

## Peer Review Overview

### **Manuscript Title: “Supporting Generalization in Non-Human Primate Behavior by Tapping into Structural Knowledge: Examples from Sensorimotor Mappings, Inference, and Decision-Making”**

Received	17-Aug-2020
1 <sup>st</sup> Decision	13-Nov-2020
Revision Submitted	15-Dec-2020
Accepted	12-Jan-2021

---

### Decision Letter

Dear Jean-Paul and Dora,

Thank you for submitting your manuscript to Progress in Neurobiology. We have received comments from reviewers on your manuscript. Your paper should become acceptable for publication pending suitable minor revision and modification of the article in light of the appended reviewer comments.

When resubmitting your manuscript, please carefully consider all issues mentioned in the reviewers' comments, outline every change made point by point, and provide suitable rebuttals for any comments not addressed.

To submit your revised manuscript go to <https://www.editorialmanager.com/proneu/> and log in as an Author where you will see a menu item called 'Submission Needing Revision'.

Please resubmit your manuscript by Jan 12, 2021.

We look forward to receiving your revised manuscript.

Kind regards,  
Sabine

Sabine Kastner, MD, PhD  
Editor-in-Chief  
Progress in Neurobiology

#### Comments from the Editors and Reviewers:

Reviewer #1: Noel et al. lays out an argument for the development of behavioral tasks that more closely mirror the complex behaviors which the neuroscience community is ultimately trying to understand, the behaviors of the natural environment. Toward that end, they propose a class of tasks that engage closed action-perception loops, simultaneously investigating active sensing, dynamical planning, and how organisms leverage existing knowledge and structural regularities in their environments. Importantly, they provide a concrete example of their proposal in the form of a virtual navigation task that preserves many of the physical laws of the natural environment. They show that non-human primates are able to learn this task relatively easily (compared to traditional behavioral tasks) and furthermore they generalize, even demonstrating zero or one-shot learning in novel sensorimotor environments. Importantly, they also demonstrate that the analysis and interpretation of such tasks is both tractable and informative about learning and sensorimotor processing.

This is a well written article giving an important and timely perspective regarding the need for complex behavioral tasks in systems neuroscience, that provides concrete examples of the advantages, flexibility, and remaining challenges of such tasks.

-----

The only overarching response I have to the perspective article (beyond an enthusiastic agreement with the authors regarding the need for more complex behavioral paradigms) is related to the statements early in the article that promise to use reinforcement learning as a theoretical framework. The authors have presented a series of complex behavioral tasks that are certainly reinforcement learning tasks (tasks where it is clear that learning is driven by the goal of maximizing cumulative reward, otherwise the monkeys wouldn't adjust their gain or choose the closer target), I guess my trouble is that the early framing of the paper left me expecting that the authors would arrive at a more specific (and perhaps more continuous time) analysis framework for the behavior that had implications for relating their results to physiology. But I suppose I may have read a bit too much into the framing. The authors do briefly address the challenge of relating complex behavior over time to neurophysiology in the discussion. And I have to admit that while I might want this aspect of the paper to be more developed, I also recognize that the issues of analyzing complex behavior over time and relating it to neural data is an issue that systems neuroscience is only just beginning to address. In fact, that reality is precisely the reason a perspective article like this is important and timely.

-----

Figure 3B. Could you add an inset that shows that shows a scatter plot of the predicted radial error vs. the actual radial error for monkey J? If I understand the text correctly there is no relationship between the gain fluctuations (3B lower) and the radial error (3B upper, black curve,  $p=.65$ ) but by eye it sort of looks like there's a negative correlation between the predicted radial error vs. the actual radial error for monkey J. If I'm following everything correctly and this negative correlation exists, it doesn't weaken the argument, it just means that the monkeys are actually doing a better job of correcting for the gain when the predicted error is larger. I would guess that the size of that predicted error is directly related to the size of the gain fluctuation, which is why I'm asking questions.

Related to the observations/questions above, might it also be important to look at the cross correlation over trials of actual radial error vs. gain as well as actual radial error vs. predicted radial error.

Minor comments:

The title uses the phrase expressive behavior and while it is mentioned in the abstract and at the end of the paper "expressive behavior" is never explicitly defined. If you choose to keep the same title, I would want you to provide a more explicit definition. I would also recommend reconsidering the title. It seems to no longer fully reflect the scope of the paper, but perhaps it's just that I don't have a clear definition of expressive in my head.

Figure 3B: consider a separate panel label for 3B, lower and upper since they are separate plots like C and D.

There were also some places in the paper that word choice made it somewhat difficult to follow the otherwise clear logic:

Page 5 — "...our peripheral sensors are seldom static...". It's really unclear to me what is meant by this phrase, even after skimming through Ahissar & Arieli, 2001. I think part of the problem is that I really don't know what is meant by peripheral in this situation (except that I'm pretty positive you don't mean the visual periphery).

Page 7 —

- \* ... integration task [including] many of the elements detailed...
- \* protracted => longer, " fulfill these desiderata" => include these elements

Page 13 —

- \* "...but we consider that a focus on natural behaviors..." => "...but we believe that a focus on natural behaviors..."
- \* remove "...the cadre of..."
- \* "This shows our astonishing intelligence..." - just be careful that you actually mean our (as in humans) here. The earlier part of the paragraph is setting up the superior performance of your NHPs compared to humans in other experiments.

Reviewer #2: Review of Progress in Neurobiology Perspective by Noel et al. entitled: Supporting Expressive Behavior in Non-Human Primates by Tapping into Structural Knowledge

In this exciting and fascinating perspective piece, Noel et al. provide a very intriguing and thought provoking take on the need to complement well-established workhorse tasks of neuroscience such as delayed match to sample, go no-go, and 2AFC that provide excellent behavioral "control" to tasks that take advantage of reinforcement learning and structural aspects in the environment. I found the premise interesting and definitely excellent as a perspective. I had a few questions and comments that might make the paper even more appealing to a broad variety of readers.

