eLife

Reviewed Preprint
v2 • March 31, 2025
Revised by authors

Reviewed Preprint
v1 • January 6, 2025

Neuroscience

Quantification of the effect of hemodynamic occlusion in two-photon imaging

Baba Yogesh, Matthias Heindorf, Rebecca Jordan, Georg B Keller

Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland • Faculty of Natural Sciences, University of Basel, Basel, Switzerland • Simons Initiative for the Developing Brain, University of Edinburgh, Edinburgh, United Kingdom

eLife Assessment

This **important** study conducted experiments to quantify how neural activity independent changes in fluorescence might affect two-photon recordings when using diverse sensors. The researchers found a widespread presence of neural-activity-independent artifacts in two-photon imaging and provide **convincing** evidence that these artifacts are most likely caused by hemodynamic occlusion. Their findings underscore the importance of accounting for these artifacts when interpreting functional two-photon recordings.

https://doi.org/10.7554/eLife.104914.2.sa4

Abstract

The last few years have seen an explosion in the number of tools available to measure neuronal activity using fluorescence imaging (Chen et al., 2013; Feng et al., 2019; Jing et al., 2019; Sun et al., 2018; Wan et al., 2021). When performed in vivo, these measurements are invariably contaminated by hemodynamic occlusion artifacts. In widefield calcium imaging, this problem is well recognized. For two-photon imaging, however, the effects of hemodynamic occlusion have only been sparsely characterized. Here we perform a quantification of hemodynamic occlusion effects using measurements of fluorescence changes observed with GFP expression using both widefield and two-photon imaging. We find that in many instances the magnitude of signal changes attributable to hemodynamic occlusion is comparable to that observed with activity sensors. Moreover, we find that hemodynamic occlusion effects were spatially heterogeneous, both over cortical regions and across cortical depth, and exhibited a complex relationship with behavior. Thus, hemodynamic occlusion is an important caveat to consider when analyzing and interpreting not just widefield but also two-photon imaging data.

Introduction

Optical imaging of neuronal activity and neuromodulator concentration often involves recording the changes in fluorescence intensity of calcium or neuromodulator sensors expressed in neurons. In vivo, the hemodynamic changes in tissue around at the imaging location can modulate light transmission. Changes in blood volume through vasodilation and vasoconstriction as well as hemoglobin oxygenation change the transmission properties of the tissue. The combination of these effects is referred to as hemodynamic occlusion. Hemodynamic changes can be driven by local neuronal activity, a phenomenon known as neurovascular coupling (Attwell et al., 2010; Iadecola and Nedergaard, 2007) and are known to contribute significantly to activity measurements in techniques like fiber photometry (Zhang et al., 2022) and widefield microscopy (Ma et al., 2016; Waters, 2020). For these techniques, several methods have been proposed to control for the hemodynamic contribution to these activity measurements (Lohani et al., 2022; Valley et al., 2020; Zhang et al., 2022). However, potential effects of hemodynamic signals on neuronal activity measurements acquired using two-photon imaging remain less well characterized. While it has been shown that arteriole dilation decreases fluorescence recorded in immediately underlying neurons, it is still unclear how large these effects are in a behaving animal where both neuronal activity dynamics and blood pressure and flow variance are likely larger than as previously measured under anesthesia (Shen et al., 2012).

Here, we characterized hemodynamic signals in two-photon imaging of mouse cortex during visuomotor behaviors. We expressed an activity independent marker (GFP) in mouse cortical neurons and measured fluorescence changes using two-photon imaging while mice were interacting with a virtual environment. We found changes in GFP fluorescence in response to locomotion and visual stimuli of a magnitude comparable to that measured with common activity sensors like GCaMP and GRAB variants. GFP signals were apparent both in population average responses and at individual neuron level. Moreover, these signals were heterogeneous over dorsal cortex and cortical layers. We compared the size of GFP signals to the calcium responses reported by GCaMP6 sensors, and to the neuromodulator signals reported by the GRAB sensors for dopamine (GRAB-DA1m), serotonin (GRAB-5HT1.0), acetylcholine (GRAB-ACh3.0), and norepinephrine (GRAB-NE1m). We found that a large fraction of the GRAB response to the different neuromodulators was explained by the GFP signal, consistent with an earlier study characterizing this relationship for GRAB-5HT1.0 and GRAB-5HT3.0 (Ocana-Santero et al., 2024). Based on our findings, we speculate that the primary driver of the hemodynamic signals in two-photon imaging is hemodynamic occlusion. If so, then this problem is likely primarily mitigated by using high dynamic range sensors with low baseline fluorescence, and best controlled for using activity independent fluorescence measurements.

Results

To quantify hemodynamic occlusion in two-photon microscopy we imaged GFP in mouse cortex during behavior. We used an AAV vector (AAV2/1-Ef1α-eGFP-WPRE) injected in primary visual cortex (V1) and anterior cingulate cortex (ACC) to express GFP in cortical neurons (**Figure S1**). Given that the Ef1α promoter biases expression to excitatory neurons, we estimate that most (95%) of the GFP-expressing neurons were excitatory (Attinger et al., 2017). We then used a two-photon microscope to record GFP signals while mice explored a virtual tunnel. Mice were head-fixed on a spherical treadmill surrounded by a toroidal screen and exposed to a set of visuomotor conditions known to activate neurons in V1 and ACC (Attinger et al., 2017; Keller et al., 2012; Leinweber et al., 2017; Zmarz and Keller, 2016). First, mice were exposed to a closed loop condition during which running speed was coupled to visual flow speed in a virtual tunnel. Mice

were also exposed to an open loop condition, during which locomotion and visual flow in the virtual tunnel were not coupled, to darkness, and to the presentation of full-field drifting gratings. Throughout all experimental conditions, mice were free to locomote on the spherical treadmill and did so voluntarily. Unless stated otherwise, we will use 'response' to refer to the observed change in GFP fluorescence.

Locomotion onset and visual stimuli resulted in strong changes in GFP population responses

At the onset of locomotion, we found a transient increase in apparent GFP fluorescence in the population response of layer 2/3 (L2/3) neurons in V1 (**Figure 1A**). This increase in fluorescence of approximately 1% $\Delta F/F_0$ was only marginally smaller than the fluorescence changes typically observed at locomotion onset with genetically encoded calcium indicators (approximately 2% $\Delta F/F_0$ for GCaMP6f (Widmer et al., 2022; Yogesh and Keller, 2023), approximately 1.5% $\Delta F/F_0$ for GCaMP3 (Keller et al., 2012), etc.). By contrast, presentation of full-field drifting gratings resulted in a decrease of the GFP population signal (**Figure 1B**). We observed a similar decrease in GFP population signal on the presentation of an optogenetic stimulation light (637 nm, 10mW after the objective) directed at V1 with either the virtual reality turned on or in the dark (**Figure 1C**). This response was likely visually driven as the decrease was twice as strong in dark than when the virtual reality was on, in line with the stimulation light being more visually salient in dark. Visuomotor mismatches (pauses in the visual flow feedback during locomotion), resulted in an average increase of apparent GFP fluorescence (**Figure 1D**). By design, visuomotor mismatch only occurs during times of locomotion. Because apparent GFP fluorescence increases during locomotion, we estimated the increase expected by chance using random triggers during locomotion to compute the 95% confidence interval in **Figures 1D, 1H**, and **1L**. GFP responses to mismatch were only briefly outside these 95% confidence intervals. Thus, most stimuli tested drive measurable GFP responses at the population level. However, responses in individual neurons varied considerably (**Figure S2**).

To test whether stimulus triggered GFP responses depend on imaging depth and cortical area, we repeated these experiments in layer 5 (L5) of V1 and in L2/3 of ACC. In L5, we found that apparent fluorescence changes at locomotion onset were smaller than those observed in L2/3 (**Figure 1E**). As in L2/3, full-field drifting grating stimuli (**Figure 1F**) and optogenetic stimulation light (**Figure 1G**) both resulted in a significant decrease in apparent GFP fluorescence, while visuomotor mismatch responses were only barely above chance (**Figure 1H**). In L2/3 of ACC, locomotion onset resulted in a strong increase of apparent GFP fluorescence (**Figure 1I**), while we found only small grating responses (**Figure 1J**) and a positive response to optogenetic stimulation light (**Figure 1K**). We again found no evidence of visuomotor mismatch responses (**Figure 1L**). Thus, apparent GFP fluorescence changes can vary both as a function of recording depth and cortical area.

The observation that average population responses of GFP and calcium indicators are similar does not mean that the dynamic range of individual neurons is similar. When comparing the distribution of peak $\Delta F/F_0$ responses, it is evident that peak responses are much larger for calcium indicators (**Figures 2A** and **2B**). This is still the case when comparing trial averaged responses (**Figures 2C-2E**), but the difference between GFP and calcium indicator response distributions is smaller, in particular for locomotion onsets (**Figure 2C**). The main driver for this effect is likely the highly variable stimulus triggered calcium responses of many neurons.

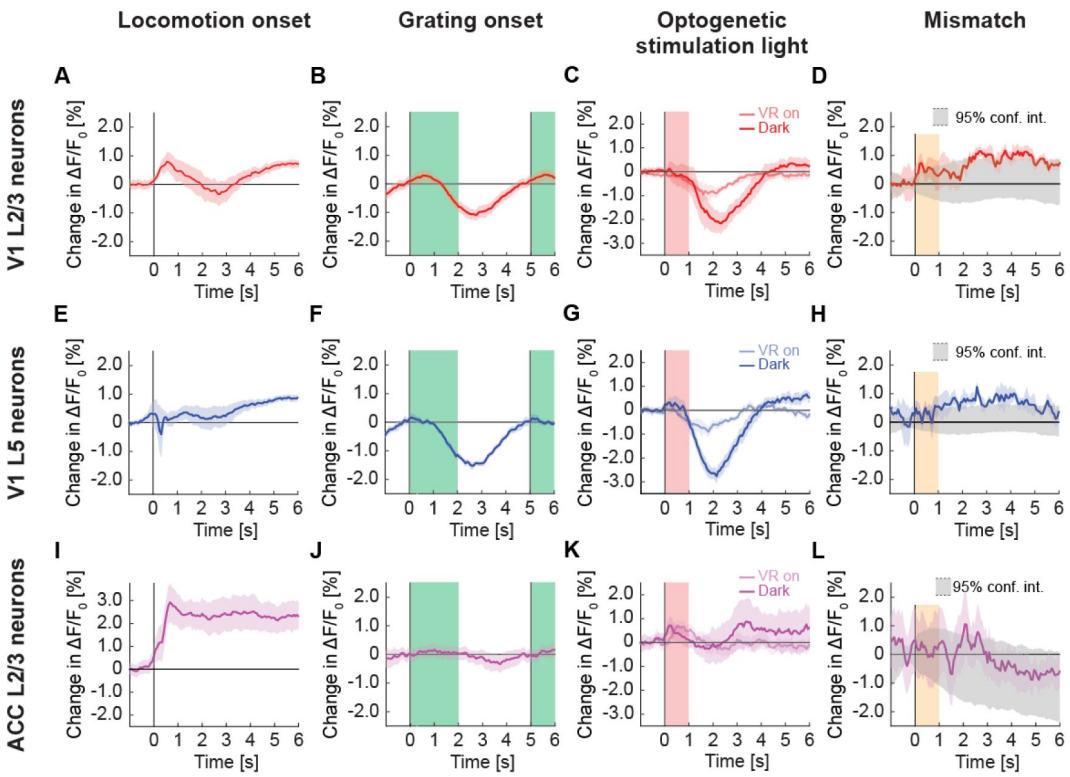


Figure 1.

GFP responses in two-photon imaging of neurons in V1 and ACC.

(**A**) Average change in apparent GFP fluorescence in L2/3 neurons in V1 on locomotion onset. Mean (solid lines) and the bootstrap standard error (shading) are calculated as hierarchical bootstrap estimate for each time bin.
(**B**) As in **A**, but for grating onsets in L2/3 neurons in V1. Green shading marks the duration of the gratings.
(**C**) As in **A**, but for responses to optogenetic stimulation light in L2/3 neurons in V1, with VR on (light red) or in dark (dark red). Pink shading marks the duration of the light stimulus.
(**D**) As in **A**, but for visuomotor mismatch in L2/3 neurons in V1. Orange shading marks the duration of visuomotor mismatch. As mismatch events are defined to occur during times of locomotion, and locomotion itself drives GFP responses, we would expect to find an increase in these signals by chance at mismatch. To correct for this, we quantified the distribution of signal amplitude on random triggers during locomotion (95% confidence interval, gray shading).
(**E**) As in **A**, but for L5 neurons in V1.
(**F**) As in **B**, but for L5 neurons in V1.
(**G**) As in **C**, but for L5 neurons in V1.
(**H**) As in **D**, but for L5 neurons in V1.
(**I**) As in **A**, but for L2/3 neurons in ACC.
(**J**) As in **B**, but for L2/3 neurons in ACC.
(**K**) As in **C**, but for L2/3 neurons in ACC.
(**L**) As in **D**, but for L2/3 neurons in ACC.

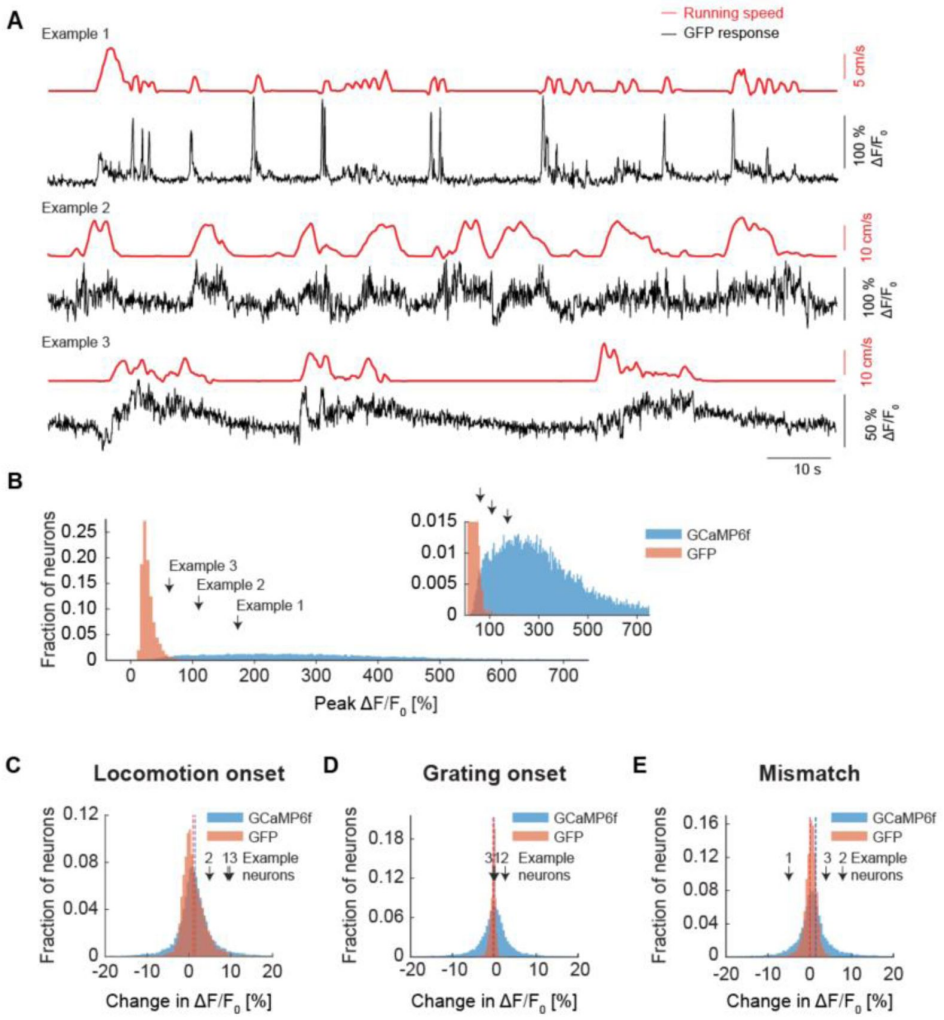


Figure 2.

Peak responses of GCaMP6f were larger than those observed with GFP.

(**A**) Example GFP response traces (black) of three neurons with corresponding running speed traces (red).
(**B**) Peak responses of GCaMP6f (blue) were typically an order of magnitude larger than those observed with GFP (orange). Inset is a zoomed-in version of the same data.
(**C**) Histogram of trial averaged responses on locomotion onset measured with two-photon imaging of calcium indicator (blue) and GFP (orange). Dotted lines indicate the mean response of the distribution of locomotion onset response in calcium indicator (blue) and GFP (orange).
(**D**) As in **C**, but for grating onset.
(**E**) As in **C**, but for visuomotor mismatch.

Individual neurons showed significant GFP responses to different stimuli

Next, we tested whether GFP signals are strong enough to result in significant responses when looking at individual neurons. Raw GFP traces exhibited fluctuations reminiscent of calcium responses but tended to be of lower peak amplitude (**Figure 2A**). It is possible that small, correlated changes in GFP signal, when averaged over the population, result in a significant response indistinguishable from calcium signals reported with GCaMP but are negligible at the level of individual neurons. Thus, we quantified the fraction of neurons that are significantly responsive to different stimuli. GFP expression levels were high enough for all neurons in our dataset such that peak $\Delta F/F_0$ response amplitudes were independent of baseline F_0 (**Figure S4**). Commensurate with the population averages (**Figure 1**), we found a surprisingly large fraction of neurons in V1, in both L2/3 (**Figure 3A**) and L5 (**Figure 3B**) responsive to locomotion onset and grating stimuli. In both locations, the fraction of neurons responsive to visuomotor mismatch was not different from chance. In ACC L2/3 (**Figure 3C**), a significant fraction of GFP labelled neurons was responsive to locomotion onset, while the fraction of neurons responsive to gratings and visuomotor mismatch was not different from chance. In all cases, the fraction of neurons responsive as measured with GFP was similar to the fraction of neurons responsive as measured with GCaMP (**Figures 3D-3F**). Thus, even at the single neuron level, hemodynamic influence on GFP remains sizable.

Changes in GFP fluorescence correlated with blood vessel dilation

One possible source of changes in GFP fluorescence is brain motion. However, we only find detectable brain motion during locomotion onsets and not the other stimuli (**Figure S5**), and even during locomotion onsets the movements are comparably small and follow different kinematics than the GFP signals. Finally, brain motion artefacts should on average decrease fluorescence, as they will move the cell out of the ROI. Thus, we speculated that the dominant change in GFP fluorescence is driven by hemodynamic occlusion (see Discussion for alternative possibilities). Blood absorbs light. This absorption changes both as a function of blood volume and blood oxygenation. Thus, the blood vessels on the cortical surface that are between the GFP labeled neurons and the objective of the microscope act to variably occlude the fluorescence measurement. As they constrict, dilate, or change oxygenation, they influence the transmission of both excitation and emission light similar to how an LCD (liquid crystal display) functions. One testable prediction of this hypothesis is that the cross section of blood vessels visible in a given field of view should correlate inversely with the apparent GFP fluorescence of surrounding cells. This makes the simplifying assumption that the cross section of surface blood vessels, which likely have the strongest contribution to occlusion, correlates with that of blood vessels in the imaging plane. To quantify the correlation between blood vessel cross section and apparent GFP fluorescence, we first estimated the cross section of selected blood vessels in the imaging plane (**Figures 4A-4C**). Blood vessels in V1 dilated systematically, for example upon light stimulation (**Figures 4A1** and **4B1**), and contracted, for example on locomotion onset in ACC (**Figure 4C1**). The estimation of blood vessel cross section worked well for high contrast boundaries but failed in many instances of low contrast boundaries due to noise in the boundary estimates in single imaging frames. A proxy for blood vessel cross section that does not rely on thresholding is the average fluorescence of a region of interest of fixed size and fully contained within the blood vessel (**Figures 4A2, 4B2, 4C2**). The intuition here is that there should be no fluorescence inside the blood vessel. Thus, fluorescence measured inside the blood vessel must come from the fact that the point spread function of the microscope extends to regions outside of the blood vessel. As the blood vessel contracts (or dilates) more (or less) of GFP in the tissue around the blood vessel will move into the point spread function. Comparing both metrics to the average apparent GFP fluorescence changes of the surrounding cells (**Figures 4A3, 4B3, 4C3**), we found that both blood

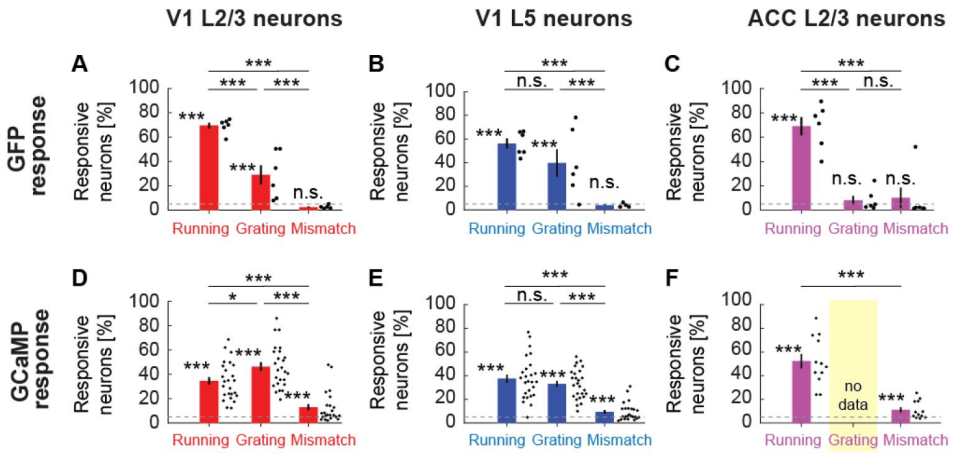


Figure 3.

Fraction of neurons with significant GFP responses to different stimuli.

(**A**) The fraction of GFP expressing V1 L2/3 neurons responsive to locomotion, full-field grating, and visuomotor mismatch onset, quantified for all imaging sites. Each dot is an imaging site, jittered along the x-axis to make them visible. Whiskers mark the SEM. Dashed line indicates chance levels (p = 0.05). n.s.: not significant; *: p < 0.05; **: p < 0.01; ***: p < 0.001; see **Table S1** for all statistical information.
(**B**) As in **A**, but for V1 L5 neurons.
(**C**) As in **A**, but for ACC L2/3 neurons.
(**D**) As in **A**, but for V1 L2/3 neurons expressing GCaMP6f.
(**E**) As in **A**, but for V1 L5 neurons expressing GCaMP6f.
(**F**) As in **A**, but for ACC L2/3 neurons expressing a GCaMP variant (data are from experiments using GCaMP3, GCaMP5, GCaMP6s, and GCaMP6f). Note, we have no grating presentation for this population.

vessel area and fluorescence explain over 80% of the variance of apparent GFP fluorescence changes (**Figures 4A4, 4B4, 4C4**) in these example data. Here, variance explained is defined as the squared linear correlation coefficient between blood vessel area and fluorescence. Quantifying this for all data, we found that the fraction of variance explained was 50% or higher for most event triggered averages (**Figure 4D**) as well as across the entire recording duration (**Figures 4E** and **S6**). The variance explained tended to be lower in ACC than in V1; we suspect that this is caused by the superior sagittal sinus contributing more strongly to hemodynamic occlusion in ACC, and that the diameters of local blood vessels are less well correlated with that of this large blood vessel. Consistent with vasoconstriction causing the increase of signal on locomotion onset, we sometimes find patterns of fluorescence changes that look like blood vessels that pass immediately above the imaging plan (**Figures S7A** and **S7B**). More typical however, is a broad increase that likely is driven by hemodynamic changes in blood vessels closer to the surface of the brain (**Figures S7C** and **S7D**).

