption***
    - ***Test behavioral disruption at different time lags*** *(relative to image onset)*
    - ***Test behavioral disruption at different time lags for MURI images***

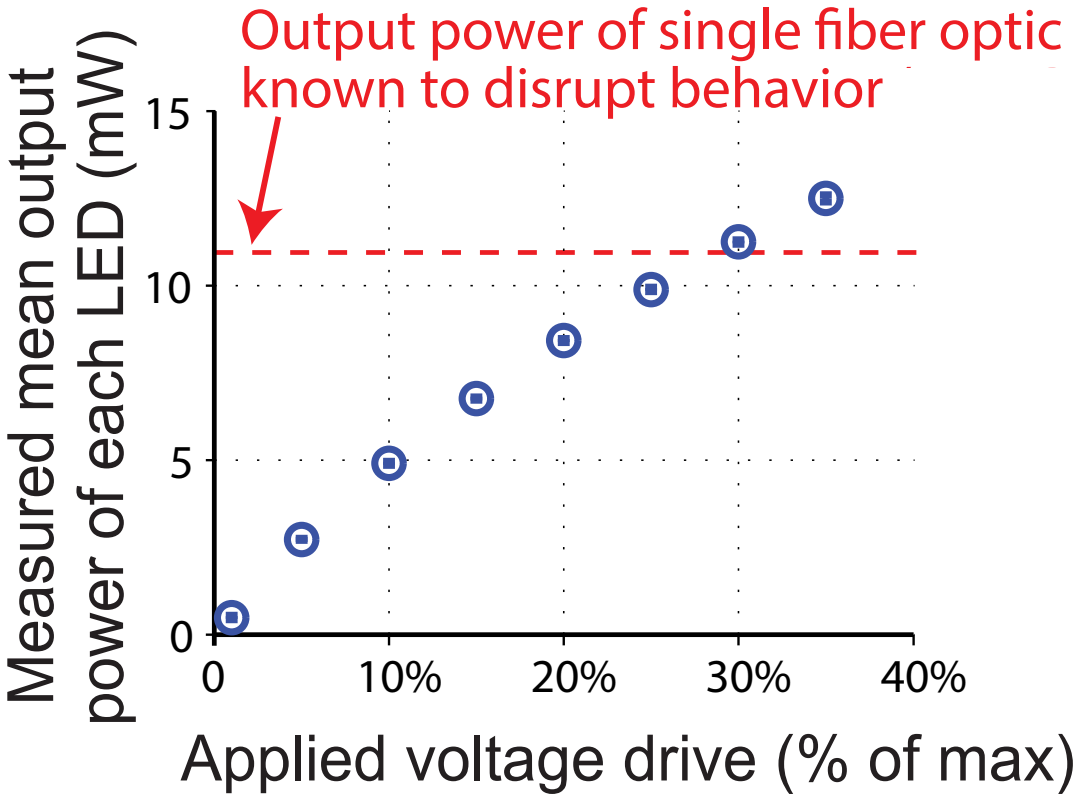
Implant for both IT population recording and V4 optogenetic suppression
- ***Test behavioral and neural disruption at different time lags for MURI images***

Develop, test, and deploy a method for delivering light to large regions of cortex