I think there is a bit of a false equivalency between visual perturbations and sensorimotor adaptation studies that use for instance velocity dependent curl force fields that take the monkeys 10-30 trials to adapt to. Adaptation needed for motor tasks involve much more complicated degrees of freedom in the arm and the elbow and the shoulder compared to a simple 2D visual task with changing gains. I am sure the authors will agree that simple lateral jumps of a cursor or speed up or slowdown of a cursor are far simpler tasks that may be easy for animals to adapt to. If the authors perturbed the joystick in a curl force field manner and the monkey learned to adapt after a single trial, one could construct such an equivalency and that their naturalistic setting led to faster learning. I perhaps worry this is a bit apples and oranges and perhaps a bit too strong and would be perceived by readers as such. I still think the zero shot to one shot adaptation is very impressive and is still worth highlighting but I am unsure that one can make direct comparisons to motor adaptation experiments and this 2D joystick control.

Similarly, I find the discussion on natural decision-making a bit too strong. The classical decision-making tasks including the ones that my lab and the Angelaki lab itself has used to great effect involve deliberation on noisy stimuli that often are hard for the monkeys to learn. There are often timeouts for the monkeys for wrong choices in addition to a lack of reward. The task here is choosing between two targets one near and one far. As the authors themselves somewhat indicate, this is a simple reward optimization problem and there are no penalties for choosing the wrong target given that both targets are equally rewarded. I am confident, if I provided a monkey two targets with equal rewards to reach to the monkey would choose the nearer target. I think one can certainly say that there is a benefit to these tasks in terms of ease of training, but they do not get at many of the questions that perceptual decision-making tasks focus on. Nevertheless, I agree on the point and this can be used for inducing rich choice behavior. Of particular interest would be if there were 3 or 4 targets then it has benefits over and above the 2AFC and allows study of multi-option decision-making.

Perhaps it would be relevant to allude to studies that use naturalistic sensory stimuli to understand the brain in combination with the tight action-perception loop the authors discuss. In the real world, we don't see arrows, we see reality. For the study of speech and auditory function, there really is no other substitute than natural stimuli. And if we want artificial general intelligence it needs to operate in the real world. Would it make sense to incorporate a rich virtual reality like world in 3D for the monkeys to explore (sort of like a video game) or audiovisual avatars like I used in my 2013 paper (Chandrasekaran et al. PNAS 2013) to get monkeys to respond to audiovisual stimuli, or faces as used by the Tsao lab and famously by Charlie Gross to understand the function of brain areas.

I also think perhaps the authors should discuss studies of motor control and brain computer interfaces where monkeys often perform pinball tasks or curved reaches. These types of studies involve free movement of the arm in 2 or 3 dimensions and provide access to rich trajectories for the experimenter like the task here. In those the monkey knows how to use its arm in 3d and might even need less training. For instance, monkeys will reach to a visual target on day 1 on a screen. It is just very natural for them to do so.

I think the authors need to also perhaps consider citing other efforts now to develop freely moving monkey experiments that have been developed by Berger et al. 2018, and OpenMonkeyStudio from Zimmerman and Hayden, and treadmill walking experiments from Foster et al 2014 (freely moving monkey treadmill model). Perhaps alluding to such efforts will also broaden the perspective of the paper.

The reason behind simple tasks is that people worry about influences on neural data from covariates. In fact, this is often a common critique of rodent experiments where the body position of the mouse/rat is highly variable from trial to trial and often contributes considerable variance to neural activity in a given region. Multiple papers have been written on this topic, but it is unclear what the solution is. I am sure if the authors were to read a paper where eye movements were unconstrained then they would doubt the neural results. Also, my suspicion is that many of these controlled behavioral tasks in monkeys (where for instance monkeys are forced to reach to a target without moving their eyes) emerged after



experimenters found that lots of covariates impact neural activity. Then disentangling different neural effects is complex. Fortunately, perhaps my next point might be of help.

I think the authors have cited dPCA and TCA as candidate methods for analyzing neural correlates of naturalistic behavior. However, I feel the authors should also consider citing methods such as LFADS (pandarinath et al. 2018), GPFA (Yu et al. 2009), and PSID (Shanechi lab, Nature Neuroscience 2020). These methods are far more suited to the formidable challenge of extracting sensible neural insights from these behaviors. dPCA demands averaging based on parameters modulated by the experimenter! Single trial analysis with rigorous measurement of covariates might allow us to extract neural insights from complex behavior especially when the behavior involves free eye and arm movements.

Chandramouli Chandrasekaran

## Author Response Letter

### Reviewer #1:

Noel et al. lays out an argument for the development of behavioral tasks that more closely mirror the complex behaviors which the neuroscience community is ultimately trying to understand, the behaviors of the natural environment. Toward that end, they propose a class of tasks that engage closed action-perception loops, simultaneously investigating active sensing, dynamical planning, and how organisms leverage existing knowledge and structural regularities in their environments. Importantly, they provide a concrete example of their proposal in the form of a virtual navigation task that preserves many of the physical laws of the natural environment. They show that non-human primates are able to learn this task relatively easily (compared to traditional behavioral tasks) and furthermore they generalize, even demonstrating zero or one-shot learning in novel sensorimotor environments. Importantly, they also demonstrate that the analysis and interpretation of such tasks is both tractable and informative about learning and sensorimotor processing.