GFP responses were visuomotor context sensitive and could not be explained by a linear combination of responses to constituent stimuli

Next, we investigated how stereotypical GFP responses are across different visuomotor conditions and cortical regions, and whether GFP responses can be explained by a linear combination of a locomotion driven component and a visually driven component. We found that in V1 L2/3, locomotion onset responses were similar in closed loop and open loop, but different in the dark (**Figure 5A**). During a closed loop locomotion onset, there is concurrent onset of visual flow with the locomotion onset. During open loop and dark locomotion onsets there is no concurrent onset of visual flow, and they only differ in the average visual input. The similarity between open and closed loop running onsets could be explained by a dominant locomotion related signal and a negligible visual onset response. If so, the smaller GFP responses in dark locomotion would need to be the consequence of the lower average visual input. This demonstrates that GFP responses, similar to neuronal calcium responses, can be non-linear combinations of component responses. This is perhaps not surprising, assuming local blood flow is primarily related to local neuronal activity. In L5 of V1, locomotion onset responses in open loop and dark conditions were comparable, while those in closed loop conditions had a delayed transient dip (**Figure 5B**). Thus, hemodynamic responses can exhibit a strong depth dependence. In ACC L2/3 neurons all three types of locomotion onsets resulted in increases of apparent fluorescence (**Figure 5C**). Thus, hemodynamic responses are difficult to predict from component signals (**Figure S3**), and do not generalize across depth or cortical areas. Finally, while there appears to be some correspondence to local neuronal activity in that locomotion onsets are more strongly modulated by visual context in V1 than in ACC, it is not immediately obvious how the GFP signals relate to responses measured with calcium indicators.

Locomotion increased correlation in GFP signals

Another metric frequently used in the analysis of two-photon data are correlation-based analyses. Locomotion, for example, has been shown to decorrelate the activity of excitatory neurons in cortex (Aydın et al., 2018; Dadarlat and Stryker, 2017; Erisken et al., 2014; Yogesh and Keller, 2023). Given the sizable fraction of neurons significantly responsive to locomotion, we speculated that correlations between neurons likely are influenced by hemodynamic responses. We tested this by quantifying the effect of locomotion on the pairwise correlations of GFP signals between individual neurons. We found that pairwise correlations were systematically increased during locomotion (**Figure 6**). This was true for neurons in V1 L2/3 (**Figure 6A**), V1 L5 (**Figure 6B**), and ACC L2/3 (**Figure 6C**). Thus, hemodynamic responses can have a strong influence on pairwise correlations, and in the case of locomotion, this effect is opposite to that observed with calcium indicators (Aydın et al., 2018; Dadarlat and Stryker, 2017; Erisken et al., 2014;

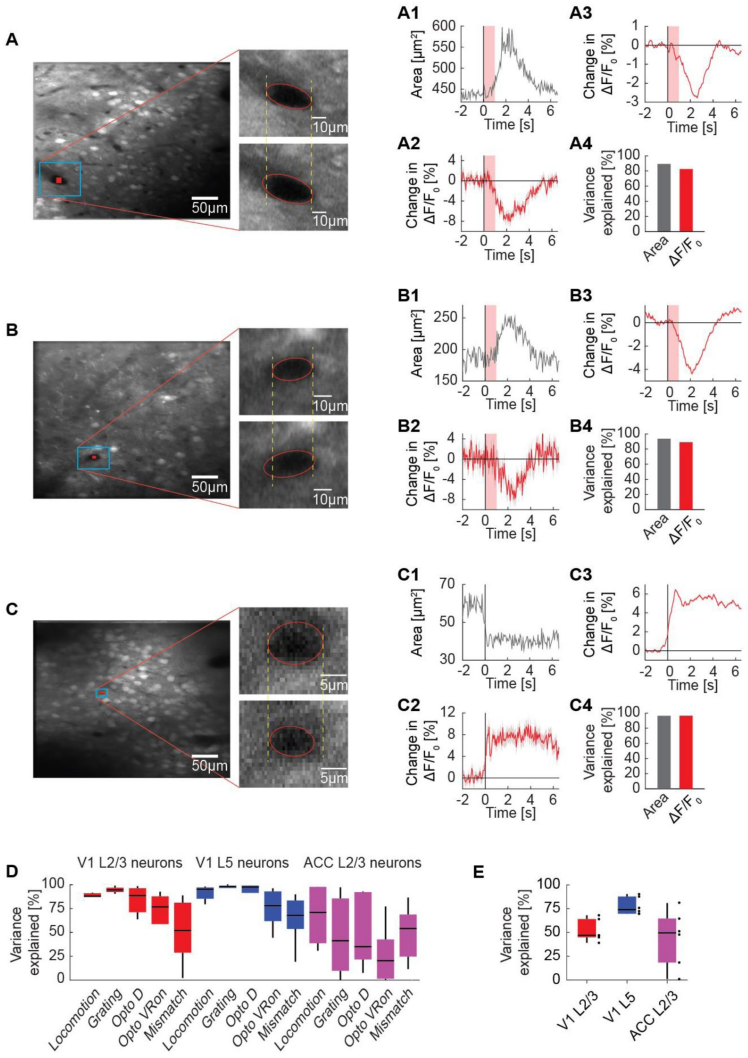


Figure 4.

Correlation between blood vessel area and GFP responses.

(**A**) Example average fluorescence image from an imaging site in L2/3 of V1 under a two-photon microscope. The fluorescence signal is averaged over 22 trials of optogenetic stimulation light directed at V1 while the mouse was in the dark. Shown on the right is an example blood vessel before (upper panel) and after (lower panel) stimulation.
(**A1**) Average area of the example blood vessel shown in **A** (in blue box) on light stimulation. The area of the indicated blood vessel is extracted from the trial averaged two-photon image. Shading in pink indicates the optogenetic light presentation.
(**A2**) Average change in GFP signal from a ROI positioned within the blood vessel (shown in **A** in red). Mean (red) and SEM (shading) were calculated over light stimulation trials.
(**A3**) Average change in GFP signal from all neurons (268 neurons) simultaneously imaged in the same site. Mean (red) and SEM (shading) were calculated over light stimulation trials.
(**A4**) Variance in average GFP signal in this example L2/3 site in V1 explained by area of the example blood vessel and by the change in GFP signal from a ROI within the blood vessel.
(**B**) As in **A**, from an imaging site in L5 of V1.
(**C**) As in **A**, from an imaging site in L2/3 of ACC on locomotion onset.
(**D**) Percentage of variance in neuronal GFP signal to individual events explained by the change in GFP signal from blood vessel ROIs. Boxes show 25th and 75th percentile, central mark is the median, and the whiskers extend to the most extreme datapoints not considered outliers.
(**E**) Percentage of variance in neuronal GFP signal over the entire recording explained by blood vessel GFP signal per imaging location. Each datapoint is one imaging site. Boxes show 25th and 75th percentile, central mark is the median, and the whiskers extend to the most extreme datapoints not considered outliers.

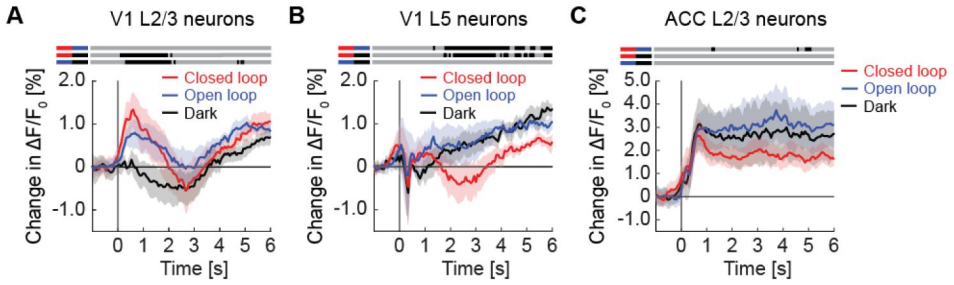


Figure 5.

GFP responses at locomotion onset depended on visuomotor condition and cortical region.

(**A**) Average GFP response in L2/3 neurons in V1 on locomotion onset in closed loop (in red), in open loop (in blue), and in darkness (in grey). Here and in subsequent panels, bins with a significant difference ($p < 0.01$) are indicated by a black line above the plot (lines compared are marked on the left by pairs of colored horizontal bars); those with $p > 0.01$ are marked gray. Mean (solid lines) and the bootstrap standard error (shading) are calculated as hierarchical bootstrap estimate for each time bin.
(**B**) As in **A**, but for L5 neurons in V1.
(**C**) As in **A**, but for L2/3 neurons in ACC.

Yogesh and Keller, 2023). This is consistent with the interpretation that hemodynamic changes on locomotion occur at a larger spatial scale than changes in neuronal activity patterns and thus tend to increase correlations between neurons.

GFP responses in widefield imaging were similar in magnitude to those in two-photon imaging

In widefield calcium imaging, hemodynamic responses are better characterized and widely recognized as a confounding problem (Allen et al., 2017; Ma et al., 2016; Valley et al., 2020). To directly contrast hemodynamic widefield signals with those observed in two-photon imaging, we repeated our experiments by performing widefield imaging of GFP responses. We expressed GFP pan-neuronally across dorsal cortex using either a retro-orbital injection of AAV-PHP.eB-EF1α-eGFP, or transgenic mice that express GFP under a *c-fos* promoter (**Table S2**). As with two-photon imaging (**Figure 1**), and consistent with previous work (Allen et al., 2017; Valley et al., 2020), we found strong GFP responses across dorsal cortex (**Figure 7**). Locomotion and grating onsets resulted in GFP responses in both V1 (**Figures 7A** and **7B**) and ACC (**Figures 7D** and **7E**). During visuomotor mismatch, we found no significant response in either of these regions (**Figures 7C** and **7F**). All these responses were greatly attenuated or absent when we imaged the cortex without expressing a fluorescent indicator (**Figure S8**), arguing in favor of hemodynamic occlusion of fluorescence from the genetically expressed indicators as the underlying cause for most of these responses in the GFP signal.

GRAB sensors in cortex displayed similar responses to those recorded in GFP imaging

Finally, we compared the magnitude of hemodynamic responses to those recorded with the relatively low dynamic range GRAB sensors using both two-photon and widefield imaging. For this we used the GRAB-DA1m dopamine sensor (Sun et al., 2018), the GRAB-5HT1.0 serotonin sensor (Wan et al., 2021), the GRAB-ACh3.0 acetylcholine sensor (Jing et al., 2019), and the GRAB-NE1m(Feng et al., 2019) norepinephrine sensor (Feng et al., 2019). We first imaged GRAB-DA1m, GRAB-5HT1.0, and GRAB-ACh3.0 using two-photon imaging in L2/3 of V1. To estimate hemodynamic responses from GRAB sensors, we compared responses of regions of interest (ROI) selected in neuropil and those selected in blood vessels (**Figures 8A, 8E** and **8I**). We reasoned that the signal in neuropil regions is a combination of a true GRAB signal and hemodynamic occlusion, while the signal in a ROI selected inside of a blood vessel should be dominated by the blood vessel diameter. We found that the GRAB-DA1m response to locomotion and full-field grating onsets (**Figure 8B** and **8C**) was similar to GFP responses in V1 L2/3 (**Figure 1**). There were responses to locomotion and grating onset in the GRAB-DA1m signal (**Figure 8B** and **8C**). In both cases neuropil and blood vessel ROI responses were similar. Also, similar to the GFP responses, we found no evidence of a response to visuomotor mismatch in the GRAB-DA1m signal (**Figure 8D**). We observed similar responses when imaging GRAB-5HT1.0 in V1 (**Figures 8F-8H**). However, these responses were smaller in amplitude, likely driven by the overall dimmer fluorescence of the GRAB-5HT1.0 sensor. Thus, it is likely that both GRAB-DA1m and GRAB-5HT1.0 responses recorded with two-photon imaging are dominated by hemodynamic occlusion.

A response that was distinct from GFP, GRAB-DA1m, and GRAB-5HT1.0 signals was the response in GRAB-ACh3.0 on locomotion onset (**Figure 8J**). Here we found a prominent increase in GRAB-ACh3.0 signal that was three-fold higher than either the GFP signal or the GRAB-ACh3.0 signals measured in blood vessel ROIs. This makes it more likely that the observed fluorescence changes are driven by a mixture of hemodynamic occlusion and GRAB-ACh3.0 responses. The GRAB-ACh3.0 response to grating onset (**Figure 8K**), however, was similar to GFP responses. And again, we found no response to visuomotor mismatch in GRAB-ACh3.0 (**Figure 8L**). To test whether the absence of response differences between GRAB-DA1m, GRAB-5HT1.0 and GFP signals was simply a

Figure 6.

Locomotion increased pairwise correlation of GFP signals between neurons.

(**A**) Average pairwise correlation in GFP signal of V1 L2/3 neurons while mice were stationary and while they were locomoting. Each dot is the mean pairwise correlation of one neuron to all other neurons in the same imaging site. Colors indicate data from different mice.
(**B**) As in **A**, but for V1 L5 neurons.
(**C**) As in **A**, but for ACC L2/3 neurons.

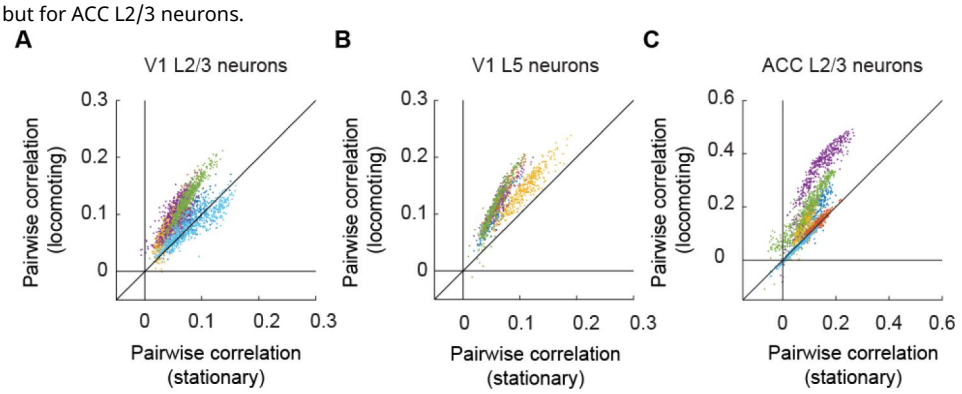
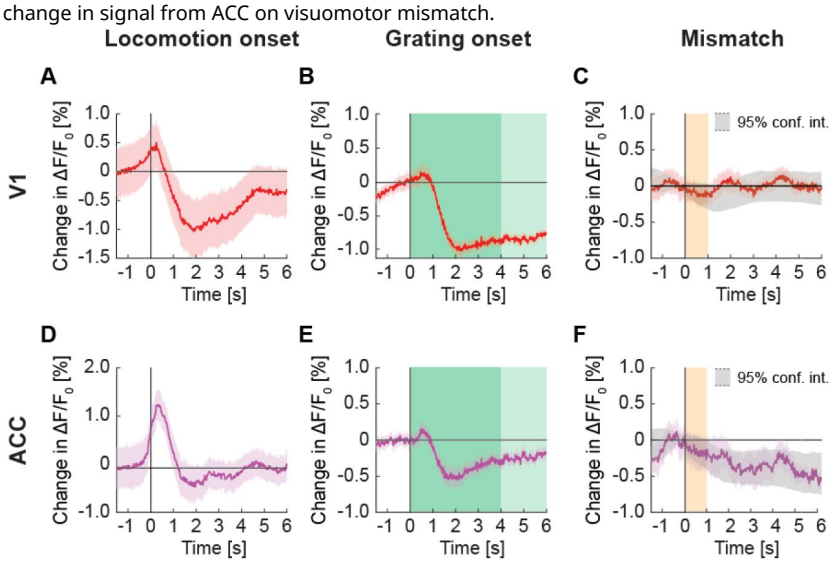


Figure 7.

Signals in widefield fluorescence imaging of V1 and ACC in mice expressing GFP.

(**A**) Average change in signal from V1 on locomotion onset. Mean (solid lines) and the bootstrap standard error (shading) are calculated as hierarchical bootstrap estimate for each time bin.
(**B**) As in **A**, but for grating onset in V1. Grating duration were randomly selected between 4 s and 8 s (green shading).
(**C**) As in **A**, but on visuomotor mismatch in V1. Orange shading marks the duration of visuomotor mismatch. Gray shading marks the 95% confidence interval.
(**D**) As in **A**, but for change in signal from ACC on locomotion onset.
(**E**) As in **B**, but for change in signal from ACC on grating onset.
(**F**) As in **C**, but for change in signal from ACC on visuomotor mismatch.



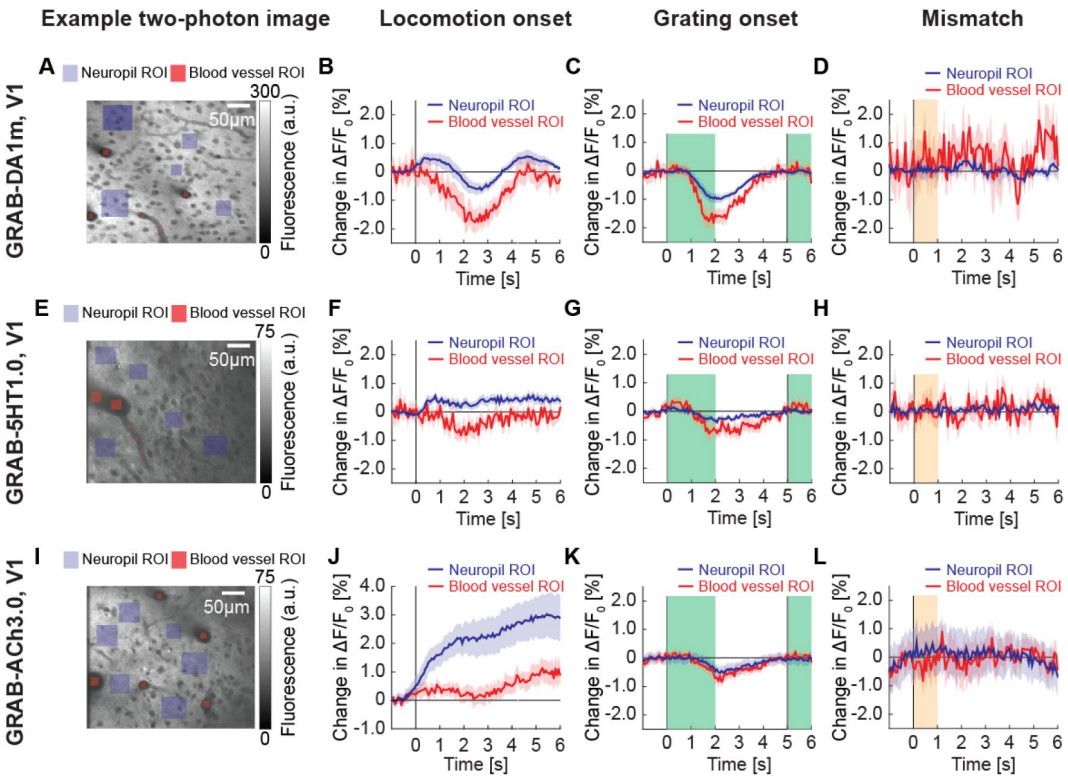


Figure 8.

GRAB signals in V1 with two-photon imaging.

(**A**) Example two-photon image taken in V1 of mouse expressing GRAB-DA1m, with neuropil ROIs (in violet) and blood vessel ROIs (in red).
(**B**) Average change in GRAB-DA1m signal in V1 on locomotion onset in ROIs over neuropil (in black) and in ROIs in blood vessels (in red). Mean (solid lines) and the bootstrap standard error (shading) are calculated as hierarchical bootstrap estimate for each time bin.
(**C**) As in **B**, but for response to grating onset. Green shading marks the duration of the gratings.
(**D**) As in **B**, but for response to visuomotor mismatch. Orange shading marks the duration of visuomotor mismatch.
(**E**) Example two-photon image taken in V1 of mouse expressing GRAB-5HT1.0, with neuropil ROIs (in green) and blood vessel ROIs (in red).
(**F**) Average change in GRAB-5HT1.0 signal in V1 on locomotion onset in ROIs over neuropils (in black) and in ROIs in blood vessels (in red).
(**G**) As in **F**, but for response to grating onset.
(**H**) As in **F**, but for response to visuomotor mismatch.
(**I**) Example two-photon image of mouse expressing GRAB-ACh3.0 in V1, with neuropil ROIs (in green) and blood vessel ROIs (in red).
(**J**) Average change in GRAB-ACh3.0 fluorescence in V1 in response to locomotion onset in ROIs over neuropil (in black) and in ROIs over blood vessels (in red).
(**K**) As in **J**, but for response to grating onset.
(**L**) As in **J**, but for response to visuomotor mismatch.

result of imaging in V1, where both dopamine and serotonin release is weaker than in frontal areas of cortex (Berger et al., 1991; Hamada et al., 2023), we repeated these experiments in ACC. Also here, we found no evidence of GRAB responses that could not be explained by hemodynamic occlusion (**Figure S9**). A strong influence of hemodynamic signals was also apparent in widefield imaging of GRAB-NE1m. We found that the widefield GRAB-NE1m responses (**Figure 9**) were similar to GFP responses (**Figure 7**). Thus, in certain cases of low dynamic range sensors, one can readily measure responses, but these are not always easily distinguished from those driven by hemodynamic occlusion.

Discussion

Our data demonstrate that in vivo two-photon fluorescence imaging during behavior is subject to relatively large and fast changes in measured fluorescence independent of any sensor typically used to measure functional calcium, voltage, or peptide changes. The surprising finding to us, was not that these changes are observed, but rather how large they are on a population average compared to fluorescence changes measured with functional sensors. There are several possible sources of these fluorescence changes:

Brain movement artifacts

Many body movements result in the movement of the brain relative to the skull. In our experience, locomotion onset as well chewing and licking (or other movements that involve the masseter muscles) are the strongest drivers of such movements. Use of a cranial window prevents much of the movement in the axis perpendicular to the window (which typically is the optical axis). To mitigate this problem, we discard any experimental preparation with visible z-motion, and correct x/y motion using a full-frame registration algorithm. For locomotion onsets, brain movement kinematics are different in shape than the GFP responses we observe, and in the case of visual stimulation, we find no evidence of stimulus driven brain movements (**Figure S5**). Finally, given that the regions of interest used to quantify fluorescence changes are selected based on a mean position of the neuron, movement away from the mean position should result in an average decrease in observed fluorescence. Thus, while there certainly are brain movement related artifacts that can strongly influence two-photon measurements if not properly mitigated, these artifacts cannot explain visual stimulation induced fluorescence changes, and likely only explain the initial, fast decrease of fluorescence observed in a subset of neurons at locomotion onset (**Figure S2**).