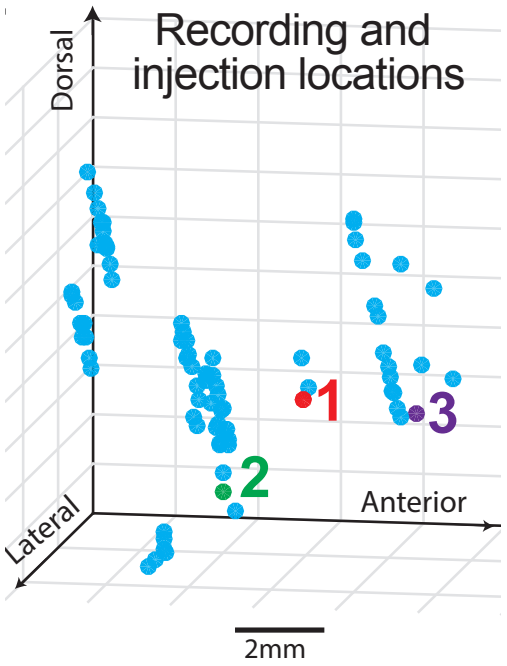


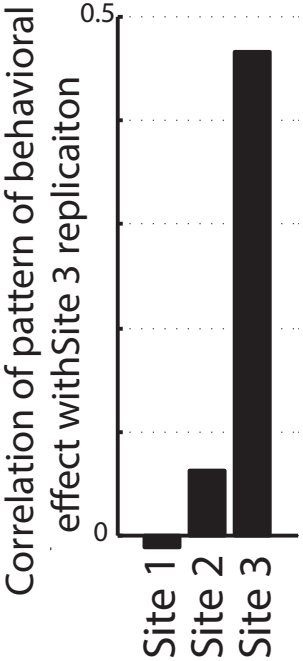
Develop, test, and deploy a method for delivering light to large regions of cortex

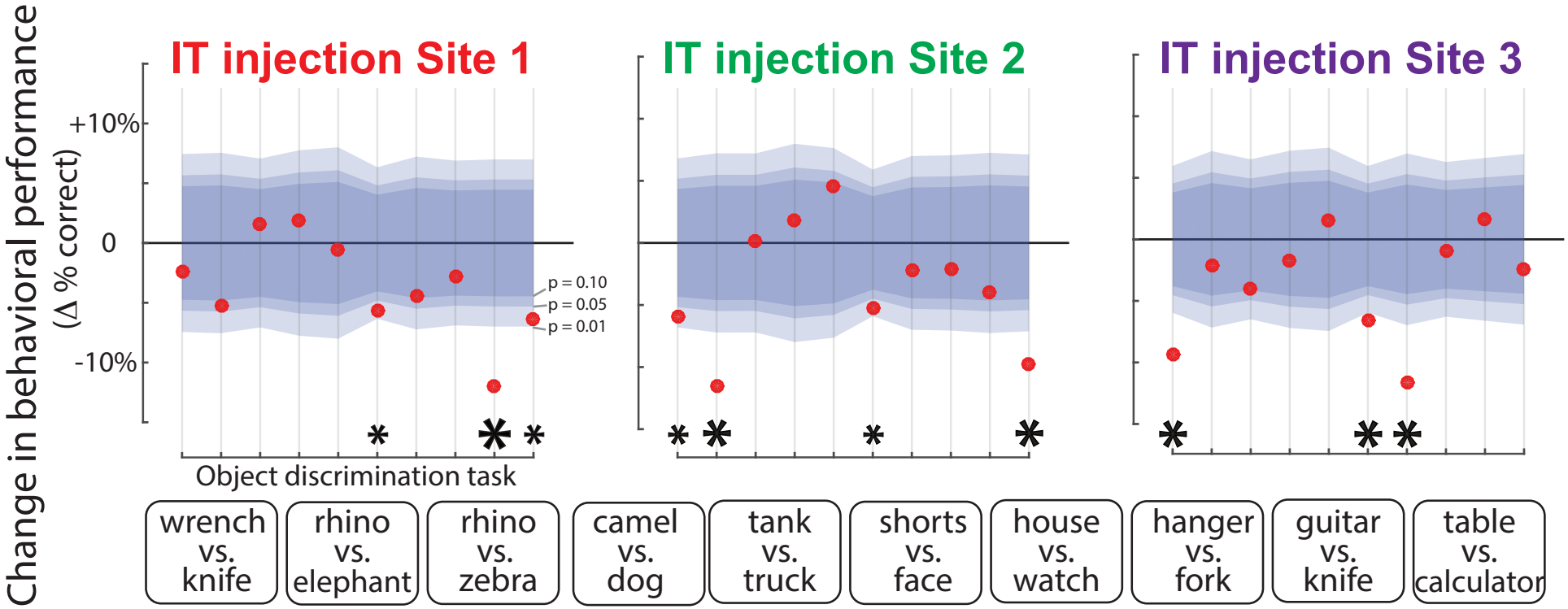




Related IT object code science likely to result from this work…
(take advantage of spatially patterned LED stimulation)