This is a well written article giving an important and timely perspective regarding the need for complex behavioral tasks in systems neuroscience, that provides concrete examples of the advantages, flexibility, and remaining challenges of such tasks.

We thank the reviewer for his/her overall positive and constructive feedback.

The only overarching response I have to the perspective article (beyond an enthusiastic agreement with the authors regarding the need for more complex behavioral paradigms) is related to the statements early in the article that promise to use reinforcement learning as a theoretical framework. The authors have presented a series of complex behavioral tasks that are certainly reinforcement learning tasks (tasks where it is clear that learning is driven by the goal of maximizing cumulative reward, otherwise the monkeys wouldn't adjust their gain or choose the closer target). I guess my trouble is that the early framing of the paper left me expecting that the authors would arrive at a more specific (and perhaps more continuous time) analysis framework for the behavior that had implications for relating their results to physiology. But I suppose I may have read a bit too much into the framing. The authors do briefly address the challenge of relating complex behavior over time to neurophysiology in the discussion. And I have to admit that while I might want this aspect of the paper to be more developed, I also recognize that the issues of analyzing complex behavior over time and relating it to neural data is an issue that systems neuroscience is only just beginning to address. In fact, that reality is precisely the reason a perspective article like this is important and timely.

We thank the reviewer for pointing out that the framing of the piece in the first few sections – heavily relying on reinforcement learning – may mislead readers and cause false expectations. In fact, our intention was always to distill *conceptual* teaching from that field, and leverage



those in experimental design, and not to further reinforcement learning itself (a field we are not experts in). To address this important comment we have taken the following 3 measures.

First, we have added a paragraph at the end of section 1. *Introduction: The Neurosciences of Tomorrow*, in order to explicitly state the goal and roadmap of the current piece. The section reads as follows:

Similarly, authors (Botvinick et al., 2020) have eloquently written about a new wave of artificial intelligence (e.g., deep reinforcement learning) that is poised to inform next-generation neuroscience. Here, it is not our intention to discuss reinforcement learning and control broadly, nor to illustrate how these frameworks may guide data analyses (see Inverse Reinforcement Learning, Ng & Russell, 2000, Choi & Kim, 2011; Inverse Rational Control; Daptardar et al., 2019; Wu et al., 2020; Kwon et al., 2020). Instead, we very specifically distill a number of the conceptual contributions from this framework, and attempt to translate these to experimental choices that may accelerate the study of brain function by allowing animals to express intelligent behavior akin to that of human everyday life. Most importantly, we provide a concrete example of an experimental ecosystem that allows for such generalization, demonstrating generalization to three distinct and fundamental computations.

And amended the *Abstract* accordingly:

The framework of reinforcement learning naturally wades across action and perception, and thus is poised to inform the neurosciences of tomorrow, not only from a data analyses and modeling framework, but also in guiding experimental design.

Second, we have considerably reduced sections 1. 1. *How Should Agents Interact with Their Surrounding? Active Sensing and Planning* and 1.2. *What Type of Task? Cognitive Maps, Structural Knowledge, and Learning Sets* in order to further emphasize the example experimental ecosystem and the three generalization examples, as opposed to reinforcement learning. The two sections abovementioned have been reduced from ~1500 words to ~1100 words. This section has also been broadened to include reference to the field of control theory, another likely important theoretical contributor to the neurosciences of tomorrow.

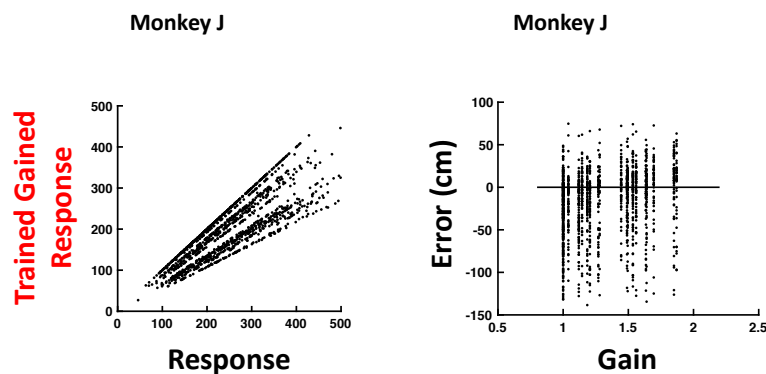
Lastly, we add to the discussion mention to the fact that continuing the development of continuous time-analyses (both model-free and model-based) is critical in scaffolding the study of natural behavior, and vice-versa. This paragraph reads as follows:

Now, it is true that the study of natural behaviors comes at the expense of needing more sophisticated behavioral tracking (Bala et al., 2020; Pereira et al., 2020; Wu et al., 2020) and data analysis tools (see Huk et al., 2018, for an insightful perspective on this topic). For example, one of the pillars of data analyses in neurophysiology, i.e., averaging across trials, breaks down in naturalistic tasks with continuous action-perception loops. On the bright side, it is unlikely that the brain computes averages. Further, powerful techniques for the efficient estimation of single-units tuning functions are already underway (Balzani et al., 2020; Dowling et al., 2020), and a number of techniques for inferring the latent dynamics of populations of neurons exists, even at the single trial level (e.g., GPFA; Yu et al., 2009; LFADS: Sussillo et al., 2016; Pandarinath et al., 2018; PSID: Sani et al., 2020) and when requiring time-warping (Williams et al., 2020). Further, while in the current piece we have not discussed nor leveraged a reinforcement learning perspective for model-based data analyses (see Choi & Kim, 2011; Daptardar et al., 2019; Kwon et al., 2020; Wu et al., 2020), it is our hope that already starting to study natural and generalizable behaviors in systems neuroscience will precisely demand for developments in this area.