Stimulation light contamination

Light from the visual display system can leak into the optical pathway of the two-photon microscope. This results in stimulus locked changes in fluorescence that can occur both for visual stimulation, but also for closed loop locomotion onsets, as locomotion triggers an onset of visual flow. This can be partially mitigated by physically shielding the imaging site from stimulation light. However complete shielding is difficult (e.g. some of the stimulation light likely travels through the eye to the brain and into the microscope – the fact that light quite efficiently travels this route is what has enabled convenient eye tracking in two photon imaging as the stimulation light from the two-photon laser illuminates the pupil from the inside). In two photon imaging this problem can however be almost completely solved by synchronizing all light sources in the experimental setup to the dead times of the microscope (e.g. the turnaround times of the resonant scanner) (Leinweber et al., 2014), as we have also done here. Finally, stimulation light contamination exhibits essentially immediate onset and offset kinetics and is relatively simple to recognize. Thus, we are confident that stimulation light contamination does not significantly contribute to the effects we describe here.

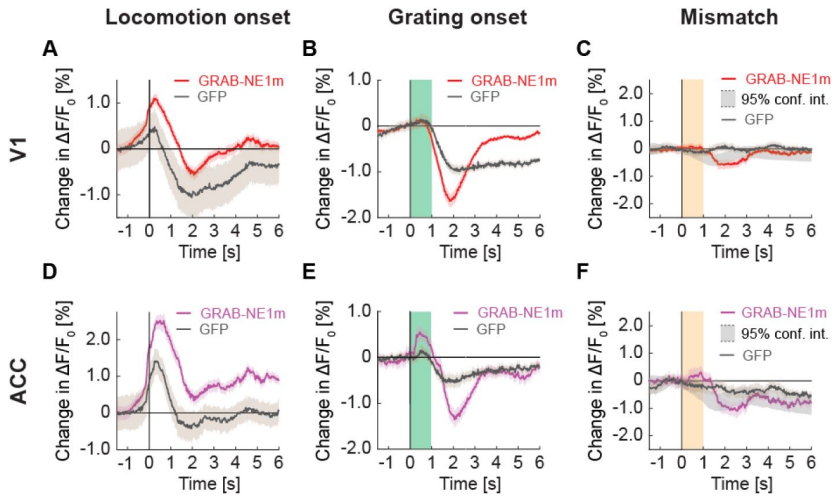


Figure 9.

Signals of V1 and ACC in widefield imaging in mice expressing GRAB-NE1m.

(**A**) Average change in signal from V1 on locomotion onset. Mean (solid lines) and the bootstrap standard error (shading) are calculated as hierarchical bootstrap estimate for each time bin. Here and below, GFP response (in gray) from **Figure 7** overlaid.
(**B**) As in **A**, but for grating onset in V1. Green shading marks the duration of the gratings.
(**C**) As in **A**, but for visuomotor mismatch in V1. Orange shading marks the duration of visuomotor mismatch. Gray shading marks the 95% confidence interval.
(**D**) As in **A**, but for change in signal from ACC on locomotion onset.
(**E**) As in **B**, but for change in signal from ACC on grating onset.
(**F**) As in **C**, but for change in signal from ACC on visuomotor mismatch.

GFP's sensitivity to pH changes

GFP changes its fluorescence with changes in pH (Kneen et al., 1998). And intracellular pH systematically changes with neuronal activity. While the relationship between neuronal activity and intracellular pH is complex and not fully understood, increases in neuronal activity tend to result in decreases of intracellular pH. Thus, the activity driven pH change should decrease GFP fluorescence. What we find, however, in many cases is a positive correlation between neuronal activity and GFP fluorescence. More importantly, however, intracellular pH changes driven by neuronal activity are on the order of 0.1 around a resting pH somewhere between 7 and 8 (Chesler, 2003; Raimondo et al., 2013). While GFP is indeed sensitive to pH, this is primarily the case for pH values below 6. The fluorescence changes above a pH of 7 are comparably small, and likely negligible when the changes are in the range of 0.1 (Dos Santos et al., 2020; Kneen et al., 1998). Thus, we don't think the pH sensitivity of GFP substantially contributes to the effects we observe.

Hemodynamic tissue deformation

A second type of movement artifact is produced by local tissue deformations caused by blood vessel diameter changes. Cells in close proximity to a blood vessel are displaced when blood vessels dilate and contract. These effects are primarily visible in cells in the immediate vicinity (a few cell body diameters) of blood vessels. As with brain movement driven artefacts, the net effect of such tissue deformations is to displace the cell away from the region of interest used to measure fluorescence and hence result they tend to result in apparent fluorescence decreases on average. These effects certainly also contribute to the signals we report but given the observed increase in fluorescence on running onset, the effect of hemodynamic tissue deformation are likely smaller compared to that of occlusion. Moreover, these effects are primarily visible when there is a high fluorescence inside a ROI and low fluorescence immediately surrounding it (e.g. as is the case when expressing GFP or GCaMP). When expressing GRAB sensors, labelling is much more homogeneous, and there are no distinct contrast changes at the edges of ROIs. Thus, the fact that the hemodynamic driven signals look very similar when labelling is primarily confined to somata (**Figure 1**), compared to using GRAB sensors (**Figure 8**), is consistent with our interpretation that hemodynamic tissue deformations are not the primary driver of the signals we report.

Hemodynamic occlusion

Blood absorbs stimulation and emission light. As a function of volume and composition of blood, light absorption can change dynamically. Changes in neuronal activity in the cortex are associated with hemodynamic changes through neurovascular coupling (Ruff et al., 2024) that are layer specific (Mächler et al., 2021). Cells directly under arterioles have been shown to be occluded by changes in arteriole diameter (Shen et al., 2012). In brain structures like cortex that have a relatively dense blood vessel network on the surface, this problem is augmented as any blood vessel within the imaging light cone between the surface of the brain and the recorded cell will contribute to the occlusion. Using an objective with a numerical aperture of 1, this is a light cone with a half angle of approximately 45 degrees. Thus, when e.g. imaging a neuron at a depth of 500 µm, all surface vessels in circle around location of the neuron of 1 mm diameter will contribute to the occlusion. Even ignoring occlusion contributions that arise from changes to the composition of blood (e.g. oxygenation changes), the changes driven by blood vessel cross section changes are likely sufficient to explain the effect sizes we observe. Thus, we conclude that hemodynamic occlusion is the primary contributor to the effects we observe. Some of the other factors discussed above certainly also contribute to the effects we describe, and there may yet be other influences we have not thought of yet. Importantly however, all these factors similarly would contribute to signals measured with functional indicators.

How to mitigate the problem

There are likely two approaches: use a second imaging channel for concurrent isosbestic illumination of the sensor or estimate the hemodynamic contribution in a separate set of experiments using GFP. For the calcium indicator GCaMP, isosbestic imaging has been shown to work in widefield imaging (Allen et al., 2017; Couto et al., 2021). However, imaging at the isosbestic wavelength concurrently with the functional measurement would require a second tunable two-photon light source. In the case of fiber photometry (Zhang et al., 2022) and widefield imaging (Ma et al., 2016; Scott et al., 2018; Valley et al., 2020) hemodynamic influence is well documented, and there are correction methods published (Allen et al., 2017; Ma et al., 2016). But even here methods based on isosbestic illumination can be difficult to implement accurately, either because light sources used for isosbestic illumination are not tunable or not available at the exact wavelength necessary, or when using LEDs, have emission spectra are too broad. One could consider trying a model based approach to estimate hemodynamic occlusion effects based on GFP measurements, as can be done for widefield imaging (Valley et al., 2020). However, because hemodynamic occlusion changes across the cortical surface and depth, separate GFP calibrations are likely necessary for different regions and layers imaged. Finally, a simple but laborious approach is to repeat key experiments using GFP to test how much of the functional effects can be explained by non-functional effects. While arguably more than a decade too late (Keller et al., 2012), this is what our aim was here.

Finally, given that hemodynamic changes are likely driven by neuronal activity through neurovascular coupling (Ruff et al., 2024), which can be highly specific to cortical layer (Mächler et al., 2021), these signals could be a useful proxy to estimate neuronal activity. Assuming one would identify a neuronal cell-type with activity that correlates well with local hemodynamic changes, blood vessel diameter measurements could be used as a relatively simple secondary measurement in functional two-photon imaging.

Methods

Key Resources Table

Reagent type (species) or resource	Designation	Source or reference	Identifier	Additional information
Strain, strain background (adeno-associated virus)	AAV2/1-Ef1α-eGFP-WPRE (10^{12} GC/ml)	FMI vector core	vector.fmi.ch	
Strain, strain background (adeno-associated virus)	AAV2/1-Ef1α-GCaMP6f-WPRE (10^{11}-10^{14} GC/ml)	FMI vector core	vector.fmi.ch	
Strain, strain background (adeno-associated virus)	AAV2/9-hSyn-GRAB-DA1m (10^{12} GC/ml)	FMI vector core	vector.fmi.ch	
Strain, strain background (adeno-associated virus)	AAV2/9-hSyn-GRAB-5HT1.0 (10^{13} GC/ml)	FMI vector core	vector.fmi.ch	
Strain, strain background (adeno-associated virus)	AAV2/9-hSyn-GRAB-ACh3.0 (10^{13} GC/ml)	FMI vector core	vector.fmi.ch	
Strain, strain background (adeno-associated virus)	AAV-PHP.eB-EF1α-eGFP (10^{13} GC/ml)	FMI vector core	vector.fmi.ch	
Strain, strain background (adeno-associated virus)	AAV-PHP.eB-hSyn-GRAB-NE1m (10^{13} GC/ml)	FMI vector core	vector.fmi.ch	
Chemical compound, drug	Fentanyl citrate	Actavis	CAS 990-73-8	Anesthetic compound
Chemical compound, drug	Midazolam (Dormicum)	Roche	CAS 59467-96-8	Anesthetic compound
Chemical compound, drug	Medetomidine (Domitor)	Orion Pharma	CAS 86347-14-0	Anesthetic compound
Chemical compound, drug	Ropivacaine	Presenius Kabi	CAS 132112-35-7	Analgesic compound
Chemical compound, drug	Lidocaine	Bichsel	CAS 137-58-6	Analgesic compound
Chemical compound, drug	Buprenorphine	Reckitt Benckiser Healthcare	CAS 52485-79-7	Analgesic compound
Chemical compound, drug	Humigel	Virbac	-	Ophthalmic gel
Chemical compound, drug	Flumazenil (Anexate)	Roche	CAS 78755-81-4	Anesthetic antagonist
Chemical compound, drug	Atipamezole (Antisedan)	Orion Pharma	CAS 104054-27-5	Anesthetic antagonist
Chemical compound, drug	Metacam	Boehringer Ingelheim	CAS 71125-39-8	Analgesic compound
Chemical compound, drug	N-Butyl-2-cyanoacrylate	Braun	CAS 6606-65-1	Histoacryl
Chemical compound, drug	Dental cement (Paladur)	Heraeus Kulzer	CAS 9066-86-8	

Genetic reagent *Mus musculus*	*C57BL/6*	Charles River	-	
Genetic reagent *Mus musculus*	*B6J.129S6-Chat$^{tm2(Cre)Lowl}$/MwarJ* Alias used here: ChAT-IRES-Cre	Jackson Laboratories	RRID:IMSR_JAX:028861	Cre expression in cholinergic neurons
Genetic reagent *Mus musculus*	*B6J.FVB(Cg)-Tg(Tlx3-cre)PL56Gsat/Mmucd* Alias used here: Tlx3-Cre	MMRRC	RRID:MMRRC_041158-UCD	Cre expression in a subset of L5 neurons
Genetic reagent *Mus musculus*	*B6.Cg-Tg(Fos/EGFP)1-3Brth/J* Alias used here: fosGFP	Jackson Laboratories	RRID:IMSR_JAX:014135	Cre expression in a subset of cortical neurons
Software, algorithm	MATLAB (2023b)	The MathWorks	RRID:SCR_001622	Data analysis
Software, algorithm	LabVIEW	National Instruments	RRID:SCR_014325	Hardware control
Software, algorithm	Two-photon acquisition software	Keller laboratory	sourceforge.net/p/iris-scanning/	Data acquisition
Software, algorithm	Image data processing software	Keller laboratory	sourceforge.net/p/iris-scanning/calliope	Data processing
Software, algorithm	Python	python.org	RRID:SCR_008394	Virtual reality
Software, algorithm	Panda3D	panda3d.org	N/A	Virtual reality
Other	Virtual reality and two-photon setup	(Leinweber et al., 2014, 2017)	DOI: 10.3791/50885, 10.1016/j.neuron.2017.08.036	Hardware setup
Other	OBIS 673 nm LX	Coherent	Cat#1187194	Optogenetic stimulation laser
Other	LED	Prizmatix	UHP-T-595	Sham stimulation
Other	Titanium headplate	FMI/ETHZ workshop	N/A	Mice head-fixation
Other	Dental drill	Meisinger	N/A	For craniotomy

Mice

We used a total of 35 C57BL/6 mice, 4 fosGFP mice, 6 Tlx3-Cre mice, and 23 ChAT-IRES-Cre mice. Both male and female mice, 6 - 16 weeks old at the start of the experiment, were used. See **Table S2** for details of mouse inclusion for different figures. Between experiments, mice were group-housed in a vivarium (light/dark cycle: 12/12 hours). All animal procedures were approved by and carried out in accordance with the guidelines laid by the Veterinary Department of the Canton of Basel-Stadt, Switzerland.

Surgery

For all surgical procedures, mice were anesthetized with a mixture of fentanyl (0.05 mg/kg; Actavis), midazolam (5.0 mg/kg; Dormicum, Roche), and medetomidine (0.5 mg/kg; Domitor, Orion) injected intraperitoneally. Analgesics were applied perioperatively (2% lidocaine gel, meloxicam 5 mg/kg) and postoperatively (buprenorphine 0.1 mg/kg, meloxicam 5 mg/kg). Eyes were covered with ophthalmic gel (Virbac Schweiz AG). Depending on the experiment, we either implanted a cranial window to perform two-photon imaging or used crystal skull (or clear skull preparation in case of GRAB-NE1m imaging) for widefield imaging. Cranial windows were implanted over V1 and ACC as previously described (Keller et al., 2012; Leinweber et al., 2014). Briefly, using a dental drill, a 4 mm craniotomy was made over the right V1, centered 2.5 mm lateral and 0.5 mm anterior to lambda. A second craniotomy was made over right ACC, centered at midline, 0.5 mm anterior to bregma. After injection of an AAV vector carrying the reporter (see **Table S2** for details of virus),

the exposed cortex was sealed with a 3 mm or 4 mm circular glass coverslip and glued in place using gel superglue (Ultra Gel, Pattex). The remaining exposed surface of the skull was covered with Histoacryl (B. Braun), and a titanium head bar was fixed to the skull using dental cement (Paladur, Heraeus Kulzer). For widefield experiments, we injected an AAV vector with PHP.eB capsid retro-orbitally (6 µl per eye of at least 10^{13} GC/ml) to drive expression throughout cortex, or imaged GFP expressed in a fosGFP mice, or imaged without any fluorophore. For crystal skull surgery, we surgically removed the skull plate overlying the dorsal cortex and superglued a crystal skull coverslip over the craniotomy. Prior to removing the skull plate overlying dorsal cortex, the location of bregma relative to other landmarks on the skull was recorded. For clear skull preparation, the skull was cleared with a three-component polymer (C&B Metabond, Parkell), and the crystal skull coverslip was directly attached to the skull. An epifluorescence image was taken to mark reference points on cortical surface. As with 4mm cranial window implantation, a titanium head bar was fixed to the skull using dental cement. After surgery, anesthesia was antagonized by a mixture of flumazenil (0.5 mg/kg; Anexate, Roche) and atipamezole (2.5 mg/kg; Antisedan, Orion Pharma) injected intraperitoneally. Imaging commenced earliest 2 weeks after head bar implantation or 3 weeks after retro-orbital AAV injection.

Virtual reality environment

The virtual reality setup is based on the design of Dombeck and colleagues (Dombeck et al., 2007). Briefly, mice were head-fixed and free to run on an air-supported spherical treadmill. The rotation of the ball was restricted around the vertical axis with a pin. The virtual reality environment was projected onto a toroidal screen covering approximately 240 degrees horizontally and 100 degrees vertically of the mouse's visual field, using a projector (Samsung SP-F10M) synchronized to the resonant scanner of the two-photon microscope such that visual stimulation occurs at the turnaround time of the resonant scanner, during which any photon reaching the PMT is discarded, thereby allowing for near-simultaneous stimulation-artifact-free imaging. The virtual environment consisted of an infinite corridor with walls patterned with vertical sinusoidal gratings with a spatial frequency of approximately 0.04 cycles per degree (Leinweber et al., 2014). In closed loop sessions, the locomotion of the mouse was coupled to movement along a virtual tunnel. In open loop sessions, we uncoupled the two and replayed the visual flow from a preceding closed loop session. In grating sessions, we presented full-field drifting gratings (0°, 45°, 90°, 270°, moving in either direction) in a pseudo-random sequence. Grating stimuli were presented for between 1 s and 8 s depending on the experiment. In the inter-stimulus interval, mice were shown a gray screen with average luminance matched to that of the grating stimuli. For optogenetic light stimulation, we used 1 s pulses (10 mW after the objective) of a red laser (637 nm) directed at visual cortex through a cranial window.

Two-photon imaging

Two-photon calcium imaging was performed using custom-built microscopes (Leinweber et al. 2014). The illumination source was a tunable femtosecond laser (Insight, Spectra Physics or Chameleon, Coherent) tuned to 930 nm. Emission light was band-pass filtered using a 525/50 filter for GCaMP and a 607/70 filter for tdTomato/mCherry (Semrock) and detected using a GaAsP photomultiplier (H7422, Hamamatsu). Photomultiplier signals were amplified (DHPCA-100, Femto), digitized (NI5772, National Instruments) at 800 MHz, and band-pass filtered at 80 MHz using a digital Fourier-transform filter implemented in custom-written software on an FPGA (NI5772, National Instruments). The scanning system of the microscopes was based on a 12 kHz resonant scanner (Cambridge Technology). Images were acquired at a resolution of 750 x 400 pixels (60 Hz frame rate), and a piezo-electric linear actuator (P-726, Physik Instrumente) was used to move the objective (Nikon 16x, 0.8 NA) in steps of 15 µm between frames to acquire images at 4 different depths. This resulted in an effective frame rate of 15 Hz. We imaged layer 2/3 at a depth of 100-250 µm, and layer 5 at a depth of 400-600 µm. The field of view was 375 µm x 300 µm.

Widefield imaging

Widefield imaging experiments were conducted on a custom-build macroscope with commercially available objectives mounted face-to-face (Nikon 85 mm/f1.8 sample side, Nikon 50 mm/f1.4 sensor side). A 470 nm LED (Thorlabs) powered by a custom-build LED driver was used to excite GFP through an excitation filter (SP490, Thorlabs) reflected off a dichroic mirror (LP490, Thorlabs) placed parfocal to the objectives. The fluorescence was collected through a 525/50 emission filter on a sCMOS camera (PCO edge 4.2). LED illumination was adjusted with a collimator (Thorlabs SM2F32-A) to homogenously illuminate cortical surface through the cranial window. An Arduino board (Arduino Mega 2560) was used to synchronize LED onset with frame trigger signal of the camera. The duty cycle of the 470 nm LED was 90%. Images were acquired at 50 Hz or 100 Hz effective frame rate. Raw images were cropped on the sensor and data was stored to disk with custom-written LabVIEW (National Instruments) software, resulting in an effective pixel size of 60 μm^2 at a resolution of 1108 pixels x 1220 pixels (1.35 MP).

Extraction of neuronal and GRAB activity

Two-photon calcium imaging data were processed as previously described (Keller et al., 2012) and all data analysis was done in MATLAB (MathWorks). Briefly, raw images were full-frame registered to correct for lateral brain motion. All data with visible z-motion were excluded. Neurons were manually selected based on mean and maximum fluorescence images. For GRAB imaging, neuropil and blood vessels were manually selected. Raw fluorescence traces were corrected for slow drift in fluorescence using an 8^{th}-percentile filtering with a 66 s (or 1000 frames) window (Dombeck et al., 2007). $\Delta F/F_0$ traces were calculated as mean fluorescence in a selected region of every imaging frame, subtracted and normalized by the overall median fluorescence. No neuropil subtraction was done. All two-photon calcium and GRAB imaging data was acquired at 15 Hz.

For widefield macroscope imaging, raw movie data was manually registered across days by aligning subsequent mean projections of the data to the first recorded image sequence. ROIs were placed relative to readily identifiable anatomical landmarks that had been previously noted during cranial window surgery (see above). This resulted in the selection of six 20 pixels x 20 pixels ROIs per hemisphere. Of those ROIs, we calculated the activity as the $\Delta F/F_0$, wherein F_0 was the median fluorescence of the recording (approximately 30 000 frames in a 5 min recording). $\Delta F/F_0$ was corrected for slow fluorescence drift caused by thermal brightening of the LED using 8^{th} percentile filtering with a 62.5 s moving window similar to what's described above for two-photon imaging.

Extraction of blood vessel area

To measure the area of blood vessels, we manually identified rectangular ROIs encompassing the blood vessels and a surrounding margin of neuropil, as depicted in **Figures 4A** and **4B**. We generated a trigger-averaged video from these ROIs by averaging the video across all instances of optogenetic stimulation under dark condition. We applied a threshold set at the 0.5^{th} percentile of all pixel intensities in the video to binarize each frame. Subsequently, frame-wise we fitted an ellipse to the connected pixels that fell below this threshold, isolating the blood vessel. The area of this ellipse served as an estimate of the blood vessel area.

Data analysis

All data analysis was done using custom scripts written in MATLAB (MathWorks). To quantify the average population response traces, we first calculated the average event-triggered fluorescence trace for each ROI. The responses of all ROIs were then averaged and baseline-subtracted.

Locomotion onset was defined as running speed crossing a threshold of 0.25 cm/s for at least 1 s, while having been below the threshold during the preceding 1 s. The same criteria were used to define visual flow onsets in the open loop condition using visual flow speed. Visuomotor mismatch responses were probed by presenting brief 1 s full-field visual flow halts in the closed loop condition. For a visuomotor mismatch event to be included in analysis, mice had to be locomoting uninterrupted above threshold (0.25 cm/s) from −0.5 s to +1 s after the event onset. Additionally, for a ROI to be included for analysis of the response to a particular event, it had to have at least 3 onsets to the event. The GFP response was baseline subtracted using a −0.5 s to 0 s window relative to onset. The same was −1 s to −0.5 s in case of locomotion onset to account for preparatory activity.