DiCarlo Aim 2 *(MURI 2.4.4) Test causality: When do IT object representations depend on bidirectional computing? (DiCarlo) .*

Technical steps and progress thus far:

Demonstrate that optogenetic control can: ✓
- ***suppress ventral stream neurons with rapid temporal onset*** ✓
- ***produce predictable behavioral disruption*** ✓

Develop, test, and deploy a method for delivering light to large regions of cortex
- ***Chronic LED arrays: Design and construct*** ✓
  - ***Heat and power testing*** ✓
  - ***Test optogenetic control performance in simpler cortical area (V1)***
    - ***Surgical implantation techniques*** ✓
    - ***Demonstrate predictable behavioral suppression***
  - ***Test optogenetic control performance in ventral stream (V4, IT)***
    - ***Confirm neural suppression***
    - ***Confirm behavioral disruption***
    - ***Test behavioral disruption at different time lags*** *(relative to image onset)*
    - ***Test behavioral disruption at different time lags for MURI images***

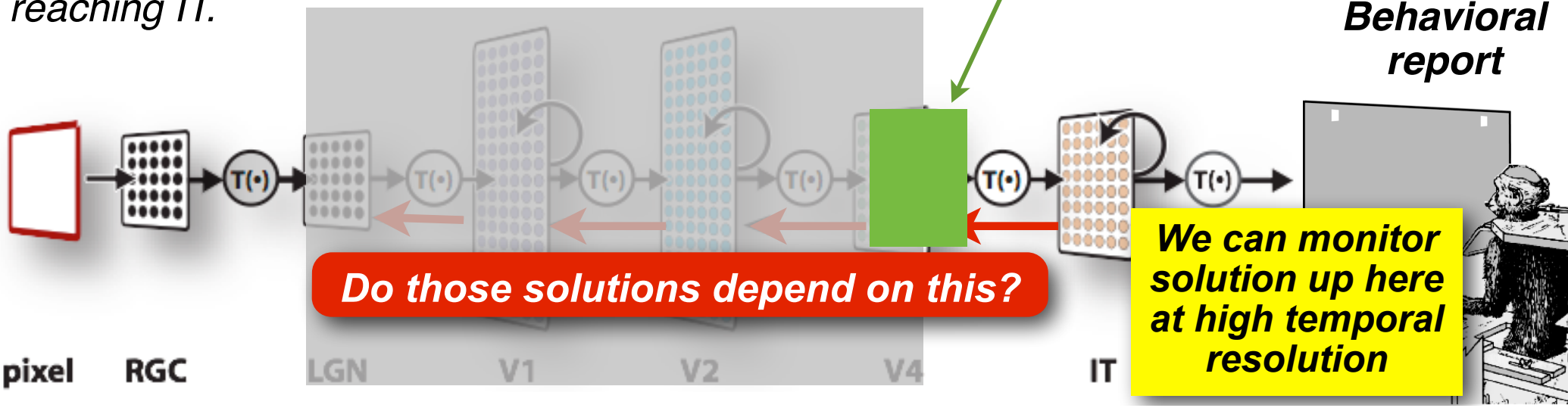
Implant for both IT population recording and V4 optogenetic suppression
- ***Test behavioral and neural disruption at different time lags for MURI images***

(MURI 2.4.4) Test causality: When do IT object representations depend on bidirectional computing? (DiCarlo) .

Approach 1: Temporally-delayed blockade of dominant cortical input to IT cortex

Apply optogentic blockade of central V4 or PIT at time lags after image onset on randomly-interleaved trials.

Logic: information built from cortical-cortical bidirectional computing is prevented from reaching IT.