Figure 3B. Could you add an inset that shows that shows a scatter plot of the predicted radial error vs. the actual radial error for monkey J? If I understand the text correctly there is no relationship between the gain fluctuations (3B lower) and the radial error (3B upper, black curve,  $p=.65$ ) but by eye it sort of looks like there's a negative correlation between the predicted radial error vs. the actual radial error for monkey J. If I'm following everything correctly and this negative correlation exists, it doesn't weaken the argument, it just means that the monkeys are actually doing a better job of correcting for the gain when the predicted error is larger. I would guess that the size of that predicted error is directly related to the size of the gain fluctuation, which is why I'm asking questions.

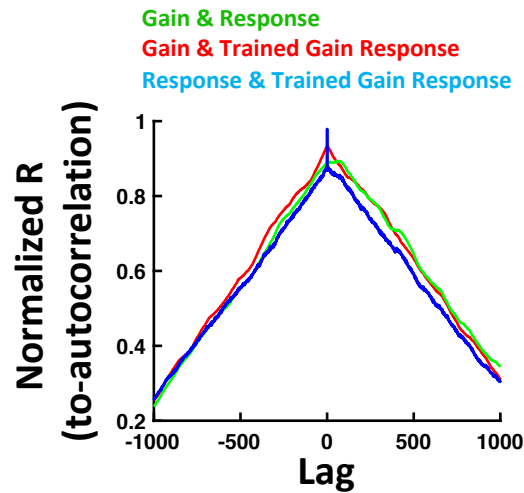
The reviewer is correct, as indicated in the manuscript, we saw no relation between the relatively rapid (~50 trials) changes in gain, and the slower fluctuations in performance. Further, we performed the analyses requested by the reviewer here. The apparent negative correlation between predicted and actual error the reviewer observed in Figure 3 is due to the fact that at small gains (close to 1), there won't be much of a difference between predicted and actual end locations, and thus error. Conversely, at larger gains, the difference will be exacerbated. This becomes clear when plotting a scatter plot of predicted vs. actual endpoints, as it can be observed in the left-most panel below, this simply reflects the varying gains.

When we plot error as a function of gain (right panel), we see no correlation between the gain applied and error. We decided not to include these figures in the main text, as their explanation is straightforward and the key lack of a correlation between error and gain changes is already included in the main text.



Related to the observations/questions above, might it also be important to look at the cross correlation over trials of actual radial error vs. gain as well as actual radial error vs. predicted radial error.

We performed a cross-correlation between gain and response (green), gain and predicted responses (trained gain, red), and between the actual and predicted response (blue). These are all normalized such that the auto-correlation at lag = 0 is 1. The notable feature of these cross-correlations is that it jumps and takes a value close to 1 ( $=0.96$ ) at zero lag for the correlation between actual and predicted response. The cross correlation is not perfect ( $=1$ ) as on a few occasions the monkey overshoots the target, such that real and predicted end-locations are on opposite sides of the target (one is overshoot and the other is undershot).



Minor comments:

The title uses the phrase expressive behavior and while it is mentioned in the abstract and at the end of the paper "expressive behavior" is never explicitly defined. If you choose to keep the same title, I would want you to provide a more explicit definition. I would also recommend reconsidering the title. It seems to no longer fully reflect the scope of the paper, but perhaps it's just that I don't have a clear definition of expressive in my head.

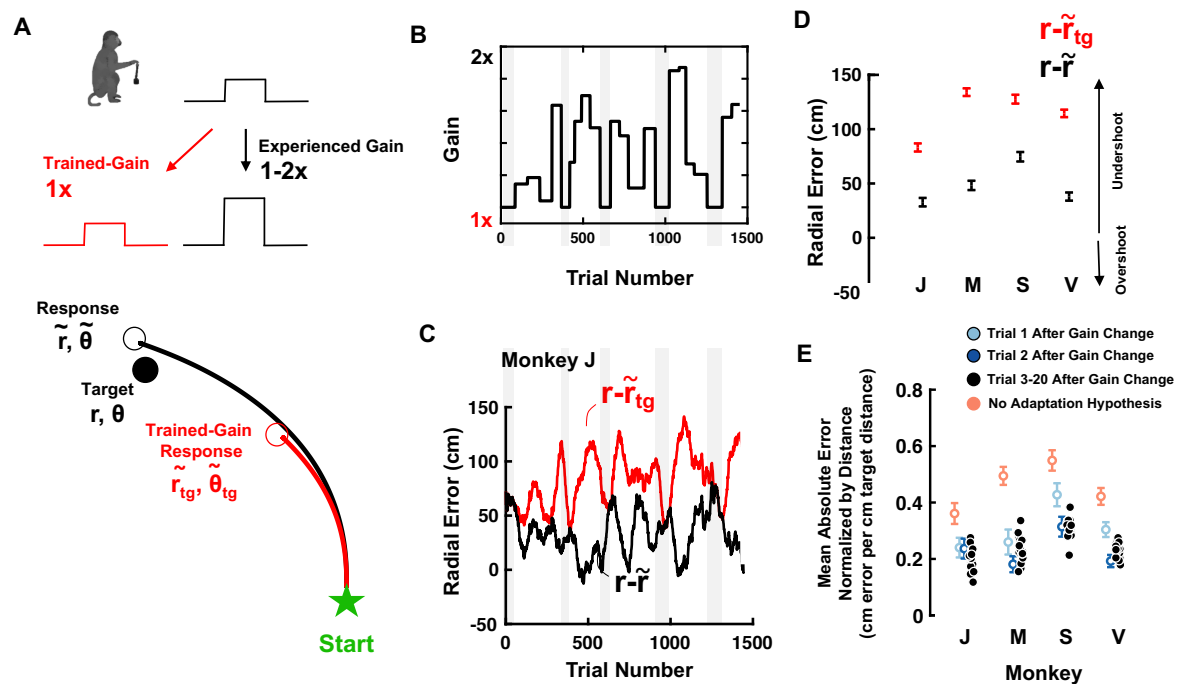
We thank the author for this suggestion. We have indeed changed the title and eliminated all reference to "expressive" behavior. The title now reads:

**Supporting Generalization in Non-Human Primate Behavior by Tapping into Structural Knowledge: Examples from Sensorimotor Mappings, Inference, and Decision-Making**

Figure 3B: consider a separate panel label for 3B, lower and upper since they are separate plots like C and D.