In **Figures 2C-2E**, to compare GFP and GCaMP responses, we trial averaged the response of individual neurons and used the mean response over a time-window of 0 s to +2.0 s relative to locomotion or visuomotor mismatch onset, with a baseline subtraction window of −1 s to −0.5 s. For grating responses, we used the mean of the trial averaged response over +0.5 s to +3.0 s relative to stimulus onset, with a baseline subtraction window of −0.5 s to 0 s.

In **Figure 3**, for each neuron, we averaged the response from +0.5 s to +1.5 s relative to onset for individual locomotion onset and visuomotor mismatch trials and tested if the mean of this distribution is significantly different from 0 using a t-test with a chance threshold of 5%. In a similar analysis, the response window for grating responsive neurons was set at +0.5 s to +2.5 s to account for the 2 s long grating stimuli.

In **Figure 4**, the blood vessel diameter and the average GFP signal from both blood vessel ROIs and neuropil ROIs were smoothed with a 0.5 s moving average and the variance explained was computed on the −2.0 s to +6.0 s window relative to stimulus onset with a linear model (i.e. R^2, where R is the linear correlation coefficient between the two signals). In **Figure 4D**, data points were drawn as outliers if they were larger than 75th percentile + 1.5*(75th percentile - 25th percentile) or smaller than 25th percentile - 1.5*(75th percentile - 25th percentile) of the distribution and were omitted from the figure.

In **Figures 7** and **S8**, due to insufficient triggers, we relaxed the locomotion threshold for mismatch to 0.12 cm/s, and in **Figure S8**, we relaxed the sitting window before locomotion onset to −0.5 s to 0 s.

Statistical Analysis

All statistical information for the tests performed in the manuscript is provided in **Table S1**. Unless stated otherwise, the shading indicates the error in mean estimation to within 1 standard deviation. For analysis where the experimental unit was neurons (or ROIs), we used hierarchical bootstrap (Saravanan et al., 2020) for statistical testing due to the nested nature (neurons (or ROIs) and mice) of the data. Briefly, we trial averaged the data for the respective event per neuron (or ROI), and first resampled the data (with replacement) at the level of imaging sites, and then, from the selected sites, resampled for neurons. We then computed the mean of this bootstrap sample and repeated this N times to generate a bootstrap distribution of the mean estimate. For all statistical testing the number of bootstrap samples (N) was 10 000, for plotting bootstrap mean and standard error response curves it was 1000. The bootstrap standard error is the 68% confidence interval (1 SD, 15.8th percentile to 84.2nd percentile) in the bootstrap distribution of the mean.

Supplementary information

Figure panel	Value compared	Test	P-value	N (ROIs)	N (Sites)	N (Mice)
1A	Locomotion onset response in V1 L2/3	-	-	2647	6	6
1B	Grating onset response in V1 L2/3	-	-	2647	6	6
1C	Visual response to opto stim light in V1 L2/3	-	-	2647	6	6
1D	Visuomotor mismatch response in V1 L2/3	-	-	2218	5	5
1E	Locomotion onset response in V1 L5	-	-	1463	6	6
1F	Grating onset response in V1 L5	-	-	1463	6	6
1G	Visual response to opto stim light in V1 L5	-	-	1463	6	6
1H	Visuomotor mismatch response in V1 L5	-	-	947	4	4
1I	Locomotion onset response in ACC L2/3	-	-	2417	6	6
1J	Grating onset response in ACC L2/3	-	-	2417	6	6
1K	Visual response to opto stim light in ACC L2/3	-	-	2417	6	6
1L	Visuomotor mismatch response in ACC L2/3	-	-	2417	6	6
2B-2E	GCaMP6f *vs* GFP	-	-	15 103, 6527	51, 18	15, 6
3A	Fraction of L2/3 neurons in V1 classified responsive to locomotion onsets based on GFP signals	Hierarchical bootstrap	$<10^{-4}$	2647	6	6
	Fraction of L2/3 neurons in V1 classified responsive to grating onsets based on GFP signals	Hierarchical bootstrap	$<10^{-4}$	2647	6	6
	Fraction of L2/3 neurons in V1 classified responsive to mismatch onsets based on GFP signals	Hierarchical bootstrap	0.34	2218	5	5
	Fraction of L2/3 neurons in V1 classified responsive to locomotion onsets *vs* grating onset based on GFP signals	Hierarchical bootstrap	$<10^{-4}$	2647, 2647	6,6	6,6
	Fraction of L2/3 neurons in V1 classified responsive to grating onsets *vs* mismatch onset based on GFP signals	Hierarchical bootstrap	$<10^{-4}$	2647, 2218	6,5	6,5
	Fraction of L2/3 neurons in V1 classified responsive to locomotion onsets *vs* mismatch onset based on GFP signals	Hierarchical bootstrap	$<10^{-4}$	2647, 2218	6,5	6,5
3B	Fraction of L5 neurons in V1 classified responsive to locomotion onsets based on GFP signals	Hierarchical bootstrap	$<10^{-4}$	1463	6	6
	Fraction of L5 neurons in V1 classified responsive to grating onsets based on GFP signals	Hierarchical bootstrap	$<10^{-4}$	1463	6	6
	Fraction of L5 neurons in V1 classified responsive to mismatch onsets based on GFP signals	Hierarchical bootstrap	0.09	947	4	4

Table S1.

Statistical information on all analysis.

	Fraction of L5 neurons in V1 classified responsive to locomotion onsets *vs* grating onset based on GFP signals	Hierarchical bootstrap	0.08	1463, 1463	6,6	6,6
	Fraction of L5 neurons in V1 classified responsive to grating onsets *vs* mismatch onset based on GFP signals	Hierarchical bootstrap	$<10^{-4}$	1463, 947	6,4	6,4
	Fraction of L5 neurons in V1 classified responsive to locomotion onsets *vs* mismatch onset based on GFP signals	Hierarchical bootstrap	$<10^{-4}$	1463, 947	6,4	6,4
3C	Fraction of L2/3 neurons in ACC classified responsive to locomotion onsets based on GFP signals	Hierarchical bootstrap	$<10^{-4}$	2417	6	6
	Fraction of L2/3 neurons in ACC classified responsive to grating onsets based on GFP signals	Hierarchical bootstrap	0.18	2417	6	6
	Fraction of L2/3 neurons in ACC classified responsive to mismatch onsets based on GFP signals	Hierarchical bootstrap	0.34	2417	6	6
	Fraction of L2/3 neurons in ACC classified responsive to locomotion onsets *vs* grating onset based on GFP signals	Hierarchical bootstrap	$<10^{-4}$	2417, 2417	6,6	6,6
	Fraction of L2/3 neurons in ACC classified responsive to grating onsets *vs* mismatch onset based on GFP signals	Hierarchical bootstrap	0.57	2417, 2417	6,6	6,6
	Fraction of L2/3 neurons in ACC classified responsive to locomotion onsets *vs* mismatch onset based on GFP signals	Hierarchical bootstrap	$<10^{-4}$	2417, 2417	6,6	6,6
3D	Fraction of L2/3 neurons in V1 classified responsive to locomotion onsets based on GCaMP6f signals	Hierarchical bootstrap	$<10^{-4}$	7971	25	15
	Fraction of L2/3 neurons in V1 classified responsive to grating onsets based on GCaMP6f signals	Hierarchical bootstrap	$<10^{-4}$	7971	25	15
	Fraction of L2/3 neurons in V1 classified responsive to mismatch onsets based on GCaMP6f signals	Hierarchical bootstrap	$<10^{-4}$	7653	24	15
	Fraction of L2/3 neurons in V1 classified responsive to locomotion onsets *vs* grating onset based on GCaMP6f signals	Hierarchical bootstrap	0.01	7971, 7971	25, 25	15, 15
	Fraction of L2/3 neurons in V1 classified responsive to grating onsets *vs* mismatch onset based on GCaMP6f signals	Hierarchical bootstrap	$<10^{-4}$	7971, 7653	25, 24	15, 15
	Fraction of L2/3 neurons in V1 classified responsive to locomotion onsets *vs* mismatch onset based on GCaMP6f signals	Hierarchical bootstrap	$<10^{-4}$	7971, 7653	25, 24	15, 15
3E	Fraction of L5 neurons in V1 classified responsive to locomotion onsets based on GCaMP6f signals	Hierarchical bootstrap	$<10^{-4}$	7132	26	14
	Fraction of L5 neurons in V1 classified responsive to grating onsets based on GCaMP6f signals	Hierarchical bootstrap	$<10^{-4}$	7132	26	14
	Fraction of L5 neurons in V1 classified responsive to mismatch onsets based on GCaMP6f signals	Hierarchical bootstrap	$<10^{-3}$	6812	25	13

Table S1. (continued)

	Fraction of L5 neurons in V1 classified responsive to locomotion onsets *vs* grating onset based on GCaMP6f signals	Hierarchical bootstrap	0.19	7132, 7132	26, 26	14, 14
	Fraction of L5 neurons in V1 classified responsive to grating onsets *vs* mismatch onset based on GCaMP6f signals	Hierarchical bootstrap	$<10^{-4}$	7132, 6812	26, 25	14, 13
	Fraction of L5 neurons in V1 classified responsive to locomotion onsets *vs* mismatch onset based on GCaMP6f signals	Hierarchical bootstrap	$<10^{-4}$	7132, 6812	26, 25	14, 13
3F	Fraction of L2/3 neurons in ACC classified responsive to locomotion onsets based on GCaMP signals	Hierarchical bootstrap	$<10^{-4}$	2207	12	7
	Fraction of L2/3 neurons in ACC classified responsive to mismatch onsets based on GCaMP signals	Hierarchical bootstrap	$<10^{-4}$	2207	12	7
	Fraction of L2/3 neurons in ACC classified responsive to locomotion onsets *vs* mismatch onset based on GCaMP signals	Hierarchical bootstrap	$<10^{-4}$	2207, 2207	12,12	7,7
4D-4E	Variance explained in V1 L2/3 neuron GFP signal	-	-	2647	6	6
4D-4E	Variance explained in V1 L5 neuron GFP signal	-	-	1463	6	6
4D-4E	Variance explained in ACC L2/3 neuron GFP signal	-	-	2417	6	6
5A	Locomotion onset response under different conditions in V1 L2/3	-	-	2647	6	6
5B	Locomotion onset response under different conditions in V1 L5	-	-	1463	6	6
5C	Locomotion onset response under different conditions in ACC L2/3	-	-	2417	6	6
6A	Pairwise correlation in GFP signal in V1 L2/3 neurons, stationary *vs* locomoting	Hierarchical bootstrap	$<10^{-3}$	2647	6	6
6B	Pairwise correlation in GFP signal in V1 L5 neurons, stationary *vs* locomoting	Hierarchical bootstrap	$<10^{-4}$	1276	6	6
6C	Pairwise correlation in GFP signal in ACC L2/3 neurons, stationary *vs* locomoting	Hierarchical bootstrap	$<10^{-4}$	2417	6	6
7A	Locomotion onset response in V1	-	-	48	24	11
7B	Grating onset response in V1	-	-	52	26	12
7C	Visuomotor mismatch response in V1	-	-	48	24	11
7D	Locomotion onset response in ACC	-	-	48	24	11
7E	Grating onset response in ACC	-	-	52	26	12
7F	Visuomotor mismatch response in ACC	-	-	48	24	11
8B	Locomotion onset response in GRAB-DA1m in V1 – neuropil and blood vessel ROI response	-	-	168, 34	7, 5	7, 5
8C	Grating onset response in GRAB-DA1m in V1 – neuropil and blood vessel ROI response	-	-	168, 34	7, 5	7, 5

Table S1. (continued)

8D	Visuomotor mismatch response in GRAB-DA1m in V1 – neuropil and blood vessel ROI response	-	-	124, 20	5, 3	5, 3
8F	Locomotion onset response in GRAB-5HT1.0 in V1 – neuropil and blood vessel ROI response	-	-	220, 113	10, 10	9, 9
8G	Grating onset response in GRAB-5HT1.0 in V1 – neuropil and blood vessel ROI response	-	-	205, 101	9, 9	9, 9
8H	Visuomotor mismatch response in GRAB-5HT1.0 in V1 – neuropil and blood vessel ROI response	-	-	193, 101	9, 9	8, 8
8J	Locomotion onset response in GRAB-ACh3.0 in V1 – neuropil and blood vessel ROI response	-	-	178, 118	8,8	8,8
8K	Grating onset response in GRAB-ACh3.0 in V1 – neuropil and blood vessel ROI response	-	-	178, 118	8,8	8,8
8L	Visuomotor mismatch response in GRAB-ACh3.0 in V1 – neuropil and blood vessel ROI response	-	-	178, 118	8,8	8,8
9A	Locomotion onset response in GRAB-NE1m in V1	-	-	64	32	5
9B	Grating onset response in GRAB-NE1m in V1	-	-	12	6	5
9C	Visuomotor mismatch response in GRAB-NE1m in V1	-	-	18	9	5
9D	Locomotion onset response in GRAB-NE1m in ACC	-	-	64	32	5
9E	Grating onset response in GRAB-NE1m in ACC	-	-	12	6	5
9F	Visuomotor mismatch response in GRAB-NE1m in ACC	-	-	18	9	5
S2A	Locomotion onset response in V1 L2/3	-	-	2647	6	6
	Grating onset response in V1 L2/3	-	-	2647	6	6
	Visual response to opto stim light in dark in V1 L2/3	-	-	2647	6	6
	Visual response to opto stim light in VR in V1 L2/3	-	-	2647	6	6
	Visuomotor mismatch response in V1 L2/3	-	-	2218	5	5
S2B	Locomotion onset response in V1 L5	-	-	1463	6	6
	Grating onset response in V1 L5	-	-	1463	6	6
	Visual response to opto stim light in dark in V1 L5	-	-	1463	6	6
	Visual response to opto stim light in VR in V1 L5	-	-	1463	6	6
	Visuomotor mismatch response in V1 L5	-	-	947	4	4
S2C	Locomotion onset response in ACC L2/3	-	-	2417	6	6
	Grating onset response in ACC L2/3	-	-	2417	6	6
	Visual response to opto stim light in dark in ACC L2/3	-	-	2417	6	6
	Visual response to opto stim light in VR in ACC L2/3	-	-	2417	6	6
	Visuomotor mismatch response in ACC L2/3	-	-	2417	6	6
S3A	Visual flow onset response with GFP in V1 L2/3	-	-	2647	6	6
S3B	Locomotion onset response in dark with GFP in V1 L2/3	-	-	2647	6	6
S3C	Visual flow onset response in closed loop with GFP in V1 L2/3	-	-	2647	6	6

Table S1. (continued)

S4	Average peak GFP response for low vs intermediate GFP expression level	Hierarchical bootstrap	0.88	7076, 6477	17, 15	6, 6
	Average peak GFP response for intermediate vs high GFP expression level	Hierarchical bootstrap	0.36	6477, 4857	15, 10	6, 6
	Average peak GFP response for low vs high GFP expression level	Hierarchical bootstrap	0.85	7076, 4857	17, 10	6, 6
S5A	Brain velocity on locomotion onset	-	-	-	18	6
S5B	Brain velocity on grating onset	-	-	-	18	6
S5C	Brain velocity on opto stim light onset in dark	-	-	-	18	6
S5D	Brain velocity on opto stim light onset in VR	-	-	-	18	6
S8A	Locomotion onset response with widefield and no fluorophore in V1	-	-	24	12	6
S8B	Grating onset response with widefield and no fluorophore in V1	-	-	24	12	6
S8C	Visuomotor mismatch response with widefield and no fluorophore in V1	-	-	16	8	4
S8D	Locomotion onset response with widefield and no fluorophore in ACC	-	-	24	12	6
S8E	Grating onset response with widefield and no fluorophore in ACC	-	-	24	12	6
S8F	Visuomotor mismatch response with widefield and no fluorophore in ACC	-	-	16	8	4
S9A	Locomotion onset response in GRAB-DA1m in V1 – neuropil and blood vessel ROI response	-	-	122, 61	7, 7	7, 7
S9B	Grating onset response in GRAB- DA1m in V1 – neuropil and blood vessel ROI response	-	-	122, 61	7, 7	7, 7
S9C	Visuomotor mismatch response in GRAB- DA1m in V1 – neuropil and blood vessel ROI response	-	-	80, 50	5, 5	5, 5
S9D	Locomotion onset response in GRAB-5HT1.0 in ACC – neuropil and blood vessel ROI response	-	-	123, 79	7, 7	6, 6
S9E	Grating onset response in GRAB-5HT1.0 in ACC – neuropil and blood vessel ROI response	-	-	109, 69	6, 6	6, 6
S9F	Visuomotor mismatch response in GRAB-5HT1.0 in ACC – neuropil and blood vessel ROI response	-	-	91, 56	5, 5	5, 5

Table S1. (continued)

Genotype	Virus	Site of injection	Nbr. Mice	Nbr. ROIs	Figures
C57BL/6	AAV2/1-Ef1α-eGFP-WPRE	Primary visual cortex – L2/3	6	3072	Figures 1, 2, 3A, 4, 5, 6, S1, S2, S3, S4, S5, S7
C57BL/6	AAV2/1-Ef1α-eGFP-WPRE	Primary visual cortex – L5	6	1710	Figures 1, 2, 3B, 4, 5, 6, S1, S2, S3, S4, S5, S6, S7
C57BL/6	AAV2/1-Ef1α-eGFP-WPRE	Anterior cingulate cortex – L2/3	6	2739	Figures 1, 2, 3C, 4, 5, 6, S1, S2, S3, S4, S5, S6, S7
ChAT-IRES-Cre	AAV2/1-Ef1α-GCaMP6f-WPRE	Primary visual cortex – L2/3 and L5	15	15 103	Figures 2, 3D, 3E
C57BL/6	AAV-PHP.eB-EF1α-eGFP	Retro-orbital	8	32	Figure 7
fosGFP	-	-	4	20	Figure 7
Tlx3-Cre	-	-	6	24	Figure S8
ChAT-IRES-Cre	AAV2/9-hSyn-GRAB-ACh3.0	Primary visual cortex	8	609	Figure 8
C57BL/6	AAV2/9-hSyn-GRAB-DA1m	Anterior cingulate cortex	7	183	Figures 8, S9
		Primary visual cortex	7	202	
C57BL/6	AAV2/9-hSyn-GRAB-5HT1.0	Anterior cingulate cortex	9	202	Figures 8, S9
		Primary visual cortex	6	333	
C57BL/6	AAV-PHP.eB-hSyn-GRAB-NE1m	Retro-orbital	5	156	Figure 9

Table S2.

Number of mice in different experimental groups.

Supplementary figures

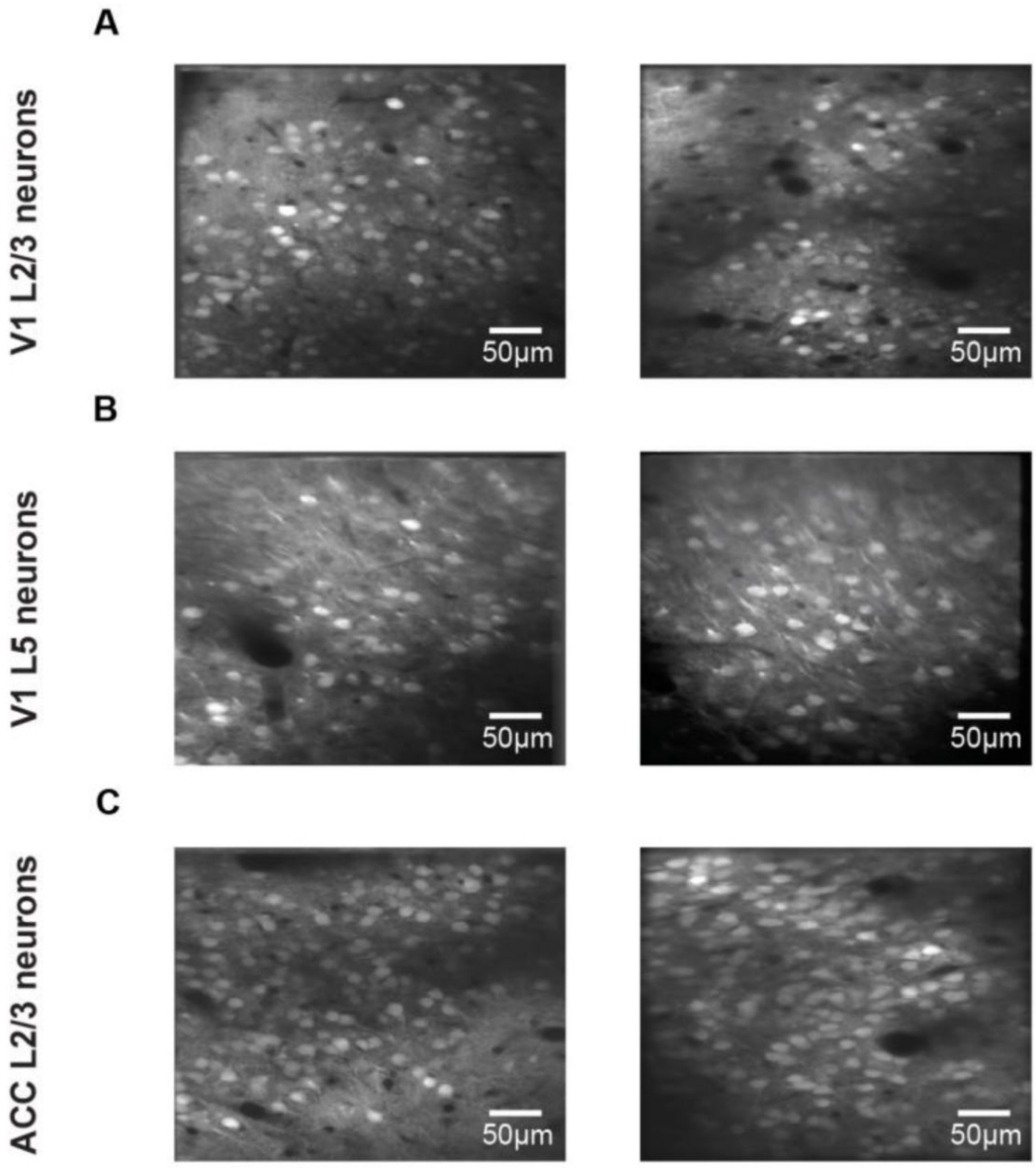


Figure S1.

Example imaging sites.

(**A**) Example of two imaging sites in L2/3 of V1.

(**B**) Example of two imaging sites in L5 of V1.

(**C**) Example of two imaging sites in L2/3 of ACC.