DiCarlo Aim 2 ***(MURI 2.4.4) Test causality: When do IT object representations depend on bidirectional computing? (DiCarlo) .***

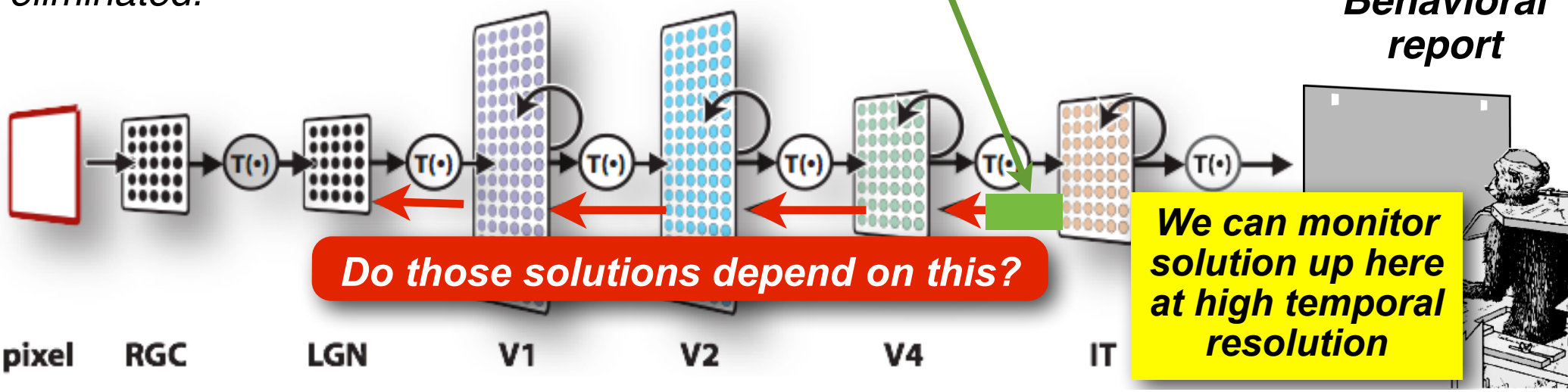
Approach 1: Temporally-delayed blockade of dominant cortical input to IT cortex

Apply optogentic blockade of central V4 or PIT at time lags after image onset on randomly-interleaved trials.

Approach 2: Trial-interleaved blockade of only cortical-cortical feedback projections

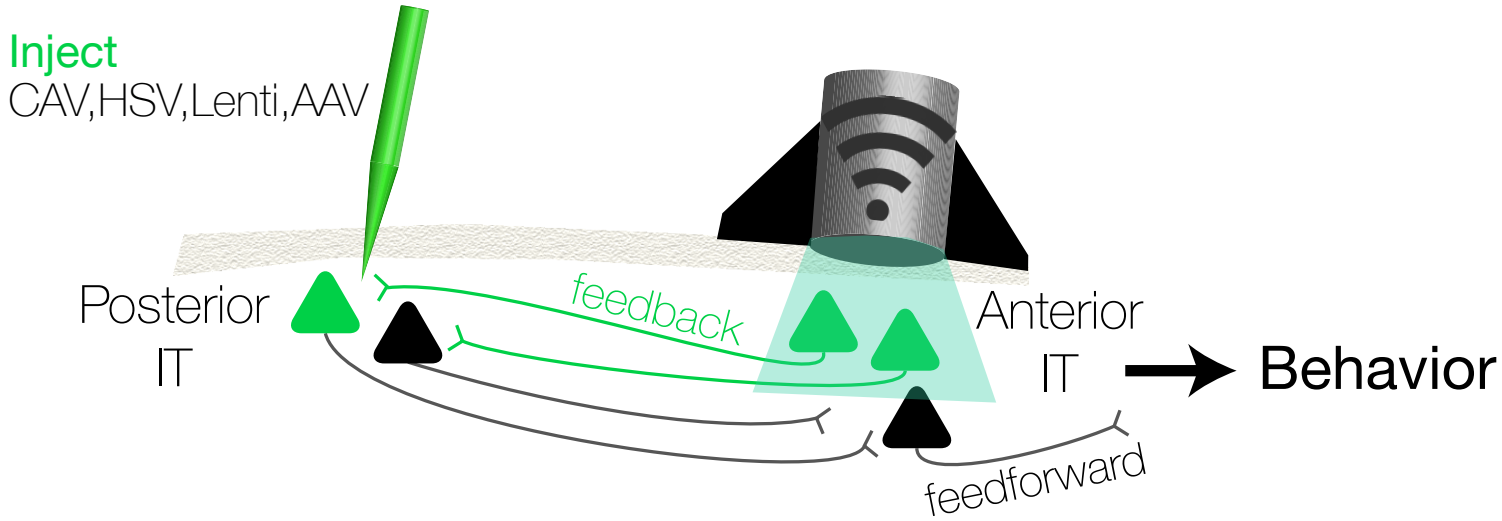
Would be a cleaner manipulation than Approach 1, and would complement our understanding from that work. But no group has yet demonstrated these more advanced optogenetic tools in a monkey.