Figure 3 has been updated to separate Figure 3B into Figure 3B and C:





There were also some places in the paper that word choice made it somewhat difficult to follow the otherwise clear logic:

Page 5 — "...our peripheral sensors are seldom static...". It's really unclear to me what is meant by this phrase, even after skimming through Ahissar & Arieli, 2001. I think part of the problem is that I really don't know what is meant by peripheral in this situation (except that I'm pretty positive you don't mean the visual periphery).

This section has been amended to read:

This compartmentalization may seem a sensible first order approach to describe the distinct components forming behavior, and this intent likely drove the development of what are today classic paradigms in the study of brain function. However, in natural conditions, and thus throughout evolution, perception and action do not occur serially. In fact, eye movements not only dictate the content and relative resolution of visual input, but perhaps most importantly, they also dictate the relative timing of sensory input. Critically, this natural "rhythm" (Leopold & Park, 2020) of the visual system seemingly guides its functioning.

Page 7 —

\* ... integration task [including] many of the elements detailed...

\* protracted => longer, " fulfill these desiderata" => include these elements



These have been corrected.

Page 13 —

\* "...but we consider that a focus on natural behaviors..." => "...but we believe that a focus on natural behaviors..."

Corrected.

\* remove "...the cadre of..."

Corrected.

\* "This shows our astonishing intelligence..." - just be careful that you actually mean our (as in humans) here. The earlier part of the paragraph is setting up the superior performance of your NHPs compared to humans in other experiments.

We thank the reviewer for noting this typo. The sentence has been modified as follows:

Remarkably, the animals were able to continuously estimate the likely position of the unseen firefly and appropriately navigate to its reward boundary already within the very first session they encountered moving fireflies. **This shows both the astonishing intelligence of these experimental animals**, being able to traverse a large space of potential actions and states **while** concurrently estimating how other agents are moving through their own space of potential actions and states, **but also emphasizes our duty as experimentalist to allow them to show this behavior**.

**Reviewer #2:**

Review of Progress in Neurobiology Perspective by Noel et al. entitled: Supporting Expressive Behavior in Non-Human Primates by Tapping into Structural Knowledge

In this exciting and fascinating perspective piece, Noel et al. provide a very intriguing and thought provoking take on the need to complement well-established workhorse tasks of neuroscience such as delayed match to sample, go no-go, and 2AFC that provide excellent behavioral "control" to tasks that take advantage of reinforcement learning and structural aspects in the environment. I found the premise interesting and definitely excellent as a perspective. I had a few questions and comments that might make the paper even more appealing to a broad variety of readers.

We thank Dr. Chandrasekaran for his constructive comments and aiding us in broadening the perspective.

I think there is a bit of a false equivalency between visual perturbations and sensorimotor adaptation studies that use for instance velocity dependent curl force fields that take the monkeys 10-30 trials to adapt to. Adaptation needed for motor tasks involve much more complicated degrees of freedom in the arm and the elbow and the shoulder compared to a simple 2D visual task with changing gains. I am sure the authors will agree that simple lateral jumps of a cursor or speed up or slowdown of a cursor are far simpler tasks that may be easy for animals to adapt to. If the authors perturbed the joystick in a curl force field manner and the monkey learned to adapt after a single trial, one could construct such an equivalency and that their naturalistic setting led to faster learning. I perhaps worry this is a bit apples and oranges and perhaps a bit too strong and would be perceived by readers as such. I still think the zero shot to one shot adaptation is very impressive and is still worth highlighting but I am unsure that one can make direct comparisons to motor adaptation experiments and this 2D joystick control.

The reviewer is correct in that comparing changes in sensorimotor gains and visual perturbations is perhaps a false equivalence. As such, we have eliminated the section of the discussion that compared the speed of adaptation during the gain manipulation from that of force-field tasks in humans.

As an aside, we consider this comment very interesting and perhaps something to scrutinize in more detail in the future provided we record the pertinent variables (e.g., joint angles). As Dr. Chandrasekaran notes, the manipulation here was a 1-dimensional change (gain) applied in 2-dimensions (forward and lateral velocity). Yet, to navigate in this space, the animals are using a joystick controlled by their hands, arms, etc. Thus, while the visual perturbation is 2-dimensional, correction for this perturbation may be higher-dimensional. This is not in the purview of the current perspective, but we appreciate the food-for-thought.

Similarly, I find the discussion on natural decision-making a bit too strong. The classical decision-making tasks including the ones that my lab and the Angelaki lab itself has used to great effect involve deliberation on noisy stimuli that often are hard for the monkeys to learn. There are often timeouts for the monkeys for wrong choices in addition to a lack of reward. The task here is choosing between two targets one near and one far. As the authors themselves somewhat indicate, this is a simple reward optimization problem and there are no penalties for choosing the wrong target given that both targets are equally rewarded. I am confident, if I provided a monkey two targets with equal rewards to reach to the monkey would choose the nearer target. I think one can certainly say that there is a benefit to these tasks in



terms of ease of training, but they do not get at many of the questions that perceptual decision-making tasks focus on. Nevertheless, I agree on the point and this can be used for inducing rich choice behavior. Of particular interest would be if there were 3 or 4 targets then it has benefits over and above the 2AFC and allows study of multi-option decision-making.