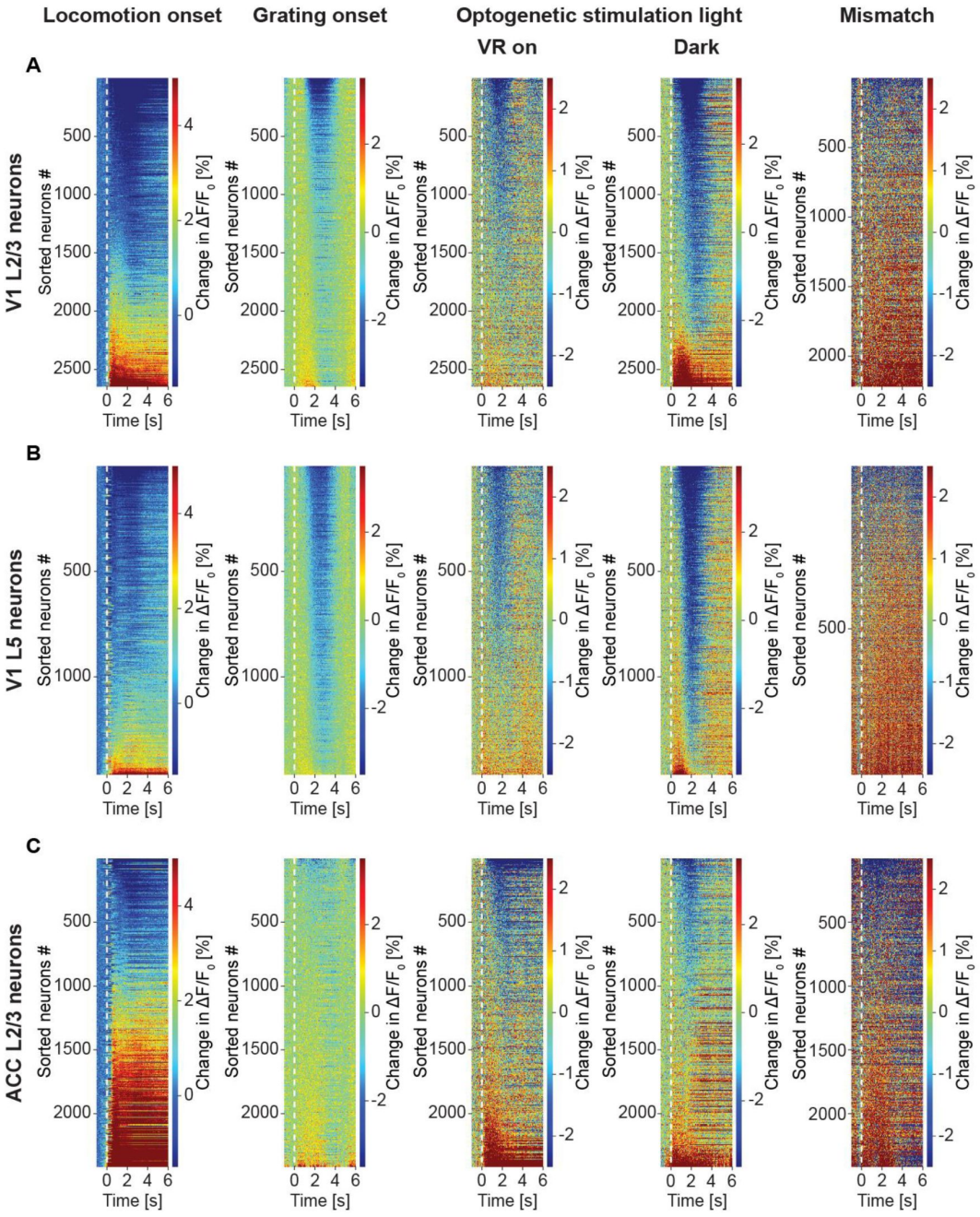


Figure S2.

GFP response heatmaps. Related to Figure 1.

(**A**) Trial-averaged change in apparent GFP fluorescence in L2/3 neurons in V1 to 5 different stimuli: locomotion onset, grating onset, optogenetic light stimulation with the VR on and in dark, and visuomotor mismatch. Note, the colormap is different for different panels and not always symmetric around 0 to illustrate all changes.
(**B**) As in **A**, but for L5 neurons in V1.
(**C**) As in **A**, but for L2/3 neurons in ACC.

Figure S3.

GFP response in V1 L2/3 on locomotion onset in closed loop was not a linear sum of component responses. Related to Figure 1.

(**A**) Average change in GFP signal from L2/3 in V1 in response to visual flow onset in open loop during stationary periods. Mean (solid lines) and the bootstrap standard error (shading) are calculated as hierarchical bootstrap estimate for each time bin.
(**B**) As in **A**, but locomotion onset in dark.
(**C**) As in **A**, but for locomotion onset in closed loop.

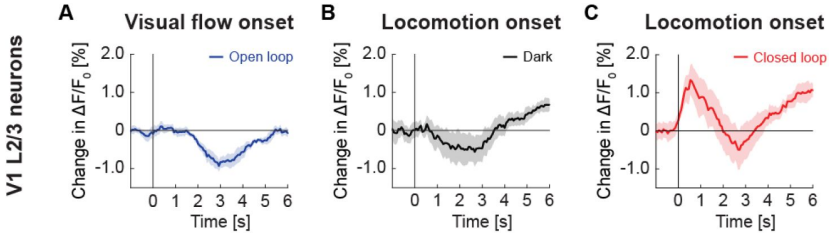


Figure S4.

Peak GFP response of neurons cannot be explained by GFP expression levels. Related to Figure 3.

The peak $\Delta F/F_0$ response of each neuron is plotted against the mean GFP fluorescence intensity. Each dot is a neuron. Superimposed are the hierarchical bootstrap estimates of the median (dark red line) and 95% confidence intervals (light red lines) for each GFP fluorescence bin. We found no evidence of a difference in peak $\Delta F/F_0$ response between low, intermediate, and high GFP fluorescence expression levels (indicated by green shading) in our dataset. This is as one would expect, assuming the baseline fluorescence F_0 is above the noise floor.

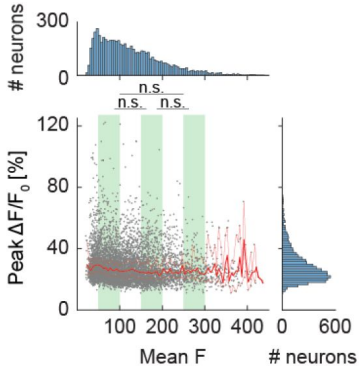


Figure S5.

Brain motion estimate. Related to Figure 4.

(**A**) Estimate of average absolute brain velocity in pixels per frame at locomotion onset across all imaging sites. This estimate is based on the pixel-wise frame shift correction in the x and y directions. Note, these shifts are corrected in the data. Mean (solid lines) and the bootstrap standard error (shading) are calculated as hierarchical bootstrap estimate for each time bin. Euclidean displacement was calculated as the room-mean-squared x- and y-displacement of the imaging site.
(**B**) As in **A**, but for grating onsets. Green shading marks the duration of the gratings. There is no appreciable change in brain velocity on visual stimulation.
(**C**) As in **A**, but for optogenetic stimulation light presentation in dark. Pink shading marks the duration of the light stimulus.
(**D**) As in **A**, but for optogenetic stimulation light presentation in VR. Pink shading marks the duration of the light stimulus.

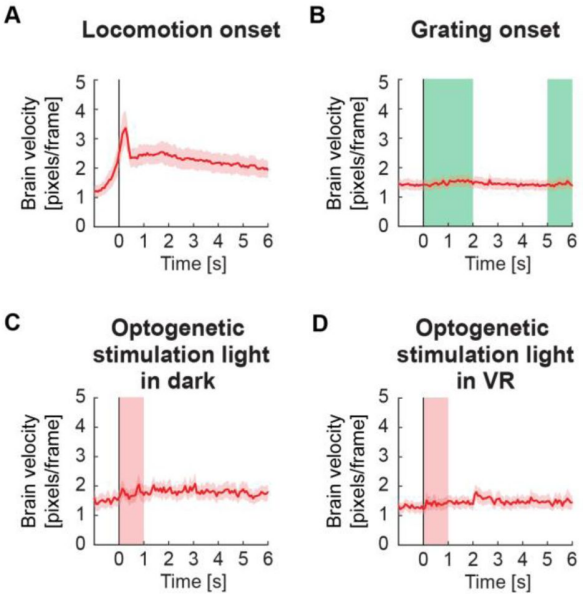
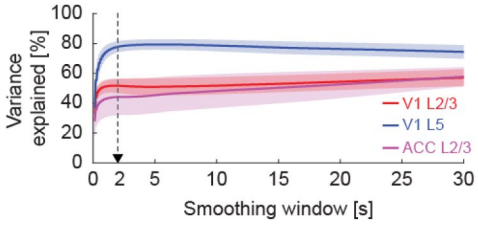


Figure S6.

The variance of neuronal GFP signal explained by blood vessel GFP signal as a function of smoothing window. Related to Figure 4E.

Black arrow indicates the smoothing window used in **Figure 4E**.



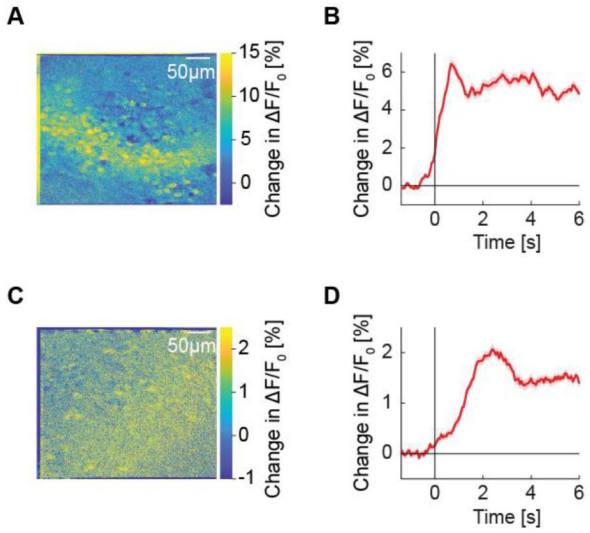


Figure S7.

Average locomotion onset response during two-photon GFP imaging. Related to Figure 4.

(**A**) Putative blood vessel shadow cast on an example imaging plane on locomotion onset. Shown is the baseline subtracted average pixelwise map of GFP response on locomotion onset.
(**B**) Average GFP response of all neurons from the imaging site in **A**.
(**C**) Example of a typical response on locomotion onset. Here shadowing is likely caused by blood vessels closer to the surface of the brain.
(**D**) Average GFP response of all neurons from the imaging site in **C**.

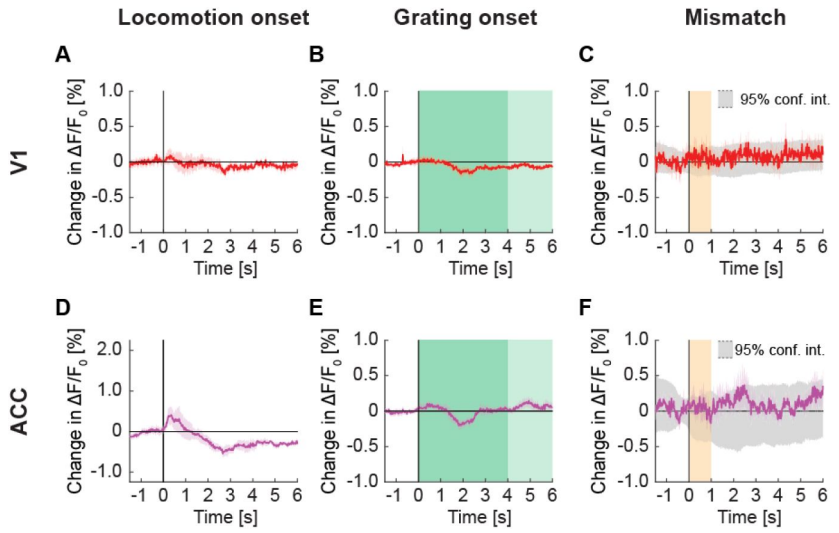


Figure S8.

Signals in widefield fluorescence imaging of V1 and ACC in mice not expressing a fluorophore. Related to Figure 7.

(**A**) Average change in signal from V1 on locomotion onset. Mean (solid lines) and the bootstrap standard error (shading) are calculated as hierarchical bootstrap estimate for each time bin, in mice not expressing a fluorophore.
(**B**) As in **A**, but on grating onset in V1. Grating duration were randomly selected between 4 s and 8 s (green shading).
(**C**) As in **A**, but on visuomotor mismatch in V1. Orange shading indicates the duration of visuomotor mismatch.
(**D**) As in **A**, but for change in signal from ACC on locomotion onset.
(**E**) As in **B**, but for change in signal from ACC on grating onset.
(**F**) As in **C**, but for change in signal from ACC on visuomotor mismatch.

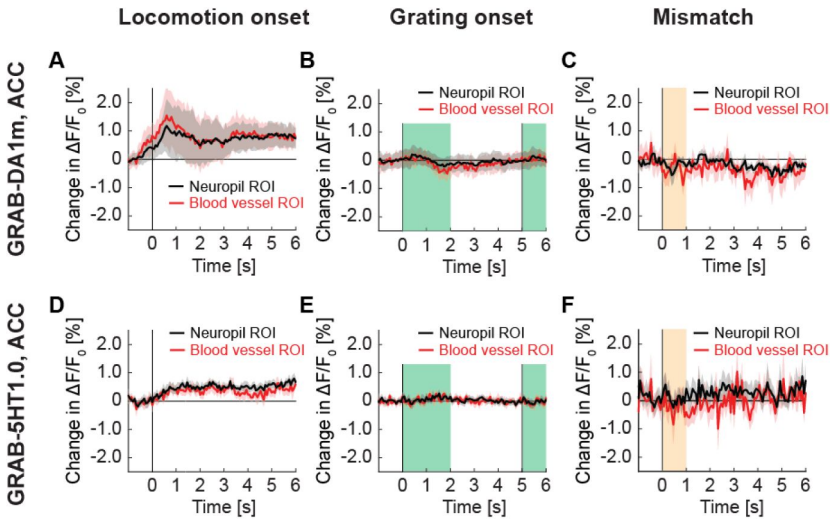


Figure S9.

GRAB signals in ACC recorded with two-photon imaging. Related to Figure 8.

(**A**) Average change in GRAB-DA1m signal in ACC on locomotion onset in ROIs over neuropil (in black) and in ROIs over blood vessels (in red). Mean (solid lines) and the bootstrap standard error (shading) are calculated as hierarchical bootstrap estimate for each time bin.
(**B**) As in **A**, but for response to full-field grating onset. Green shading marks the duration of the gratings.
(**C**) As in **A**, but for response to visuomotor mismatch.
(**E**) Average change in GRAB-5HT1.0 signal in ACC on locomotion onset in ROIs over neuropil (in black) and in ROIs over blood vessels (in red).
(**F**) As in **G**, but for response to grating onset.
(**G**) As in **G**, but for response to visuomotor mismatch.

Acknowledgements

We thank all the members of the Keller lab for discussion and support. This project has received funding from the Swiss National Science Foundation (GBK), the Novartis Research Foundation (GBK), and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 865617) (GBK).

Additional information

Author contributions

BY designed the experiments. BY performed the two-photon imaging experiments, RJ performed the widefield imaging of GRAB-NE1m experiments, and RJ and MH performed the widefield imaging of GFP and with no fluorophore experiments. BY analyzed the data. BY and GK wrote the manuscript.

References

Allen W.E., Kauvar I.V., Chen M.Z., Richman E.B., Yang S.J., Chan K., Gradinaru V., Deverman B.E., Luo L., Deisseroth K (2017) **Global Representations of Goal-Directed Behavior in Distinct Cell Types of Mouse Neocortex** *Neuron* **94**:891–907 https://doi.org/10.1016/j.neuron.2017.04.017

Attinger A., Wang B., Keller G.B (2017) **Visuomotor Coupling Shapes the Functional Development of Mouse Visual Cortex** *Cell* **169**:1291–1302 https://doi.org/10.1016/j.cell.2017.05.023

Attwell D., Buchan A.M., Charpak S., Lauritzen M., MacVicar B.A., Newman E.A (2010) **Glial and neuronal control of brain blood flow** *Nature* **468**:232–243 https://doi.org/10.1038/nature09613

Aydın Ç., Couto J., Giugliano M., Farrow K., Bonin V (2018) **Locomotion modulates specific functional cell types in the mouse visual thalamus** *Nat. Commun* **9**:4882 https://doi.org/10.1038/s41467-018-06780-3

Berger B., Gaspar P., Verney C (1991) **Dopaminergic innervation of the cerebral cortex: unexpected differences between rodents and primates** *Trends Neurosci* **14**:21–27 https://doi.org/10.1016/0166-2236(91)90179-x

Chen T.-W., Wardill T.J., Sun Y., Pulver S.R., Renninger S.L., Baohan A., Schreiter E.R., Kerr R.A., Orger M.B., Jayaraman V., Looger L.L., Svoboda K., Kim D.S (2013) **Ultrasensitive fluorescent proteins for imaging neuronal activity** *Nature* **499**:295–300 https://doi.org/10.1038/nature12354

Chesler M (2003) **Regulation and modulation of pH in the brain** *Physiol. Rev* **83**:1183–1221 https://doi.org/10.1152/physrev.00010.2003

Couto J., Musall S., Sun X.R., Khanal A., Gluf S., Saxena S., Kinsella I., Abe T., Cunningham J.P., Paninski L., Churchland A.K (2021) **Chronic, cortex-wide imaging of specific cell populations during behavior** *Nat. Protoc* **16**:3241–3263 https://doi.org/10.1038/s41596-021-00527-z

Dadarlat M.C., Stryker M.P (2017) **Locomotion enhances neural encoding of visual stimuli in mouse V1** *J. Neurosci* **37**:3764–3775 https://doi.org/10.1523/JNEUROSCI.2728-16.2017

Dombeck D.A., Khabbaz A.N., Collman F., Adelman T.L., Tank D.W (2007) **Imaging Large-Scale Neural Activity with Cellular Resolution in Awake, Mobile Mice** *Neuron* **56**:43–57 https://doi.org/10.1016/j.neuron.2007.08.003

Dos Santos N.V., Saponi C.F., Ryan T.M., Primo F.L., Greaves T.L., Pereira J.F.B. (2020) **Reversible and irreversible fluorescence activity of the Enhanced Green Fluorescent Protein in pH: Insights for the development of pH-biosensors** *Int. J. Biol. Macromol* **164**:3474–3484 https://doi.org/10.1016/j.ijbiomac.2020.08.224

Erisken S., Vaiceliunaite A., Jurjut O., Fiorini M., Katzner S., Busse L (2014) **Effects of Locomotion Extend throughout the Mouse Early Visual System** *Curr. Biol* **24**:2899–2907 https://doi.org/10.1016/j.cub.2014.10.045

Feng J., Zhang C., Lischinsky J.E., Jing M., Zhou J., Wang H., Zhang Y., Dong A., Wu Z., Wu H., Chen W., Zhang P., Zou J., Hires S.A., Zhu J.J., Cui G., Lin D., Du J., Li Y (2019) **A Genetically Encoded Fluorescent Sensor for Rapid and Specific In Vivo Detection of Norepinephrine** *Neuron* **102**:745–761 https://doi.org/10.1016/j.neuron.2019.02.037

Hamada H.T., Abe Y., Takata N., Taira M., Tanaka K.F., Doya K (2023) **Optogenetic activation of dorsal raphe serotonin neurons induces brain-wide activation** *bioRxiv* https://doi.org/10.1101/2022.08.07.503074

Iadecola C., Nedergaard M (2007) **Glial regulation of the cerebral microvasculature** *Nat. Neurosci* **10**:1369–1376 https://doi.org/10.1038/nn2003

Jing M., Li Yuexuan, Zeng J., Huang P., Skirzewski M., Peng W., Qian T., Tan K., Wu R., Zhang S., Pan S., Xu M., Li H., Saksida L.M., Prado V.F., Bussey T., Prado M.A.M., Cheng H., Li Yulong (2019) **An optimized acetylcholine sensor for monitoring in vivo cholinergic activity** *Nature Methods* **17**:1139–1146

Keller G.B., Bonhoeffer T., Hübener M (2012) **Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse** *Neuron* **74**:809–815 https://doi.org/10.1016/j.neuron.2012.03.040

Kneen M., Farinas J., Li Y., Verkman A.S (1998) **Green fluorescent protein as a noninvasive intracellular pH indicator** *Biophys. J* **74**:1591–1599 https://doi.org/10.1016/S0006-3495(98)77870-1

Leinweber M., Ward D.R., Sobczak J.M., Attinger A., Keller G.B (2017) **A Sensorimotor Circuit in Mouse Cortex for Visual Flow Predictions** *Neuron* **95**:1420–1432 https://doi.org/10.1016/j.neuron.2017.08.036

Leinweber M., Zmarz P., Buchmann P., Argast P., Hübener M., Bonhoeffer T., Keller G.B (2014) **Two-photon Calcium Imaging in Mice Navigating a Virtual Reality Environment** *J. Vis. Exp* :1–6 https://doi.org/10.3791/50885

Lohani S., Moberly A.H., Benisty H., Landa B., Jing M., Li Y., Higley M.J., Cardin J.A (2022) **Spatiotemporally heterogeneous coordination of cholinergic and neocortical activity** *Nat. Neurosci* **25**:1706–1713 https://doi.org/10.1038/s41593-022-01202-6

Ma Y., Shaik M.A., Kim S.H., Kozberg M.G., Thibodeaux D.N., Zhao H.T., Yu H., Hillman E.M.C (2016) **Wide-field optical mapping of neural activity and brain haemodynamics: considerations and novel approaches** *Philos. Trans. R. Soc. B Biol. Sci* **371**:20150360 https://doi.org/10.1098/rstb.2015.0360

Mächler P., Fomin-Thunemann N., Thunemann M., Sætra M.J., Desjardins M., Kılıç K., Şencan I., Li B., Saisan P., Cheng Q., Weldy K.L., Boas D.A., Buxton R.B., Einevoll G.T., Dale A.M., Sakadžić S., Devor A. (2021) **Microscopic Quantification of Oxygen Consumption across Cortical Layers** *bioRxiv* https://doi.org/10.1101/2021.10.13.464176

Ocana-Santero G., Packer A.M., Sharp T., Butt S.J.B (2024) **In Vivo Two-Photon Microscopy Reveals Sensory-Evoked Serotonin (5-HT) Release in Adult Mammalian Neocortex** *ACS Chem. Neurosci* **15**:456–461 https://doi.org/10.1021/acschemneuro.3c00725

Raimondo J.V., Joyce B., Kay L., Schlagheck T., Newey S.E., Srinivas S., Akerman C.J (2013) **A genetically-encoded chloride and pH sensor for dissociating ion dynamics in the nervous system** *Front. Cell. Neurosci* **7**:202 https://doi.org/10.3389/fncel.2013.00202

Ruff C.F., Anaya F.J., Dienel S.J., Rakymzhan A., Altamirano-Espinoza A., Couey J.J., Fukuda M., Watson A.M., Su A., Fish K.N., Rubio M.E., Hooks B.M., Ross S.E., Vazquez A.L (2024) **Long-range inhibitory neurons mediate cortical neurovascular coupling** *Cell Rep* **43** https://doi.org/10.1016/j.celrep.2024.113970

Saravanan V., Berman G.J., Sober S.J. (2020) **Application of the hierarchical bootstrap to multi-level data in neuroscience** *Neurons Behav. Data Anal. Theory* **3** https://nbdt.scholasticahq.com/article/13927-application-of-the-hierarchical-bootstrap-to-multi-level-data-in-neuroscience