Logic: information building from cortical-cortical bidirectional computing (over one stage) is eliminated.



DiCarlo
Aim 2

Retrograde Approach - Inhibit cell bodies



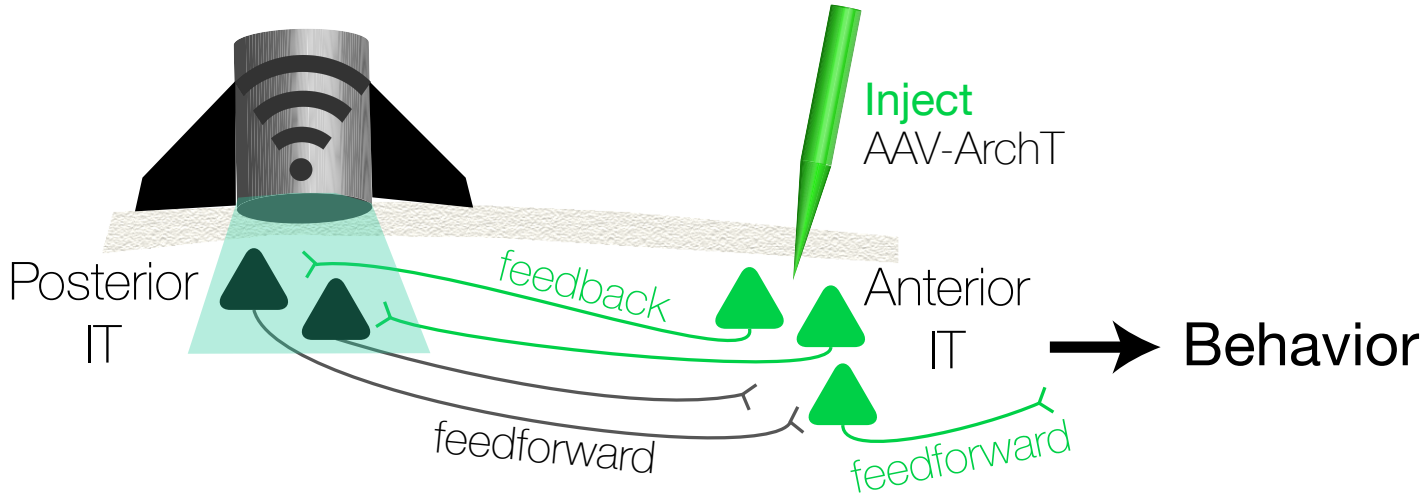
Viruses testing:
CAV-GFP-Cre (retrograde)
HSV-LT-Ef1a-eGFP (retrograde)