We agree with the reviewer that the decision-making task we employed here was very simple. The main purpose was to show that researchers can still employ 2AFC in a naturalistic setting (though admittedly we did not examine the full gamut of what is typical in most current day studies of decision making; uncertainty, cost-functions, etc.). We have modified the text in the following manners.

### Section 3. 3. *Natural Decisions – Choice of Target*

This sort of decision-making is typically studied in laboratories by training animals to perform a predefined motor response (e.g. a saccade) to report an arbitrary stimulus category (e.g. a grating tilted clockwise relative to vertical; e.g., Glimcher, 2001; Romo & Salinas, 2003; Shadlen & Kiani, 2013). **These standard** tasks are purposely kept as simple as possible in order to isolate decision-making signals uncorrupted by signals such as memory, attention, and other. **Further, these protocols typically involve additional manipulations, such as noise-levels and the magnitude of rewards and punishments, in order to isolate different components of the decision process, i.e., uncertainty, sensitivity, cost and reward functions, decision criteria, and confidence (see e.g., Chandrasekaran et al., 2017).**

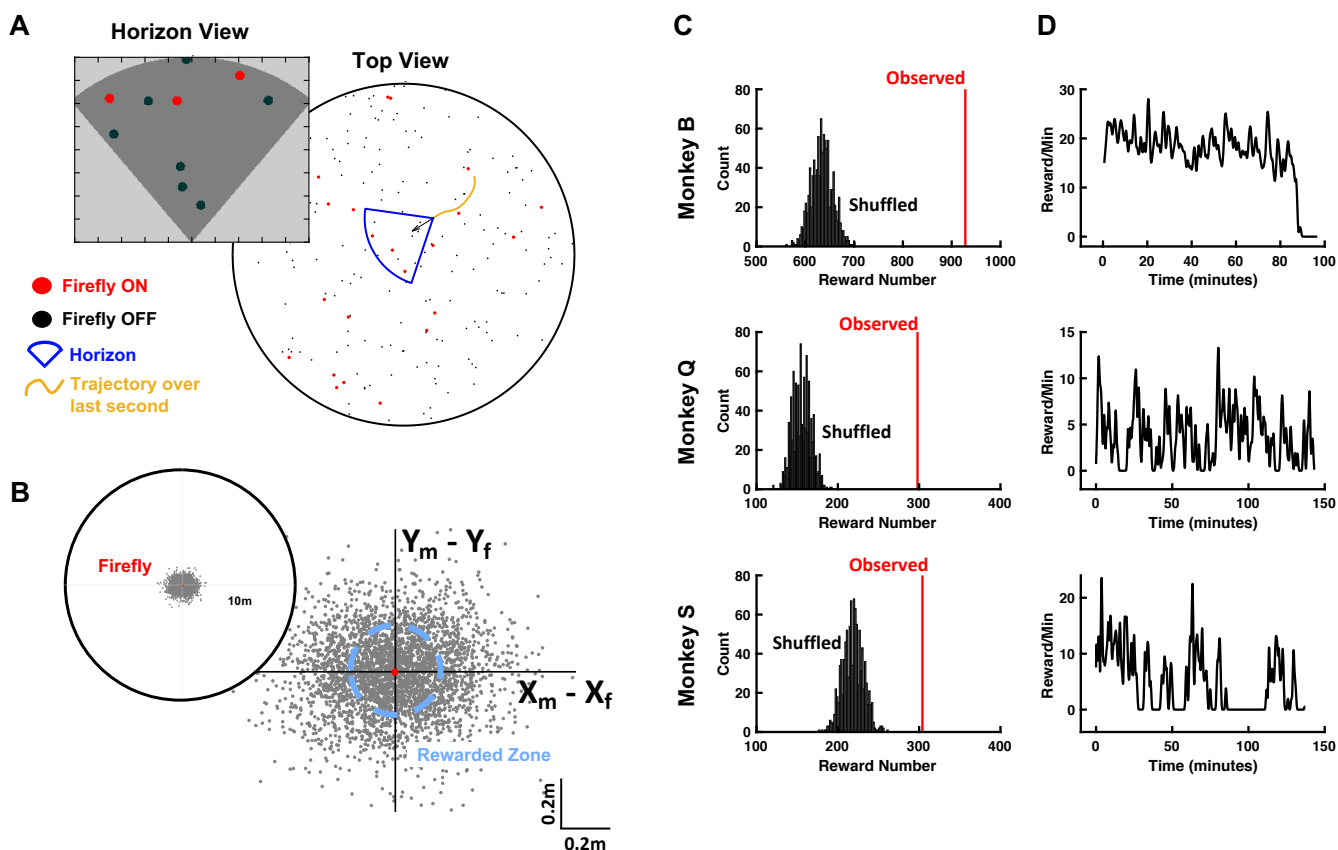
#### *Outlook*

Lastly, we were able to show that **within a binary decision-making task** monkeys instinctively maximize their expected reward per unit time, by choosing to navigate to targets associated with a higher reward rate. **Further, decision-making naturally generalized to the case of hundreds of fireflies in a world with no trial structure. The presented examples of decision making did not include all aspects of state-of-the-art decision making studies, where observers are typically asked to deliberate on noisy and often ambiguous sensory stimuli. However, within the framework of the firefly task we could easily manipulate the contrast of fireflies, as well as the reliability and congruency of optic flow vis-à-vis self-motion. The general strategy of training animals on a “reporting mechanism” – stopping at the firefly location – can be used to study a large set of naturalistic behaviors given that animals naturally attempt to maximize their reward rate.**