Scott B.B., Thiberge S.Y., Guo C., Tervo D.G.R., Brody C.D., Karpova A.Y., Tank D.W (2018) **Imaging cortical dynamics in GCaMP transgenic rats with a head-mounted widefield macroscope** *Neuron* **100**:1045–1058 https://doi.org/10.1016/j.neuron.2018.09.050

Shen Z., Lu Z., Chhatbar P.Y., O'Herron P., Kara P (2012) **An artery-specific fluorescent dye for studying neurovascular coupling** *Nat. Methods* **9**:273–276 https://doi.org/10.1038/nmeth.1857

Sun F., Zeng J., Jing M., Zhou J., Feng J., Owen S.F., Luo Y., Li F., Wang H., Yamaguchi T., Yong Z., Gao Y., Peng W., Wang L., Zhang S., Du J., Lin D., Xu M., Kreitzer A.C., Cui G., Li Y (2018) **A Genetically Encoded Fluorescent Sensor Enables Rapid and Specific Detection of Dopamine in Flies, Fish, and Mice** *Cell* **174**:481–496 https://doi.org/10.1016/j.cell.2018.06.042

Valley M.T., Moore M.G., Zhuang J., Mesa N., Castelli D., Sullivan D., Reimers M., Waters J (2020) **Separation of hemodynamic signals from GCaMP fluorescence measured with wide-field imaging** *J. Neurophysiol* **123**:356–366 https://doi.org/10.1152/jn.00304.2019

Wan J., Peng W., Li X., Qian T., Song K., Zeng J., Deng F., Hao S., Feng J., Zhang P., Zhang Y., Zou J., Pan S., Shin M., Venton B.J., Zhu J.J., Jing M., Xu M., Li Y (2021) **A genetically encoded sensor for measuring serotonin dynamics** *Nat. Neurosci* **24**:746–752 https://doi.org/10.1038/s41593-021-00823-7

Waters J (2020) **Sources of widefield fluorescence from the brain** *eLife* **9**:1–13 https://doi.org/10.7554/eLife.59841

Widmer F.C., O'Toole S.M., Keller G.B (2022) **NMDA receptors in visual cortex are necessary for normal visuomotor integration and skill learning** *eLife* **11**:e71476 https://doi.org/10.7554/eLife.71476

Yogesh B., Keller G.B (2023) **Cholinergic input to mouse visual cortex signals a movement state and acutely enhances layer 5 responsiveness** *eLife* **12** https://doi.org/10.7554/eLife.89986

Zhang W.-T., Chao T.-H.H., Yang Y., Wang T.-W., Lee S.-H., Oyarzabal E.A., Zhou J., Nonneman R., Pegard N.C., Zhu H., Cui G., Shih Y.-Y.I (2022) **Spectral fiber photometry derives hemoglobin concentration changes for accurate measurement of fluorescent sensor activity. Cell Rep** *Methods* **2**:100243 https://doi.org/10.1016/j.crmeth.2022.100243

Zmarz P., Keller G.B (2016) **Mismatch Receptive Fields in Mouse Visual Cortex** *Neuron* **92**:766–772 https://doi.org/10.1016/j.neuron.2016.09.057

Author information

Baba Yogesh
Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland, Faculty of Natural Sciences, University of Basel, Basel, Switzerland
ORCID iD: 0000-0003-1224-6035

Matthias Heindorf
Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland
ORCID iD: 0000-0002-7866-7601

Rebecca Jordan
Simons Initiative for the Developing Brain, University of Edinburgh, Edinburgh, United Kingdom
ORCID iD: 0000-0002-4871-6265

Georg B Keller
Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland, Faculty of Natural Sciences, University of Basel, Basel, Switzerland
ORCID iD: 0000-0002-1401-0117

For correspondence: georg.keller@fmi.ch

Editors

Reviewing Editor
Leopoldo Petreanu
Champalimaud Center for the Unknown, Lisbon, Portugal

Senior Editor
John Huguenard
Stanford University School of Medicine, Stanford, United States of America

Reviewer #1 (Public review):

Summary:

Fluorescence imaging has become an increasingly popular technique for monitoring neuronal activity and neurotransmitter concentrations in the living brain. However, factors such as brain motion and changes in blood flow and oxygenation can introduce significant artifacts, particularly when activity-dependent signals are small. Yogesh et al. quantified these effects using GFP, an activity-independent marker, under two-photon and wide-field imaging conditions in awake behaving mice. They report significant GFP responses across various brain regions, layers, and behavioral contexts, with magnitudes comparable to those of commonly used activity sensors. These data highlight the need for robust control strategies and careful interpretation of fluorescence functional imaging data.

Strengths:

The effect of hemodynamic occlusion in two-photon imaging has been previously demonstrated in sparsely labeled neurons in V1 of anesthetized animals (see Shen and Kara et al., Nature Methods, 2012). The present study builds on these findings by imaging a

substantially larger population of neurons in awake, behaving mice across multiple cortical regions, layers, and stimulus conditions. The experiments are extensive, the statistical analyses are rigorous, and the results convincingly demonstrate significant GFP responses that must be accounted for in functional imaging experiments.

In the revised version, the authors have provided further methodological details that were lacking in the previous version, expanded discussions regarding alternative explanations of these GFP responses as well as potential mitigation strategies. They also added a quantification of brain motion (Fig. S5) and the fraction of responsive neurons when conducting the same experiment using GCaMP6f (Fig. 3D-3F), among other additional information.

Weaknesses:

(1) The authors have now included a detailed methodology for blood vessel area quantification, where they detect blood vessels as dark holes in GFP images and measure vessel area by counting pixels below a given intensity threshold (line 437-443). However, this approach has a critical caveat: any unspecific decrease in image fluorescence will increase the number of pixels below the threshold, leading to an apparent increase in blood vessel area, even when the actual vessel size remains unchanged. As a result, this method inherently introduces a positive correlation between fluorescence decrease and vessel dilation, regardless of whether such a relationship truly exists.

To address this issue, I recommend labelling blood vessels with an independent marker, such as a red fluorescence dye injected into the bloodstream. This approach would allow vessel dilation to be assessed independently of GFP fluorescence -- dilation would cause opposite fluorescence changes in the green and red channels (i.e., a decrease in green due to hemodynamic occlusion and an increase in red due to the expanding vessel area). In my opinion, only when such ani-correlation is observed can one reliably infer a relationship between GFP signal changes and blood vessel dynamics.

Because this relationship is central to the author's conclusion regarding the nature of the observed GFP signals, including this experiment would greatly strengthen the paper's conclusion.

(2) Regarding mitigation strategy, the authors advocate repeating key functional imaging experiments using GFP, and state that their aim here is to provide a control for their 2012 study (Keller et al., Neuron). Given this goal, I find it important to discuss how these new findings impact the interpretation of their 2012 results, particularly given the large GFP responses observed.

For example, Keller et al. (2012) concluded that visuomotor mismatch strongly drives V1 activity (Fig. 3A in that study). However, in the present study, mismatch fails to produce any hemodynamic/GFP response (Fig. 3A, 3B, rightmost bar), and the corresponding calcium response is also the weakest among the three tested conditions (Fig. 3D). How do these findings affect their 2012 conclusions?

Similarly, the present study shows that GFP reveals twice as many responsive neurons as GCaMP during locomotion (Fig. 3A vs. Fig. 3D, "running"). Does this mean that their 2012 conclusions regarding locomotion-induced calcium activity need reconsideration? Given that more neurons responded with GFP than with GCaMP, the authors should clarify whether they still consider GCaMP a reliable tool for measuring brain activity during locomotion.

(3) More generally, the author should discuss how functional imaging data should be interpreted going forward, given the large GFP responses reported here. Even when key experiments are repeated using GFP, it is not entirely clear how one could reliably estimate underlying neuronal activity from the observed GFP and GCaMP responses.

For example, consider the results in Fig. 3A vs. 3D: how should one assess the relative strength of neuronal activity elicited by running, grating, or visuomotor mismatch? Does mismatch produce the strongest neuronal activity, since it is least affected by the hemodynamic/GFP confounds (Fig. 3A)? Or does mismatch actually produce the weakest neuronal activity, given that both its hemodynamic and calcium responses are the smallest?

In my opinion, such uncertainty makes it difficult to robustly interpret functional imaging results. Simply repeating experiments with GFP does not fully resolve this issue, as it does not provide a clear framework for quantifying the underlying neuronal activity. Does this suggest a need for a better mitigation strategy? What could these strategies be?

In my opinion, addressing these questions is critical not only for the authors' own work but also for the broader field to ensure a robust and reliable interpretation of functional imaging data.

(4) The authors now discuss various alternative sources of the observed GFP signals. However, I feel that they often appear to dismiss these possibilities too quickly, rather than appreciating their true potential impacts (see below).

For example, the authors argue that brain movement cannot explain their data, as movement should only result in a decrease in observed fluorescence. However, while this might hold for x-y motion, movement in the axial (z) direction can easily lead to both fluorescence increase and decrease. Neurons are not always precisely located at the focal plane -- some are slightly above or below. Axial movement in a given direction will bring some cells into focus while moving others out of focus, leading to fluorescence changes in both directions, exactly as observed in the data (see Fig. S2).

Furthermore, the authors state that they discard data with 'visible' z-motion. However, subtle axial movements that escape visual detection could still cause fluorescence fluctuations on the order of a few percent, comparable to the reported signal amplitudes.

Finally, the authors state that "brain movement kinematics are different in shape than the GFP responses we observe". However, this appears to contradict what they show in Fig. 2A. Specifically, the first example neuron exhibits fast GFP transients locked to running onset, with rapid kinematics closely matching the movement speed signals in Fig. S5A. These fast transients are incompatible with slower blood vessel area signals (Fig. 4), suggesting that alternative sources could contribute significantly.

In sum, the possibility that alternative signal sources could significantly contribute should be taken seriously and more thoroughly discussed.

(5) The authors added a quantification of brain movement (Fig. S5) and claim that they "only find detectable brain motion during locomotion onsets and not the other stimuli." However, Fig. S5 presents brain 'velocity' rather than 'displacement'. A constant (non-zero) velocity in Fig. S5 B-D indicates that the brain continues to move over time, potentially leading to significant displacement from its initial position across all conditions. While displacement in the x-y plane are corrected, similar displacement in the z direction likely occurs concurrently and cannot be easily accounted for. To assess this possibility, the authors should present absolute displacement relative to pre-stimulus frames, as displacement -- not velocity -- determines the size of movement-related fluorescence changes.

(6) In line 132-133, the authors draw an analogy between the effect of hemodynamic occlusion and liquid crystal display (LCD) function. However, there are fundamental differences between the two. LCDs modulate light transmission by rotating the polarization of light, which then passes through a crossed polarizer. In contrast, hemodynamic occlusion alters light transmission by changing the number and absorbance properties of hemoglobin.

Additionally, LCDs do not involve 'emission' light - back-illumination travels through the liquid crystal layer only once, whereas hemodynamic occlusion affects both incoming excitation light and the emitted fluorescence. Given these fundamental differences, the LCD analogy may not be entirely appropriate.

https://doi.org/10.7554/eLife.104914.2.sa3

Reviewer #2 (Public review):

- Approach

In this study, Yogesh et al. aimed at characterizing hemodynamic occlusion in two photon imaging, where its effects on signal fluctuations are underappreciated compared to that in wide field imaging and fiber photometry. The authors used activity-independent GFP fluorescence, GCaMP and GRAB sensors for various neuromodulators in two-photon and widefield imaging during a visuomotor context to evaluate the extent of hemodynamic occlusion in V1 and ACC. They found that the GFP responses were comparable in amplitude to smaller GCaMP responses, though exhibiting context-, cortical region-, and depth-specific effects. After quantifying blood vessel diameter change and surrounding GFP responses, they argued that GFP responses were highly correlated with changes in local blood vessel size. Furthermore, when imaging with GRAB sensors for different neuromodulators, they found that sensors with lower dynamic ranges such as GRAB-DA1m, GRAB-5HT1.0, and GRAB-NE1m exhibited responses most likely masked by the hemodynamic occlusion, while a sensor with larger SNR, GRAB-ACh3.0, showed much more distinguishable responses from blood vessel change. They thoroughly investigate other factors that could contribute to these signals and demonstrate hemodynamic occlusion is the primary cause.

- Impact of revision

This is an important update to the initial submission, adding much supplemental imaging and population data that provide greater detail to the analyses and increase the confidence in the authors conclusions.

Specifically, inclusion of the supplemental figures 1 and 2 showing GFP expression across multiple regions and the fluorescence changes of thousands of individual neurons provides a clearer picture of how these effects are distributed across the population. Characterization of brain motion across stimulation conditions in supplemental figure 5 provides strong evidence that the fluorescence changes observed in many of the conditions are unlikely to be primarily due to brain motion associated imaging artifacts. The role of vascular area on fluorescence is further supported by addition of new analyses on vasoconstriction leading to increased fluorescence in Figures 4C1-4, complementing the prior analyses of vasodilation.

The expansion of the discussion on other factors that could lead to these changes is thorough and welcome. The arguments against pH playing a factor in fluorescence changes of GFP, due to insensitivity to changes in the expected pH range are reasonable, as are the other discussed potential factors.

With respect to the author's responses to prior critique, we agree that activity dependent hemodynamic occlusion is best investigated under awake conditions. Measurement of these dynamics under anesthesia could lead to an underestimation of their effects. Isoflurane anesthesia causes significant vasodilation and a large reduction in fluorescence intensity in non-functional mutant GRABs. This could saturate or occlude activity dependent effects.

- Strengths

This work is of broad interest to two photon imaging users and GRAB developers and users. It thoroughly quantifies the hemodynamic driven GFP response and compares it to previously published GCaMP data in a similar context, and illustrates the contribution of hemodynamic occlusion to GFP and GRAB responses by characterizing the local blood vessel diameter and fluorescence change. These findings provide important considerations for the imaging community and a sobering look at the utility of these sensors for cortical imaging.

Importantly, they draw clear distinctions between the temporal dynamics and amplitude of hemodynamic artifacts across cortical regions and layers. Moreover, they show context dependent (Dark versus during visual stimuli) effects on locomotion and optogenetic light-triggered hemodynamic signals.

The authors suggest that signal to noise ratio of an indicator likely affects the ability to separate hemodynamic response from the underlying fluorescence signal. With a new analysis (Supplemental Figure 4) They show that the relative degree of background fluorescence does not affect the size of the artifact.

Most of the first generation neuromodulator GRAB sensors showed relatively small responses, comparable to blood vessel changes in two photon imaging, which emphasizes a need for improved the dynamic range and response magnitude for future sensors and encourages the sensor users to consider removing hemodynamic artifacts when analyzing GRAB imaging data.

- Weaknesses

The largest weakness of the paper remains that, while they convincingly quantify hemodynamic artifacts across a range of conditions, they provide limited means of correcting for them. However they now discuss the relative utility of some hemodynamic correction methods (e.g. from Ocana-Santero et al., 2024).

The paper attributes the source of 'hemodynamic occlusion' primarily to blood vessel dilation, but leaves unanswered how much may be due to shifts in blood oxygenation. Figure 4 directly addresses the question of how much of the signal can be attributed to occlusion by measuring the blood vessel dilation, and has been improved by now showing positive fluorescence effects with vasoconstriction. They now also discuss the potential impact of oxygenation.

Along these lines, the authors carefully quantified the correlation between local blood vessel diameter and GFP response (or neuropil fluorescence vs blood vessel fluorescence with GRAB sensors). We are left to wonder to what extent does this effect depend on proximity to the vessels? Do GFP/ GRAB responses decorrelate from blood vessel activity in neurons further from vessels (refer to Figure 5A and B in Neyhart et al., Cell Reports 2024)? The authors argue that the primary impact of occlusion is from blood vessels above the plane of imaging, but without a vascular reconstruction, their evidence for this is anecdotal.

The choice of ACC as the frontal region provides a substantial contrast in location, brain movement, and vascular architecture as compared to V1. As the authors note, ACC is close to the superior sagittal sinus and thus is the region where the largest vascular effects are likely to occur. A less medial portion of M2 may have been a more appropriate comparison. The authors now include example imaging fields for ACC and interesting out-of-plane vascular examples in the supplementary figures that help assess these impacts.

-Overall Assessment

This paper is an important contribution to our understanding of how hemodynamic artifacts may corrupt GRAB and calcium imaging, even in two-photon imaging modes. While it would

be wonderful if the authors were able to demonstrate a reliable way to correct for hemodynamic occlusion which did not rely on doing the experiments over with a non-functional sensor or fluorescent protein, the careful measurement and reporting of the effects here is, by itself, a substantial contribution to the field of neural activity imaging. It's results are of importance to anyone conducting two-photon or widefield imaging with calcium and GRAB sensors and deserves the attention of the broader neuroscience and in-vivo imaging community.

https://doi.org/10.7554/eLife.104914.2.sa2

Reviewer #3 (Public review):

Summary:

In this study, the authors aimed to investigate if hemodynamic occlusion contributes to fluorescent signals measured with two-photon microscopy. For this, they image the activity-independent fluorophore GFP in 2 different cortical areas, at different cortical depths and in different behavioral conditions. They compare the evoked fluorescent signals with those obtained with calcium sensors and neuromodulator sensors and evaluate their relationship to vessel diameter as a readout of blood flow.
They find that GFP fluorescence transients are comparable to GCaMP6f stimuli-evoked signals in amplitude, although they are generally smaller. Yet, they are significant even at the single neuronal level. They show that GFP fluorescence transients resemble those measured with the dopamine sensor GRAB-DA1m and the serotonin sensor GRAB-5HT1.0 in amplitude an nature, suggesting that signals with these sensors are dominated by hemodynamic occlusion. Moreover, the authors perform similar experiments with wide-field microscopy which reveals the similarity between the two methods in generating the hemodynamic signals. Together the evidence presented calls for the development and use of high dynamic range sensors to avoid measuring signals that have another origin from the one intended to measure. In the meantime, the evidence highlights the need to control for those artifacts such as with the parallel use of activity independent fluorophores.

Strengths:

- Comprehensive study comparing different cortical regions in diverse behavioral settings in controlled conditions.
- Comparison to the state-of-the-art, i.e. what has been demonstrated with wide-field microscopy.
- Comparison to diverse activity-dependent sensors, including the widely used GCaMP.

Comments on revisions:

The authors have addressed my concerns well. I have no further comments.

https://doi.org/10.7554/eLife.104914.2.sa1

Author response:

The following is the authors' response to the current reviews.

We thank you for the time you took to review our work and for your feedback! We have made only minor changes in this submission and primarily wanted to respond to the concerns raised by reviewer 1.

Reviewer #1 (Public review):

Summary:

Fluorescence imaging has become an increasingly popular technique for monitoring neuronal activity and neurotransmitter concentrations in the living brain. However, factors such as brain motion and changes in blood flow and oxygenation can introduce significant artifacts, particularly when activitydependent signals are small. Yogesh et al. quantified these effects using GFP, an activity-independent marker, under two-photon and wide-field imaging conditions in awake behaving mice. They report significant GFP responses across various brain regions, layers, and behavioral contexts, with magnitudes comparable to those of commonly used activity sensors. These data highlight the need for robust control strategies and careful interpretation of fluorescence functional imaging data.

Strengths:

The effect of hemodynamic occlusion in two-photon imaging has been previously demonstrated in sparsely labeled neurons in V1 of anesthetized animals (see Shen and Kara et al., Nature Methods, 2012). The present study builds on these findings by imaging a substantially larger population of neurons in awake, behaving mice across multiple cortical regions, layers, and stimulus conditions. The experiments are extensive, the statistical analyses are rigorous, and the results convincingly demonstrate significant GFP responses that must be accounted for in functional imaging experiments.

In the revised version, the authors have provided further methodological details that were lacking in the previous version, expanded discussions regarding alternative explanations of these GFP responses as well as potential mitigation strategies. They also added a quantification of brain motion (Fig. S5) and the fraction of responsive neurons when conducting the same experiment using GCaMP6f (Fig. 3D-3F), among other additional information.

Weaknesses:

(1) The authors have now included a detailed methodology for blood vessel area quantification, where they detect blood vessels as dark holes in GFP images and measure vessel area by counting pixels below a given intensity threshold (line 437-443). However, this approach has a critical caveat: any unspecific decrease in image fluorescence will increase the number of pixels below the threshold, leading to an apparent increase in blood vessel area, even when the actual vessel size remains unchanged. As a result, this method inherently introduces a positive correlation between fluorescence decrease and vessel dilation, regardless of whether such a relationship truly exists.

To address this issue, I recommend labelling blood vessels with an independent marker, such as a red fluorescence dye injected into the bloodstream. This approach would allow vessel dilation to be assessed independently of GFP fluorescence -- dilation would cause opposite fluorescence changes in the green and red channels (i.e., a decrease in green due to hemodynamic occlusion and an increase in red due to the expanding vessel area). In my opinion, only when such ani-correlation is observed can one reliably infer a relationship between GFP signal changes and blood vessel dynamics.

Because this relationship is central to the author's conclusion regarding the nature of the observed GFP signals, including this experiment would greatly strengthen the paper's conclusion.

This is correct – a more convincing demonstration that blood vessels dilate or constrict anticorrelated with apparent GFP fluorescence would be a separate blood vessel marker. However, we don't think this experiment is worth doing, as it is also not conclusive in the sense the reviewer may have in mind. The anticorrelation does not mean that occlusion drives all of the observed effect. Our main argument is instead that there is no other potential source than hemodynamic occlusion with sufficient strength that we can think of. The experiment one would want to do is block hemodynamic changes and demonstrate that the occlusion explains all of the observed changes.

> *(2) Regarding mitigation strategy, the authors advocate repeating key functional imaging experiments using GFP, and state that their aim here is to provide a control for their 2012 study (Keller et al., Neuron). Given this goal, I find it important to discuss how these new findings impact the interpretation of their 2012 results, particularly given the large GFP responses observed.*

We are happy to discuss how the conclusions of our own work are influenced by this (see more details below), but the important response of the field should probably be to revisit the conclusions of a variety of papers published in the last two decades. This goes far beyond what we can do here.

> *For example, Keller et al. (2012) concluded that visuomotor mismatch strongly drives V1 activity (Fig. 3A in that study). However, in the present study, mismatch fails to produce any hemodynamic/GFP response (Fig. 3A, 3B, rightmost bar), and the corresponding calcium response is also the weakest among the three tested conditions (Fig. 3D). How do these findings affect their 2012 conclusions?*

The average calcium response of L2/3 neurons to visuomotor mismatch is probably roughly similar to the average calcium response at locomotion onset (both are on the order of 1% to 5%, depending on indicator, dataset, etc.). In the Keller et al. (2012) paper, locomotion onset was about 1.5% and mismatch about 3% (see Figure 3A in that paper). What we quantify in Figure 3 of the paper here is the fraction of responsive neurons. Thus, mismatch drives strong responses in a small subset of neurons (approx. 10%), while locomotion drives a combination of a weak responses in a large fraction of the neurons (roughly 70%) and also large responses in a subset of neurons. A strong signal in a subset of neurons is what one would expect from a neuronal response, a weak signal from many neurons would be indicative of a contaminating signal. This all appears consistent.