Lenti-Fug E-GFP
Lenti-Fug B2-GFP

Lenti-CAG-Chrimson-TdTomato
Lenti-Tloop-Chrimson-TdTomato

DiCarlo
Aim 2

Anterograde Approach - Inhibit axons

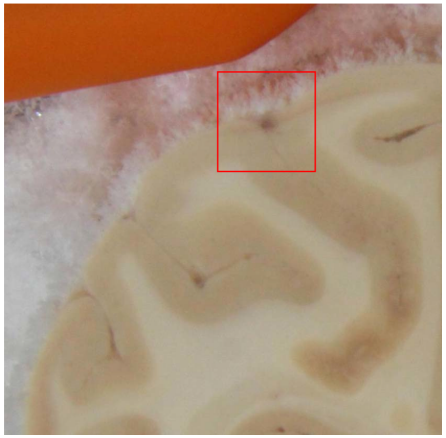


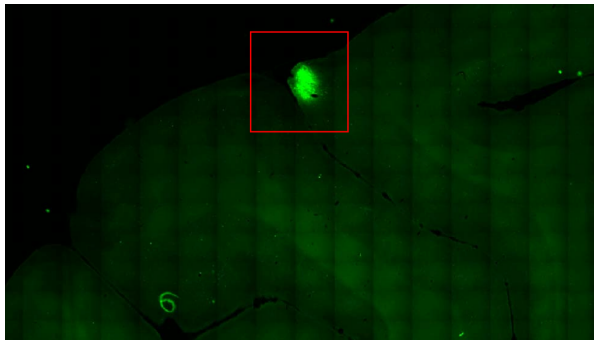
Viruses testing:

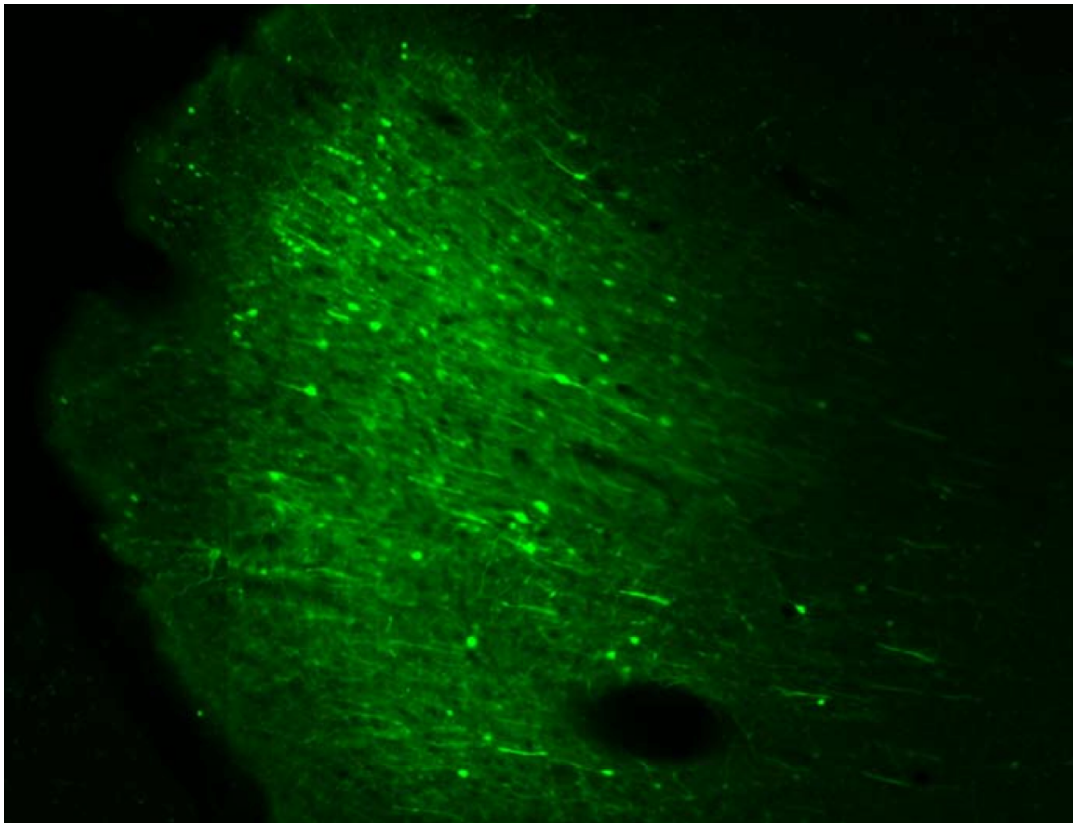
AAV2/1.CB7.CI.mCherry.WPRE.rBG (2/1. mCherry)
AAV2/9.hSyn.hChR2(H134R)-eYFP.WPRE.hGH
AAV2/8 Jaws-GFP-ER2 (anterograde)

DiCarlo
Aim 2

Pilot injection & histology







HSV-LT-Ef1a-eGFP (retrograde)
4 days post-injection into LIP