Further, we agree with the reviewer that expanding the decision-making example to a multi-option task would very nicely show generalization. Thus, we had 3 monkeys navigate a virtual environment inhabited by 200 fireflies that would flash at random times. We have modified the text to show that animals (i) deliberately stop at the location of fireflies even within this very complex world, and (ii) show interesting idiosyncrasies (e.g., when they decided to engage or not in the task) that could be exploited in the future. The added text and figure are the following:

**After demonstrating the ability to study the relatively simple case of 2 alternative choices under the firefly task, we questioned whether macaques would generalize – still within the very first session – to a much more complex scenario, one with multiple options and where the number of targets is constantly changing (ranging from 0 to 7 concurrently visible fireflies, with many more that could be in the vicinity but invisible, Fig. 6A inset, shows 10 in total with 3 visible). We embedded these animals within a large virtual environment (10 meters in diameter) that would repeat to infinite (see *Supplementary Videos S1 & S2*). This space was inhabited by 200 fireflies flashing at random times (Fig. 6A, red = firefly on, black = firefly off, blue = horizon, what is currently visible from an egocentric perspective). When we visualize the stopping**

locations of an example subject, we can observe that 43% of the times an animal stopped (linear velocity < 5cm/s), it did so within the boundaries of a reward zone (Fig. 6B). To quantitatively assess whether animals were deliberately stopping at the location of fireflies, we computed their total number of rewards (Fig. 6C, Observed), as well as the total number of rewards they would have received under different (rotated, 1000 permutations) configurations of firefly locations (Fig. 6C, Shuffled). All monkeys received significantly more rewards (all  $p < 10^{-3}$ ) than they would have if simply stopping randomly. Further, we examined their reward rate per minute (Fig. 6D), and while we observed interesting idiosyncratic fluctuations (note Monkey S decided not to forage for ~30min), animals did not consistently improve with time, suggesting they were perfectly capable of navigating this large space of potential targets from the get go. Altering between voluntarily exploiting our environment vs. resting (Fig. 6D, bottom panel) is commonplace in our daily lives, yet seldomly falls within the purview of our traditional experiments (but see Milton et al., 2020 for a recent exception).



**Figure 6: Naturalistic Foraging within a Multi-Firefly Scenario.** (A) Two-hundred fireflies were present within a large virtual environment, and flashed at random times (red = firefly on, black = firefly off, blue and inset show the horizon of what was visible during the example frame, yellow trajectory shows movement over the last second). (B) Monkey stopping location referenced to the nearest firefly. (C) Observed (red) and null distribution (black, shuffled) of total rewards within the session. (D) Rewards per minute for the three different monkeys (B, Q, S) within the first session they were exposed to the environment with hundreds of fireflies.



Perhaps it would be relevant to allude to studies that use naturalistic sensory stimuli to understand the brain in combination with the tight action-perception loop the authors discuss. In the real world, we don't see arrows, we see reality. For the study of speech and auditory function, there really is no other substitute than natural stimuli. And if we want artificial general intelligence it needs to operate in the real world. Would it make sense to incorporate a rich virtual reality like world in 3D for the monkeys to explore (sort of like a video game) or audiovisual avatars like I used in my 2013 paper (Chandrasekaran et al. PNAS 2013) to get monkeys to respond to audiovisual stimuli, or faces as used by the Tsao lab and famously by Charlie Gross to understand the function of brain areas.

Indeed, as others have eloquently written about previously, the use of naturalistic stimuli will likely also define our field moving forward. In our particular case, the task is naturalistic, but the stimuli are not. This was a conscious decision given that we still wanted to be rooted in classic neurophysiology (e.g., random dot kinematograms and the study of MSTd). Further, as alluded to by the reviewer, the use of naturalistic stimuli is most common currently in the fields of language comprehension and audio-visual binding, fields somewhat disparate from navigation via path integration (employed here). As such, we briefly mention the utility of naturalistic stimuli, but refer interested readers to appropriate venues. The text has been modified as follows:

In addition to the strong advantages afforded by this approach in the study of behavior – note that all data presented here, 4 animals in 3 different tasks, was collected within 1 week – we also believe naturalistic tasks (and stimuli; e.g., Chandrasekaran et al., 2013; see Sonkusare et al., 2019, Matusz et al., 2018) may ultimately facilitate the interpretation of neural responses.

I also think perhaps the authors should discuss studies of motor control and brain computer interfaces where monkeys often perform pinball tasks or curved reaches. These types of studies involve free movement of the arm in 2 or 3 dimensions and provide access to rich trajectories for the experimenter like the task here. In those the monkey knows how to use its arm in 3d and might even need less training. For instance, monkeys will reach to a visual target on day 1 on a screen. It is just very natural for them to do so.

Indeed, motor control, reaching behavior, and brain computer interfaces share a strong similarity with the task employed here. We agree that referring to this literature can broaden the scope of the current perspective. We have modified the text as follows:

In the Firefly Task, individual trials last on the order of 2 to 4 seconds, and the output of each trial is a two-dimensional data-rich trajectory allowing for robust model fitting (Lakshminarasimhan et al., 2018; Noel et al., 2020) and the prolonged tracking of eye-movements. In this regard, the resulting data from the Firefly Task is akin to that obtained during other naturalistic tasks, such as two (e.g., Ames et al., 2019) or three (e.g., Young et al., 2019) dimensional reaches. It provides time-varying behavioral output on a time-scale that in principle should allow for hundreds or thousands of spikes, and hence accurate decoding from neural signals (as exemplified by the ever more robust decoders built in and for brain-machine interfaces; e.g., Chaudhary et al., 2016).