Regarding influencing the conclusions of earlier work, the movement related signals described in the Keller et al. (2012) paper are probably overestimated, but are also apparent in electrophysiological recordings (Saleem et al., 2013). Thus, the locomotion responses reported in the Keller et al. (2012) paper are likely too high, but locomotion related responses in V1 are very likely real. The only conclusion we draw in the Keller et al. 2012 paper on the strength of the locomotion related responses is that they are smaller than mismatch responses (this conclusion is unaffected by hemodynamic contamination). In addition, the primary findings of the Keller et al. (2012) paper are all related to mismatch, and these conclusions are unaffected.

> *Similarly, the present study shows that GFP reveals twice as many responsive neurons as GCaMP during locomotion (Fig. 3A vs. Fig. 3D, "running"). Does this mean that their 2012 conclusions regarding locomotion-induced calcium activity need reconsideration? Given that more neurons responded with GFP than with GCaMP, the authors should clarify whether they still consider GCaMP a reliable tool for measuring brain activity during locomotion.*

Comparisons of the fraction of significantly responsive neurons between GFP and GCaMP are not straightforward to interpret. One needs to factor in the difference in signal to noise between the two sensors. (Please note, we added the GCaMP responses here upon request of the reviewers). Note, there is nothing inherently wrong with the data, and comparisons within dataset are easily made (e.g. more grating responsive neurons than running responsive neurons in GCaMP, and vice versa with GFP). The comparison across datasets is not as straightforward as we define "responsive neurons" using a statistical test that compares response to baseline activity for each neuron. GFP labelled neurons are very bright and occlusion can easily be detected. Baseline fluorescence in GCaMP recordings is much lower and often close to or below the noise floor of the data (i.e. we only see the cells when they are active). Thus occlusion in GCaMP recordings is preferentially visible for cells that have high baseline fluorescence. Thus, in the GCaMP data we are likely underestimating the fraction of responsive neurons.

Regarding whether GCaMP (or any other fluorescence indicator used in vivo) is a reliable tool, we are not sure we understand. Whenever possible, fluorescence-sensor based measurements should be corrected for hemodynamic contamination – to quantify locomotion related signals this will be more difficult than e.g. for mismatch, but that does not mean it is not reliable.

> *(3) More generally, the author should discuss how functional imaging data should be interpreted going forward, given the large GFP responses reported here. Even when key experiments are repeated using GFP, it is not entirely clear how one could reliably estimate underlying neuronal activity from the observed GFP and GCaMP responses.*

We are not sure we have a good answer to this question. The strategy for addressing this problem will depend on the specifics of the experiment, and the claims. Take the case of mismatch. Here we have strong calcium responses and no evidence of GFP responses. We would argue that this is reasonable evidence that the majority of the mismatch driven GCaMP signal is likely neuronal. For locomotion onsets, both GFP and GCaMP signals go in the same direction on average. Then one could use a response amplitude distribution comparison to conservatively exclude all neurons with a GCaMP amplitude lower than e.g. the 99th percentile of the GFP response. Etc. But we don't think there is an easy generalizable fix for this problem.

> *For example, consider the results in Fig. 3A vs. 3D: how should one assess the relative strength of neuronal activity elicited by running, grating, or visuomotor mismatch? Does mismatch produce the strongest neuronal activity, since it is least affected by the hemodynamic/GFP confounds (Fig. 3A)? Or does mismatch actually produce the weakest neuronal activity, given that both its hemodynamic and calcium responses are the smallest?*

See above, the reviewer may be confounding "response strength" with "fraction of responsive neurons" here. Regarding the relationship between neuronal activity and hemodynamics, it is very likely not just the average activity of all neurons, but a specific subset that drives blood vessel constriction and dilation. This would of course be a very interesting question to answer for the interpretation of hemodynamic based measurements of brain activity, like fMRI, but goes beyond the aim of the current paper.

> *In my opinion, such uncertainty makes it difficult to robustly interpret functional imaging results. Simply repeating experiments with GFP does not fully resolve this issue, as it does not provide a clear framework for quantifying the underlying neuronal activity. Does this suggest a need for a better mitigation strategy? What could these strategies be?*

If the reviewer has a good idea - we would be all ears. We don't have a better idea currently.

> *In my opinion, addressing these questions is critical not only for the authors' own work but also for the broader field to ensure a robust and reliable interpretation of functional imaging data.*

We agree, having a solution to this problem would be important – we just don't have one.

> *(4) The authors now discuss various alternative sources of the observed GFP signals. However, I feel that they often appear to dismiss these possibilities too quickly, rather than appreciating their true potential impacts (see below).*
>
> *For example, the authors argue that brain movement cannot explain their data, as movement should only result in a decrease in observed fluorescence. However, while this might hold for x-y motion, movement in the axial (z) direction can easily lead to both fluorescence increase and decrease. Neurons are not always precisely located at the focal plane -- some are slightly above or below. Axial movement in a given direction will bring some cells into focus while moving others out of focus, leading to fluorescence changes in both directions, exactly as observed in the data (see Fig. S2).*

The reviewer is correct that z-motion can result in an increase of apparent fluorescence (just like x-y motion can as well). On average however, just like with x-y motion, z-motion will always result in a decrease. This assumes that the user selecting regions of interest (the outlines of cells used to quantify fluorescence), will select these such that the distribution of cells selected centers on the zplane of the image. Thus, the distribution of z-location of the cell relative to the imaging plane will be some Gaussian like distribution centered on the z-plane of the image (with half the cell above the zplane and half below). Because the peak of the distribution is located on the z-plane at rest, any zmovement, up or down, will move away from the peak of the distribution (i.e. most cells will decrease in fluorescence). This is the same argument as for why x-y motion always results in decreases (assuming the user selects regions of interest centered on the location of the cells at rest).

> *Furthermore, the authors state that they discard data with 'visible' z-motion. However, subtle axial movements that escape visual detection could still cause fluorescence fluctuations on the order of a few percent, comparable to the reported signal amplitudes.*

Correct, but as explained above, z-motion will always result in average decreases of average fluorescence as explained above.

> *Finally, the authors state that "brain movement kinematics are different in shape than the GFP responses we observe". However, this appears to contradict what they show in Fig. 2A. Specifically, the first example neuron exhibits fast GFP transients locked to running onset, with rapid kinematics closely matching the movement speed signals in Fig. S5A. These fast transients are incompatible with slower blood vessel area signals (Fig. 4), suggesting that alternative sources could contribute significantly.*

We meant population average responses here. We have clarified this. Some of the signals we observed do indeed look like they could be driven by movement artifacts (whole brain motion, or probably more likely blood vessel dilation driven tissue distortion). We show this neuron to illustrate that this can also happen. However, to illustrate that this is a rare event we also show the entire distribution of peak amplitudes and the position in the distribution this neuron is from.

> *In sum, the possibility that alternative signal sources could significantly contribute should be taken seriously and more thoroughly discussed.*

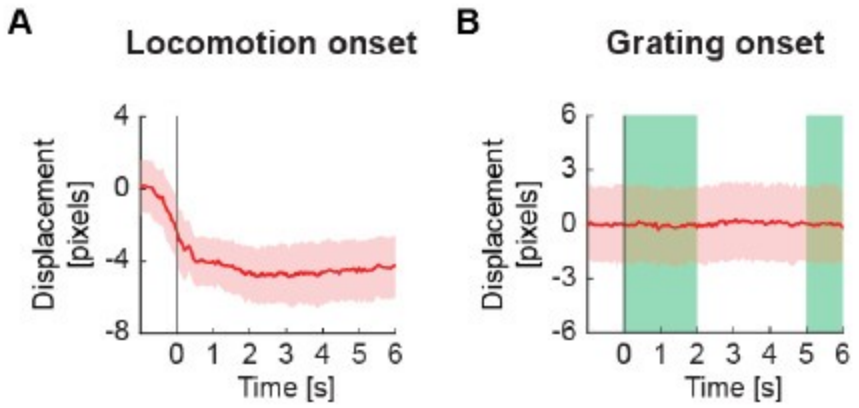
All possible sources (we could think of) are explicitly discussed (in roughly equal proportion). Nevertheless, the reviewer is correct that our focus here is almost exclusively on the what we think is the primary source of the problem. Given that – in my experience – this is also the one least frequently considered, I think the emphasis on – what we think is – the primary contributor is warranted.

> *(5) The authors added a quantification of brain movement (Fig. S5) and claim that they "only find detectable brain motion during locomotion onsets and not the other stimuli." However, Fig. S5 presents brain 'velocity' rather than 'displacement'. A constant (non-zero) velocity in Fig. S5 B-D indicates that the brain continues to move over time, potentially leading to significant displacement from its initial position across all conditions. While displacement in the x-y plane are corrected, similar displacement in the z direction likely occurs concurrently and cannot be easily accounted for. To assess this possibility, the authors should present absolute displacement relative to pre-stimulus frames, as displacement -- not velocity -- determines the size of movement-related fluorescence changes.*

We use brain velocity here as a natural measure when using frame times as time bins. The problem with using a signed displacement is that if different running onsets move the brain in opposing directions, this can average out to zero. To counteract this, one can take the absolute displacement in a response window away from the position in a baseline time window. If this is done with time bins that correspond to frame times, this just becomes displacement per frame, i.e. velocity. Using absolute changes in displacement (i.e. velocity) is more sensitive than signed displacement. The responses for signed displacement are shown below (Author response image 1), but given that we are averaging signed quantities here, the average is not interpretable.

Author response image 1.

Average signed brain displacement.



Regarding a constant drift, the reviewer might be misled by the fact that the baseline brain velocity is roughly 1 pixel per frame. The registration algorithm works in integer number of pixels only. 1 pixel per frame corresponds roughly to the noise floor of the registration

algorithm. Registrations are done independently for each frame. As a consequence, the registration oscillates between a shift of 17 and 18 pixels – frame by frame – if the actual shift is somewhere between 17 and 18 pixels. This “jitter” results in a baseline brain velocity of about 1 pixel per frame.

> *(6) In line 132-133, the authors draw an analogy between the effect of hemodynamic occlusion and liquid crystal display (LCD) function. However, there are fundamental differences between the two. LCDs modulate light transmission by rotating the polarization of light, which then passes through a crossed polarizer. In contrast, hemodynamic occlusion alters light transmission by changing the number and absorbance properties of hemoglobin. Additionally, LCDs do not involve 'emission' light - backillumination travels through the liquid crystal layer only once, whereas hemodynamic occlusion affects both incoming excitation light and the emitted fluorescence. Given these fundamental differences, the LCD analogy may not be entirely appropriate.*

The mechanism of occlusion is, as the reviewer correctly points out, different for an LCD. In both cases however, there is a variable occluder between a light source and an observer. The fact that with hemodynamic occlusion the light passes through the occluder twice (excitation and emission) does not appear to hamper the analogy to us. We have rephrased to highlight the time varying occlusion part.

> ***Reviewer #2 (Public review):***
>
> *- Approach*
>
> *In this study, Yogesh et al. aimed at characterizing hemodynamic occlusion in two photon imaging, where its effects on signal fluctuations are underappreciated compared to that in wide field imaging and fiber photometry. The authors used activity-independent GFP fluorescence, GCaMP and GRAB sensors for various neuromodulators in two-photon and widefield imaging during a visuomotor context to evaluate the extent of hemodynamic occlusion in V1 and ACC. They found that the GFP responses were comparable in amplitude to smaller GCaMP responses, though exhibiting context-, cortical region-, and depth-specific effects. After quantifying blood vessel diameter change and surrounding GFP responses, they argued that GFP responses were highly correlated with changes in local blood vessel size. Furthermore, when imaging with GRAB sensors for different neuromodulators, they found that sensors with lower dynamic ranges such as GRAB-DA1m, GRAB-5HT1.0, and GRAB-NE1m exhibited responses most likely masked by the hemodynamic occlusion, while a sensor with larger SNR, GRAB-ACh3.0, showed much more distinguishable responses from blood vessel change. They thoroughly investigate other factors that could contribute to these signals and demonstrate hemodynamic occlusion is the primary cause.*
>
> *- Impact of revision*
>
> *This is an important update to the initial submission, adding much supplemental imaging and population data that provide greater detail to the analyses and increase the confidence in the authors conclusions.*
>
> *Specifically, inclusion of the supplemental figures 1 and 2 showing GFP expression across multiple regions and the fluorescence changes of thousands of individual neurons provides a clearer picture of how these effects are distributed across the population. Characterization of brain motion across stimulation conditions in supplemental figure 5 provides strong evidence that the fluorescence changes observed in many of the conditions are unlikely to be primarily due to brain motion associated imaging artifacts. The role of vascular area on fluorescence is further supported by addition of new*

analyses on vasoconstriction leading to increased fluorescence in Figures 4C1-4, complementing the prior analyses of vasodilation.

The expansion of the discussion on other factors that could lead to these changes is thorough and welcome. The arguments against pH playing a factor in fluorescence changes of GFP, due to insensitivity to changes in the expected pH range are reasonable, as are the other discussed potential factors.

With respect to the author's responses to prior critique, we agree that activity dependent hemodynamic occlusion is best investigated under awake conditions. Measurement of these dynamics under anesthesia could lead to an underestimation of their effects. Isoflurane anesthesia causes significant vasodilation and a large reduction in fluorescence intensity in non-functional mutant GRABs. This could saturate or occlude activity dependent effects.

- Strengths

This work is of broad interest to two photon imaging users and GRAB developers and users. It thoroughly quantifies the hemodynamic driven GFP response and compares it to previously published GCaMP data in a similar context, and illustrates the contribution of hemodynamic occlusion to GFP and GRAB responses by characterizing the local blood vessel diameter and fluorescence change. These findings provide important considerations for the imaging community and a sobering look at the utility of these sensors for cortical imaging.

Importantly, they draw clear distinctions between the temporal dynamics and amplitude of hemodynamic artifacts across cortical regions and layers. Moreover, they show context dependent (Dark versus during visual stimuli) effects on locomotion and optogenetic light-triggered hemodynamic signals.

The authors suggest that signal to noise ratio of an indicator likely affects the ability to separate hemodynamic response from the underlying fluorescence signal. With a new analysis (Supplemental Figure 4) They show that the relative degree of background fluorescence does not affect the size of the artifact.

Most of the first generation neuromodulator GRAB sensors showed relatively small responses, comparable to blood vessel changes in two photon imaging, which emphasizes a need for improved the dynamic range and response magnitude for future sensors and encourages the sensor users to consider removing hemodynamic artifacts when analyzing GRAB imaging data.

- Weaknesses

The largest weakness of the paper remains that, while they convincingly quantify hemodynamic artifacts across a range of conditions, they provide limited means of correcting for them. However they now discuss the relative utility of some hemodynamic correction methods (e.g. from Ocana-Santero et al., 2024).

The paper attributes the source of 'hemodynamic occlusion' primarily to blood vessel dilation, but leaves unanswered how much may be due to shifts in blood oxygenation. Figure 4 directly addresses the question of how much of the signal can be attributed to occlusion by measuring the blood vessel dilation, and has been improved by now showing positive fluorescence effects with vasoconstriction. They now also discuss the potential impact of oxygenation.

Along these lines, the authors carefully quantified the correlation between local blood vessel diameter and GFP response (or neuropil fluorescence vs blood vessel fluorescence

with GRAB sensors). We are left to wonder to what extent does this effect depend on proximity to the vessels? Do GFP/ GRAB responses decorrelate from blood vessel activity in neurons further from vessels (refer to Figure 5A and B in Neyhart et al., Cell Reports 2024)? The authors argue that the primary impact of occlusion is from blood vessels above the plane of imaging, but without a vascular reconstruction, their evidence for this is anecdotal.

The choice of ACC as the frontal region provides a substantial contrast in location, brain movement, and vascular architecture as compared to V1. As the authors note, ACC is close to the superior sagittal sinus and thus is the region where the largest vascular effects are likely to occur. A less medial portion of M2 may have been a more appropriate comparison. The authors now include example imaging fields for ACC and interesting out-of-plane vascular examples in the supplementary figures that help assess these impacts.

-Overall Assessment

This paper is an important contribution to our understanding of how hemodynamic artifacts may corrupt GRAB and calcium imaging, even in two-photon imaging modes. While it would be wonderful if the authors were able to demonstrate a reliable way to correct for hemodynamic occlusion which did not rely on doing the experiments over with a non-functional sensor or fluorescent protein, the careful measurement and reporting of the effects here is, by itself, a substantial contribution to the field of neural activity imaging. It's results are of importance to anyone conducting two-photon or widefield imaging with calcium and GRAB sensors and deserves the attention of the broader neuroscience and invivo imaging community.

We agree with this assessment.

***Reviewer #3 (Public review)**:*

Summary:

In this study, the authors aimed to investigate if hemodynamic occlusion contributes to fluorescent signals measured with two-photon microscopy. For this, they image the activity-independent fluorophore GFP in 2 different cortical areas, at different cortical depths and in different behavioral conditions. They compare the evoked fluorescent signals with those obtained with calcium sensors and neuromodulator sensors and evaluate their relationship to vessel diameter as a readout of blood flow.

They find that GFP fluorescence transients are comparable to GCaMP6f stimuli-evoked signals in amplitude, although they are generally smaller. Yet, they are significant even at the single neuronal level. They show that GFP fluorescence transients resemble those measured with the dopamine sensor GRABDA1m and the serotonin sensor GRAB-5HT1.0 in amplitude an nature, suggesting that signals with these sensors are dominated by hemodynamic occlusion. Moreover, the authors perform similar experiments with wide-field microscopy which reveals the similarity between the two methods in generating the hemodynamic signals. Together the evidence presented calls for the development and use of high dynamic range sensors to avoid measuring signals that have another origin from the one intended to measure. In the meantime, the evidence highlights the need to control for those artifacts such as with the parallel use of activity independent fluorophores.

Strengths:

- Comprehensive study comparing different cortical regions in diverse behavioral settings in controlled conditions.

> *- Comparison to the state-of-the-art, i.e. what has been demonstrated with wide-field microscopy.*
>
> *- Comparison to diverse activity-dependent sensors, including the widely used GCaMP.*
>
> *Comments on revisions:*
>
> *The authors have addressed my concerns well. I have no further comments.*

We agree with this assessment.

The following is the authors' response to the original reviews

The major changes to the manuscript are:

(1) Re-wrote the discussion, going over all possible sources of the signals we describe.

(2) We added a quantification of brain motion as Figure S5.

(3) We added an example of blood vessel contraction as Figure 4C.

(4) We added data on the fraction of responsive neurons when measured with GCaMP as Figures 3D-3F.

(5) We added example imaging sites from all imaged regions as Figure S1.

(6) We added GFP response heatmaps of all neurons as Figure S2.

(7) We add a quantification of the relationship between GFP response amplitude and expression level Figure S4.

A detailed point-by-point response to all reviewer concerns is provided below.

> ***Public Reviews:***
>
> ***Reviewer #1 (Public Review):***
>
> *Fluorescence imaging has become an increasingly popular technique for monitoring neuronal activity and neurotransmitter concentrations in the living brain. However, factors such as brain motion and changes in blood flow and oxygenation can introduce significant artifacts, particularly when activity-dependent signals are small. Yogesh et al. quantified these effects using GFP, an activity-independent marker, under two-photon and wide-field imaging conditions in awake behaving mice. They report significant GFP responses across various brain regions, layers, and behavioral contexts, with magnitudes comparable to those of commonly used activity sensors. These data highlight the need for robust control strategies and careful interpretation of fluorescence functional imaging data.*
>
> *Strengths:*
>
> *The effect of hemodynamic occlusion in two-photon imaging has been previously demonstrated in sparsely labeled neurons in V1 of anesthetized animals (see Shen and Kara et al., Nature Methods, 2012). The present study builds on these findings by imaging a substantially larger population of neurons in awake, behaving mice across multiple cortical regions, layers, and stimulus conditions. The experiments are extensive, the statistical analyses are rigorous, and the results convincingly demonstrate significant GFP responses that must be accounted for in functional imaging experiments. However,*

> *whether these GFP responses are driven by hemodynamic occlusion remains less clear, given the complexities associated with awake imaging and GFP's properties (see below).*
>
> *Weaknesses:*
>
> *(1) The authors primarily attribute the observed GFP responses to hemodynamic occlusion. While this explanation is plausible, other factors may also contribute to the observed signals. These include uncompensated brain movement (e.g., axial-direction movements), leakage of visual stimulation light into the microscope, and GFP's sensitivity to changes in intracellular pH (see e.g., Kneen and Verkman, 1998, Biophysical Journal). Although the correlation between GFP signals and blood vessel diameters supports a hemodynamic contribution, it does not rule out significant contributions from these (or other) factors. Consequently, whether GFP fluorescence can reliably quantify hemodynamic occlusion in two-photon microscopy remains uncertain.*

We concur; our data do not conclusively prove that the effect is only driven by hemodynamic occlusion. We have attempted to make this clearer in the text throughout the manuscript. In particular we have restructured the discussion to focus on this point. Regarding the specific alternatives the reviewer mentions here:

a) Uncompensated brain motion. While this can certainly contribute, we think the effect is negligible in our interpretation for the following reasons. First, just to point out the obvious, as with all two-photon data we acquire in the lab, we only keep data with no visible z-motion (axial). Second, and more importantly, uncompensated brain motion results in a net decrease of fluorescence. As regions of interest (ROI) are selected to be centered on neurons (as opposed to be randomly selected, or next to, or above or below), movement will – on average – result in a decrease in fluorescence, as neurons are moved out of the ROIs. In the early days of awake two-photon imaging (when preps were still less stable) – we used this movement onset decrease in fluorescence as a sign that running onsets were selected correctly (i.e. with low variance). See e.g. the dip in the running onset trace at time zero in figure 3A of (Keller et al., 2012). Third, we find no evidence for any brain motion in the case of visual stimulation, while the GFP responses during locomotion and visual stimulation are of similar magnitude. We have added a quantification of brain motion (Figure S5) and a discussion of this point to the manuscript.

b) Leakage of stimulation light. First, all light sources in the experimental room (the projector used for the mouse VR, the optogenetic stimulation light, as well as the computer monitors used to operate the microscope) are synchronized to the turnaround times of the resonant scanner of the two-photon microscope. Thus, light sources in the room are turned off for each line scan of the resonant scanner and turned on in the turnaround period. With a 12kHz scanner this results in a light cycle of 24 kHz (see Leinweber et al., 2014 for details). While the system is not perfect, we can occasionally get detectable light leak responses at the image edges (in the resonant axis as a result of the exponential off kinetics of many LEDs & lasers), these are typically 2 orders of magnitude smaller than what one would get without synchronizing, and far smaller than a single digit percentage change in GFP responses, and only detectable at the image edges. Second, while in visual cortex, dark running onsets are different from running onsets with the VR turned on (Figures 5A and B), they are indistinguishable in ACC (Figure 5C). Thus, stimulation light artefacts we can rule out.

c) GFP's sensitivity to changes in pH. Activity results in a decrease in neuronal intracellular pH (https://pubmed.ncbi.nlm.nih.gov/14506304/, https://pubmed.ncbi.nlm.nih.gov/24312004/) – decreasing pH decreases GFP fluorescence (https://pubmed.ncbi.nlm.nih.gov/9512054/).