I think the authors need to also perhaps consider citing other efforts now to develop freely moving monkey experiments that have been developed by Berger et al. 2018, and OpenMonkeyStudio from Zimmerman and Hayden, and treadmill walking experiments from Foster et al 2014 (freely moving monkey treadmill model). Perhaps alluding to such efforts will also broaden the perspective of the paper.



We modified the text to reference these important developments. The very first paragraph of the perspective now reads:

It is undoubtedly an exciting time in systems neuroscience. Techniques for neural recording and perturbation are improving at a remarkable pace (Sejnowski et al, 2014; Jun et al., 2017), and the development of rigorous behavioral training procedures is permitting for psychophysics in smaller animals akin to those classically undertaken in primates (Burgess et al., 2017; IBL et al., 2020). Further, novel approaches for tracking freely moving animals (Foster et al., 2014; Berger et al., 2018, 2020; Michael et al., 2020; Bala et al., 2020; Pereira et al., 2020; Wu et al, 2020; Mao et al., 2020) is encouraging researchers to measure and leverage in their analyses additional degrees of freedom (e.g., eye, arm, and full-body movements), as opposed to artificially restricting these potential sources of explainable variance (Musall et al., 2019). Continued steps in this direction promise to further (re)shape our field, one that is and will be largely defined by novel technologies allowing for the dissection of circuit-based correlates of complex behavior.

Further, we have modified the *Outlook* as follows:

Now, it is true that the study of natural behaviors comes at the expense of needing more sophisticated behavioral tracking (Bala et al., 2020; Pereira et al., 2020; Wu et al, 2020) and data analysis tools (see Huk et al., 2018, for an insightful perspective on this topic). For example, one of the pillars of data analyses in neurophysiology, i.e., averaging across trials, breaks down in naturalistic tasks with continuous action-perception loops.

The reason behind simple tasks is that people worry about influences on neural data from covariates. In fact, this is often a common critique of rodent experiments where the body position of the mouse/rat is highly variable from trial to trial and often contributes considerable variance to neural activity in a given region. Multiple papers have been written on this topic, but it is unclear what the solution is. I am sure if the authors were to read a paper where eye movements were unconstrained then they would doubt the neural results. Also, my suspicion is that many of these controlled behavioral tasks in monkeys (where for instance monkeys are forced to reach to a target without moving their eyes) emerged after experimenters found that lots of covariates impact neural activity. Then disentangling different neural effects is complex. Fortunately, perhaps my next point might be of help.

We agree with the reviewer, the standard protocols (e.g., 2AFC tasks) were likely the best first attempt, and will likely continuously have their role to play in the study of brain function. We consider that there are indeed two choices regarding eye-movement, for example, and the interpretability of neural responses. Either we fix them (as Dr. Chandrasekaran mentions), or we measure them and leverage them in our analyses (as in the examples below). We have modified the text in the following manners.

Introduction:

This rigid experimental control arguably allows for carefully disentangling potential confounding variables (or at least the subset anticipated by experimentalists), yet tends to dissociate perception from action, guides the state-space of potential actions (i.e., actions as “reports” and not as information-sampling; Gottlieb & Oudeyer, 2018), and fixes the utility of samples we draw from the environment.

Section 2. *An Example Ecosystem for Natural Yet Controlled Behavior: Catching Fireflies*



These findings suggest that during naturalistic navigation, eye movements are an integral component of the closed-loop strategy of prediction and action, **and is not to be regarded as a nuisance variable, but instead measured and leveraged.**

I think the authors have cited dPCA and TCA as candidate methods for analyzing neural correlates of naturalistic behavior. However, I feel the authors should also consider citing methods such as LFADS (pandarinath et al. 2018), GPFA (Yu et al. 2009), and PSID (Shanechi lab, Nature Neuroscience 2020). These methods are far more suited to the formidable challenge of extracting sensible neural insights from these behaviors. dPCA demands averaging based on parameters modulated by the experimenter! Single trial analysis with rigorous measurement of covariates might allow us to extract neural insights from complex behavior especially when the behavior involves free eye and arm movements.

We thank the reviewer for highlighting these other methods. Indeed, this part of the manuscript was under-developed in the previous iteration. We have amended the text in the following manner:

**Now, it is true that the study of natural behaviors comes at the expense of needing more sophisticated behavioral tracking (Bala et al., 2020; Pereira et al., 2020; Wu et al, 2020) and data analysis tools (see Huk et al., 2018, for an insightful perspective on this topic). For example, one of the pillars of data analyses in neurophysiology, i.e., averaging across trials, breaks down in naturalistic tasks with continuous action-perception loops. On the bright side, it is unlikely that the brain computes averages. Further, powerful techniques for the efficient estimation of single-units tuning functions are already underway (Balzani et al., 2020; Dowling et al., 2020), and a number of techniques for inferring the latent dynamics of populations of neurons exists, even at the single trial level (e.g., GPFA; Yu et al., 2009; LFADS: Sussilo et al., 2016; Pandarinath et al., 2018; PSID: Sani et al., 2020) and when requiring time-warping (Williams et al., 2020). Further, while in the current piece we have not discussed nor leveraged a reinforcement learning perspective for model-based data analyses (see Choi & Kim, 2011; Daptardar et al., 2019; Kwon et al., 2020; Wu et al., 2020), it is our hope that already starting to study natural and generalizable behaviors in systems neuroscience will precisely demand for developments in this area.**

Chandramouli Chandrasekaran

Thank you for your time, constructive comments, and help in improving this manuscript. It is much appreciated.