To reiterate, we don't think hemodynamic occlusion is the only possible source to the effects we observe, but we do think it is most likely the largest.

> *(2) Regardless of the underlying mechanisms driving the GFP responses, these activity-independent signals must be accounted for in functional imaging experiments. However, the present manuscript does not explore potential strategies to mitigate these effects. Exploring and demonstrating even partial mitigation strategies could have significant implications for the field.*

We concur – however, in brief, we think the only viable mitigation strategy (we are capable of), is to repeat functional imaging with GFP imaging. To unpack this: There have been numerous efforts to mitigate these hemodynamic effects using isosbestic illumination. When we started to use such strategies in the lab for widefield imaging, we thought we would calibrate the isosbestic correction using GFP recordings. The idea was that if performed correctly, an isosbestic response should look like a GFP response. Try as we may, we could not get the isosbestic responses to look like a GFP response. We suspect this is a result of the fact that none of the light sources we used were perfectly match to the isosbestic wavelength the GCaMP variants we used (not for a lack of trying, but neither lasers nor LEDs were available for purchase with exact wavelength matches). Complicating this was then also the fact that the similarity (or dissimilarity) between isosbestic and GFP responses was a function of brain region. Importantly however, just because we could not successfully apply isosbestic corrections, of course does not mean it cannot be done. Hence for the widefield experiments we then resorted to mitigating the problem by repeating the key experiments using GFP imaging (see e.g. (Heindorf and Keller, 2024)). Note, others have also argued that the best way to correct for hemodynamic artefacts is a GFP recording based correction (Valley et al., 2019). A second strategy we tried was using a second fluorophore (i.e. a red marker) in tandem with a GCaMP sensor. The problem here is that the absorption of the two differs markedly by blood and once again a correction of the GCaMP signal using the red channel was questionable at best. Thus, we think the only viable mitigation strategy we have found is GFP recordings and testing whether the postulated effects seen with calcium indicators are also present in GFP responses. This work is our attempt at a post-hoc mitigation of the problem of our own previous two-photon imaging studies.

> *(3) Several methodology details are missing from the Methods section. These include: (a) signal extraction methods for two-photon imaging data (b) neuropil subtraction methods (whether they are performed and, if so, how) (c) methods used to prevent visual stimulation light from being detected by the two-photon imaging system (d) methods to measure blood vessel diameter/area in each frame. The authors should provide more details in their revision.*

Please excuse, this was an oversight. All details have been added to the methods.

> ***Reviewer #2 (Public Review):***
>
> *In this study, Yogesh et al. aimed at characterizing hemodynamic occlusion in two photon imaging, where its effects on signal fluctuations are underappreciated compared to that in wide field imaging and fiber photometry. The authors used activity-independent GFP fluorescence, GCaMP and GRAB sensors for various neuromodulators in two-photon and widefield imaging during a visuomotor context to evaluate the extent of hemodynamic occlusion in V1 and ACC. They found that the GFP responses were comparable in amplitude to smaller GCaMP responses, though exhibiting context-, cortical region-, and depth-specific effects. After quantifying blood vessel diameter change and surrounding GFP responses, they argued that GFP responses were highly correlated with changes in local blood vessel size. Furthermore, when imaging with GRAB sensors for different neuromodulators, they found that sensors with lower dynamic ranges such as GRAB-DA1m, GRAB5HT1.0, and GRAB-NE1m exhibited responses most*

likely masked by the hemodynamic occlusion, while a sensor with larger SNR, GRAB-ACh3.0, showed much more distinguishable responses from blood vessel change.

Strengths

This work is of broad interest to two photon imaging users and GRAB developers and users. It thoroughly quantifies the hemodynamic driven GFP response and compares it to previously published GCaMP data in a similar context, and illustrates the contribution of hemodynamic occlusion to GFP and GRAB responses by characterizing the local blood vessel diameter and fluorescence change. These findings provide important considerations for the imaging community and a sobering look at the utility of these sensors for cortical imaging.

Importantly, they draw clear distinctions between the temporal dynamics and amplitude of hemodynamic artifacts across cortical regions and layers. Moreover, they show context dependent (Dark versus during visual stimuli) effects on locomotion and optogenetic light-triggered hemodynamic signals.

Most of the first generation neuromodulator GRAB sensors showed relatively small responses, comparable to blood vessel changes in two photon imaging, which emphasizes a need for improved the dynamic range and response magnitude for future sensors and encourages the sensor users to consider removing hemodynamic artifacts when analyzing GRAB imaging data.

Weaknesses

(1) The largest weakness of the paper is that, while they convincingly quantify hemodynamic artifacts across a range of conditions, they do not quantify any methods of correcting for them. The utility of the paper could have been greatly enhanced had they tested hemodynamic correction methods (e.g. from Ocana-Santero et al., 2024) and applied them to their datasets. This would serve both to verify their findings-proving that hemodynamic correction removes the hemodynamic signal-and to act as a guide to the field for how to address the problem they highlight.

See also our response to reviewer 1 comment 2.

In the Ocana-Santero et al., 2024 paper they also first use GFP recordings to identify the problem. The mitigation strategy they then propose, and use, is to image a second fluorophore that emits at a different wavelength concurrently with the functional indicator. The authors then simply subtract (we think – the paper states “divisive”, but the data shown are more consistent with “subtractive” correction) the two signals to correct for hemodynamics. However, the paper does not demonstrate that the hemodynamic signals in the red channel match those in the green channel. The evidence presented that this works is at best anecdotal. In our hands this does not work (meaning the red channel does not match GFP recordings), we suspect this is a combination of crosstalk from the simultaneously recorded functional channel and the fact that hemodynamic absorption is strongly wavelength specific, or something we are doing wrong. Either way, we cannot contribute to this in the form of mitigation strategy.

Given that the GFP responses are a function of brain area and cortical depth – it is not a stretch to postulate that they also depend on genetic cell type labelled. Thus, any GFP calibration used for correction will need to be repeated for each cell type and brain area. Once experiments are repeated using GFP (the strategy we advocate for – we don’t think there is a simpler way to do this), the “correction” is just a subtraction (or a visual comparison).

> *(2) The paper attributes the source of 'hemodynamic occlusion' primarily to blood vessel dilation, but leaves unanswered how much may be due to shifts in blood oxygenation. Figure 4 directly addresses the question of how much of the signal can be attributed to occlusion by measuring the blood vessel dilation, but notably fails to reproduce any of the positive transients associated with locomotion in Figure 2. Thus, an investigation into or at least a discussion of what other factors (movement? Hb oxygenation?) may drive these distinct signals would be helpful.*

See also our response to reviewer 1 comment 1.

We have added to Figure 4 an example of a positive transient. At running onset, superficial blood vessels in cortex tend to constrict and hence result in positive transients.

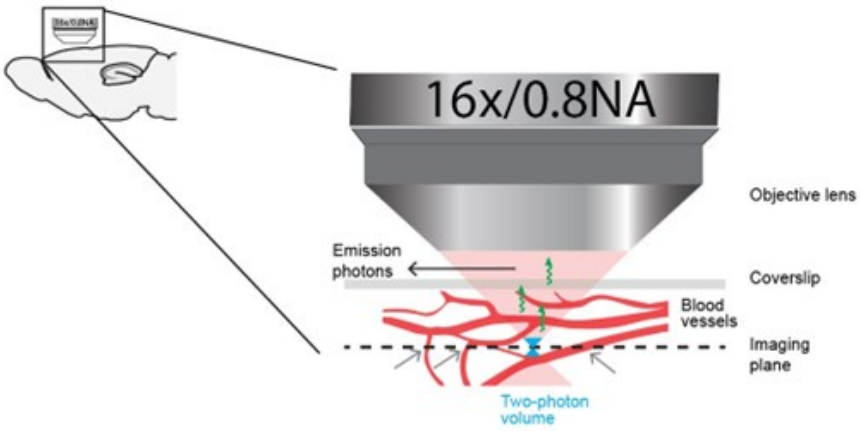
We now also mention changes in blood oxygenation as a potential source of hemodynamic occlusion. And just to be clear, blood oxygenation (or flow) changes in absence of any fluorophore, do not lead to a two-photon signal. Just in case the reviewer was concerned about intrinsic signals – these are not detectable in two photon imaging.

> *(3) Along these lines, the authors carefully quantified the correlation between local blood vessel diameter and GFP response (or neuropil fluorescence vs blood vessel fluorescence with GRAB sensors). To what extent does this effect depend on proximity to the vessels? Do GFP/ GRAB responses decorrelate from blood vessel activity in neurons further from vessels (refer to Figure 5A and B in Neyhart et al., Cell Reports 2024)?*

We indeed thought about quantifying this, but to do this properly would require having a 3d reconstruction of the blood vessel plexus above (with respect to the optical axis) the neuron of interest, as well as some knowledge of how each vessel dilates as a function of stimulus. The prime effect is likely from blood vessels that are in the 45 degrees illumination cone above the neuron (Author response image 2). Lateral proximity to a blood vessel is likely only of secondary relevance. Thus, performing such a measurement is impractical and of little benefit for others.

Author response image 2.

A schematic representation of the cone of illumination.



While imaging a neuron (the spot on the imaging plane at the focus of the cone of illumination), the relevant blood vessels that primarily contribute to hemodynamic occlusion

are those in the cone of illumination between the neuron and the objective lens. Blood vessels visible in the imaging plane (indicated by gray arrows), do not directly contribute to hemodynamic occlusion. Any distance dependence of hemodynamic occlusion in the observed response of a neuron to these blood vessels in the imaging plane is at best incidental.

> *(4) Raw traces are shown in Figure 2 but we are never presented with the unaveraged data for locomotion of stimulus presentation times, which limits the reader's ability to independently assess variability in the data. Inclusion of heatmaps comparing event aligned GFP to GCaMP6f may be of value to the reader.*

We fear we are not sure what the reviewer means by “the unaveraged data for locomotion of stimulus presentation times”. We suspect this should read “locomotion or stimulus…”. We have added heat maps of the responses of all neurons of the data shown in Figure 1 – as Figure S2.

> *(5) More detailed analysis of differences between the kinds of dynamics observed in GFP vs GCaMP6f expressing neurons could aid in identifying artifacts in otherwise clean data. The example neurons in Figure 2A hint at this as each display unique waveforms and the question of whether certain properties of their dynamics can reveal the hemodynamic rather than indicator driven nature of the signal is left open. Eg. do the decay rate and rise times differ significantly from GCaMP6f signals?*

The most informative distinction we have found is differences in peak responses (Figure 2B). Decay and rise time measurements critically depend on the identification of “events”. As a function of how selective one is with what one calls an event (e.g. easy in example 1 of Figure 2 – but more difficult in examples 2 and 3), one gets very different estimates of rise and decay times. Due to the fact that peak amplitudes are lower in GFP responses – rise and decay times will be either slower or noisier (depending on where the threshold for event detection is set).

> *(6) The authors suggest that signal to noise ratio of an indicator likely affects the ability to separate hemodynamic response from the underlying fluorescence signal. Does the degree of background fluorescence affect the size of the artifact? If there was variation in background and overall expression level in the data this could potentially be used to answer this question. Could lower (or higher!) expression levels increase the effects of hemodynamic occlusion?*

There may be a misunderstanding (i.e. we might be misunderstanding the reviewer’s argument here). Our statement from the manuscript that the signal to noise ratio of an indicator matters is based on the simple consideration that hemodynamic occlusion is in the range of 0 to 2 % ΔF/F. The larger the dynamic range of the indicator, the less of a problem 2% ΔF/F are. Imagine an indicator with average responses in the 100’s of % ΔF/F - then this would be a non-problem. For indicators with a dynamic range less than 1%, a 2% artifact is a problem.

Regarding “background” fluorescence, we are not sure what is meant here. In case the reviewer means fluorescence that comes from indicator molecules in processes (as opposed to soma) that are typically ignored (or classified as neuropil) – we are not sure how this would help. The occlusion effects are identical for both somatic and axonal or dendritic GFP (the source of the GFP fluorescence is not relevant for the occlusion effect). In case the reviewer means “baseline” fluorescence – above a noise threshold $\Delta F/F_0$ should be constant independent of F_0 (i.e. baseline fluorescence). This also holds in the data, see Figure S4. We might be stating the trivial - the normalization of fluorescence activity as $\Delta F/F_0$ has the effect that the “occluder" effect is constant for all values of all F_0.

> *(7) The choice of the phrase 'hemodynamic occlusion' may cause some confusion as the authors address both positive and negative responses in the GFP expressing neurons, and there may be additional contributions from changes in blood oxygenation state.*

Regarding the potential confusion with regards to terminology, occlusion can decrease or increase.

Only under the (incorrect) assumption that occlusion is zero at baseline would this be confusing – no? If the reviewer has a suggestion for a different term, we'd be open to changing it.

Regarding blood oxygenation – this is absolutely correct, we did not explicitly point this out in the previous version of the manuscript. Occlusion changes are driven by a combination of changes to volume and "opacity" of the blood. Oxygenation changes would be in the second category. We have clarified this in the manuscript.

> *(8) The choice of ACC as the frontal region provides a substantial contrast in location, brain movement, and vascular architecture as compared to V1. As the authors note, ACC is close to the superior sagittal sinus and thus is the region where the largest vascular effects are likely to occur. The reader is left to wonder how much of the ROI may or may not have included vasculature in the ACC vs V1 recordings as the only images of the recording sites provided are for V1. We are left unable to conclude whether the differences observed between these regions are due to the presence of visible vasculature, capillary blood flow or differences in neurovasculature coupling between regions. A less medial portion of M2 may have been a more appropriate comparison. At least, inclusion of more example imaging fields for ACC in the supplementary figures would be of value.*

Both the choice of V1 and ACC were simply driven by previous experiments we had already done in these areas with calcium indicators. And we agree, the relevant axis is likely distance from midline, not AP – i.e. RSC and ACC are likely more similar, and V1 and lateral M2 more similar. We have made this point explicitly in the manuscript and have added sample fields of view as Figure S1.

> *(9) In Figure 3, How do the proportions of responsive GFP neurons compare to GCaMP6f neurons?*

We have added the data for GCaMP responses.

> *(10) How is variance explained calculated in Figure 4? Is this from a linear model and R^2 value? Is this variance estimate for separate predictors by using single variable models? The methods should describe the construction of the model including the design matrix and how the model was fit and if and how cross validation was run.*

This is simply a linear model (i.e. R^2) – we have added this to the methods.

> *(11) Cortical depth is coarsely defined as L2/3 or L5, without numerical ranges in depth from pia.*

Layer 2/3 imaging was done at a depth of 100-250 μm from pia, and the same for layer 5 was 400-600 μm. This has been added to the methods.

> *Overall Assessment:*

This paper is an important contribution to our understanding of how hemodynamic artifacts may corrupt GRAB and calcium imaging, even in two-photon imaging modes. Certain useful control experiments, such as intrinsic optical imaging in the same paradigms, were not reported, nor were any hemodynamic correction methods investigated. Thus, this limits both mechanistic conclusions and the overall utility with respect to immediate applications by end users. Nevertheless, the paper is of significant importance to anyone conducting two-photon or widefield imaging with calcium and GRAB sensors and deserves the attention of the broader neuroscience and in-vivo imaging community.

Reviewer #3 (Public review):

In this study, the authors aimed to investigate if hemodynamic occlusion contributes to fluorescent signals measured with two-photon microscopy. For this, they image the activity-independent fluorophore GFP in 2 different cortical areas, at different cortical depths and in different behavioral conditions. They compare the evoked fluorescent signals with those obtained with calcium sensors and neuromodulator sensors and evaluate their relationship to vessel diameter as a readout of blood flow.

They find that GFP fluorescence transients are comparable to GCaMP6f stimuli-evoked signals in amplitude, although they are generally smaller. Yet, they are significant even at the single neuronal level. They show that GFP fluorescence transients resemble those measured with the dopamine sensor GRABDA1m and the serotonin sensor GRAB-5HT1.0 in amplitude an nature, suggesting that signals with these sensors are dominated by hemodynamic occlusion. Moreover, the authors perform similar experiments with wide-field microscopy which reveals the similarity between the two methods in generating the hemodynamic signals. Together the evidence presented calls for the development and use of high dynamic range sensors to avoid measuring signals that have another origin from the one intended to measure. In the meantime, the evidence highlights the need to control for those artifacts such as with the parallel use of activity independent fluorophores.

Strengths:

- Comprehensive study comparing different cortical regions in diverse behavioral settings in controlled conditions.

- Comparison to the state-of-the-art, i.e. what has been demonstrated with wide-field microscopy.

- Comparison to diverse activity-dependent sensors, including the widely used GCaMP.

Weaknesses:

(1) The kinetics of GCaMP is stereotypic. An analysis/comment on if and how the kinetics of the signals could be used to distinguish the hemodynamic occlusion artefacts from calcium signals would be useful.

We might be misunderstanding what the reviewer means by “the kinetics of GCaMP are stereotypic”. The kinetics are clearly stereotypic if one has isolated single action potential responses in a genetically identified cell type. But data recorded in vivo looks very different, see e.g. example traces in figure 1g of (Keller et al., 2012). And these are selected example traces, the average GCaMP trace looks perhaps more like the three example traces shown in Figure 2 (this is not surprising if the GCaMP signals one records in vivo are a superposition of calcium responses and hemodynamic occlusion). All quantification of kinetics relies on identifying “events”. We cannot identify events in any meaningful way for most of the data

(see e.g. examples 2 and 3 in Figure 2). The one feature we can reliably identify as differing between GCaMP and GFP responses is peak response amplitude (as quantified in Figure 2).

> *(2) Is it possible that motion is affecting the signals in a certain degree? This issue is not made clear.*

See also our response to reviewer 1 comment 1. In brief, we have added a quantification of motion artefacts as Figure S5, and argue that motion artefacts could only account for locomotion onset responses (there is no detectable brain motion to visual responses) and would predict a decrease in fluorescence (not an increase).

> *(3) The causal relationship with blood flow remains open. Hemodynamic occlusion seems a good candidate causing changes in GFP fluorescence, but this remains to be well addressed in further research.*

We agree – we have made this clearer in the manuscript.

> ***Recommendations for the authors:***
>
> ***Reviewer #1 (Recommendations for the authors):***
>
> *(1) Figure 2A shows three neurons with convincing GFP responses, with amplitudes often exceeding 100%. However, after seeing these data, I actually feel less convinced that these responses are related to hemodynamic occlusion. Blood vessel diameter changes by at most a few percent during behavior -- how could such small changes lead to >100% changes in GFP fluorescence?*
>
> *My guess is that these responses might instead be related to motion artifacts, particularly given the strong correlation between these responses and running speed (Figure 2A). One possible way to test this is by examining a pixelwise map of fluorescence changes (dF/F) during running vs. baseline. If hemodynamic effects are involved, one would likely see a shadow of the involved blood vessels in this map. Conversely, if motion artifacts are the primary factor, the map of dF/F should resemble the spatial gradients of the mean fluorescence image. Examining pixelwise maps of dF/F will likely provide insights regarding the nature of the GFP signals.*

The underlying assumption ("blood vessel diameter changes by at most a few percent") might be incorrect here. (Note also, relevant is likely the cross section, not diameter.) See Figure 4A1 and B1 for quantification of example blood vessel area changes - both example vessels change area by approximately 50%. Also note, example 1 in Figure 2 is an extreme example. The example was chosen to highlight that effects can be large. To try to illustrate that this is not typical however, we also show the distribution of all neurons in Figure 2B and mark all three example cells – example 1 is at the very tail of the distribution.

Regarding the analysis suggested, we have added examples of this for running onset to the manuscript (Figure S7). We have examples in which a blood vessel shadow is clearly visible. More typical however, is a general increase in fluorescence (on running onset) that we think is caused by blood vessels closer to the surface of the brain.

> *(2) Figure 3A shows strong GFP responses during running, while visuomotor mismatch elicit virtually no GFP-responsive neurons. This finding is puzzling, as visuomotor mismatch has been shown by the same group to activate L2/3 neurons more strongly than running (see Figure 3A, Keller et al., 2012, Neuron). Stronger neuronal activation should, in theory, result in more pronounced hemodynamic effects, and therefore, a higher proportion of GFP-responsive neurons. The absence of GFP responses during*

> *visuomotor mismatch raises questions about whether GFP signals are directly linked to hemodynamic occlusion.*
>
> *An alternative explanation is that the strong GFP responses observed during running could instead be driven by motion artifacts, e.g., those associated with the increased head or body movements during running onsets. Such artifacts could explain the observed GFP responses, rather than hemodynamic occlusion.*

This might be a misunderstanding. Mismatch responses are primarily observed in mismatch neurons. These are superficial L2/3 neurons (possibly the population that in higher mammals is L2 neurons). The fact that mismatch responses are primarily observed in this superficial population is likely the reason they were discovered using two-photon calcium imaging (which tends to have a bias towards superficial neurons as the image quality is best there), and seen in much fewer neurons when using electrophysiological techniques (Saleem et al., 2013) that are biased to deeper neurons. In response to Reviewer #2, we have now also added a quantification of the fraction of neurons responsive to these stimuli when using GCaMP (Figure 3D-F). The fraction of neurons responsive to visuomotor mismatch is smaller than those responsive on locomotion or to visual stimuli.

Thus, based on “average” responses across all cortical cell types (our L2/3 recordings here are as unbiased across all of L2/3 as possible) the response profiles (strong running onset and visual responses, and weak MM responses) are probably what one would expect in first approximation also in the blood vessel response profile. Complicating this is of course the fact that it is likely some cell type specific activity that contributes most to blood flow changes, not simply average neuronal activity.

See response to public review comment 1 for a discussion of alternative sources, including motion artefacts.

> *(3) Given the potential confound associated with brain motion, the authors might consider quantifying hemodynamic occlusion effects under more controlled conditions, such as in anesthetized animals, where brain movement is minimal. They could use drifting grating stimuli, which are known to produce wellcharacterized blood vessel and hemodynamic responses in V1. The effects of hemodynamic occlusion can then be quantified by imaging the fluorescence of an activity-independent marker. For maximal robustness, GFP should ideally be avoided, due to its known sensitivity to pH changes, as noted in the public review.*

Brain motion is negligible to visual stimuli in the awake mouse as well (Figure S5). This is likely the better control than anesthetized recordings – anesthesia has strong effects on blood pressure, heart rate, breathing, etc. all of which would introduce more confounds.

> *(4) Regardless of the precise mechanism driving the observed GFP response, these activity-independent signals must be accounted for in functional imaging experiments. This applies not only to experiments using small dynamic range sensors but also to those employing 'high dynamic range' sensors like GCaMP6, which, according to the authors, exhibit responses only ~2-fold greater than those of GFP.*
>
> *In this context, the extensive GFP imaging data are highly valuable, as they could serve as a benchmark for evaluating the effectiveness of correction methods. Ideally, effective correction methods should produce minimal responses when applied to GFP imaging data. With these data at hand, I strongly encourage the authors to explore potential correction methods, as such methods could have far-reaching impact on the field.*

As discussed above, we have tested a number of such correction approaches for both widefield and two-photon imaging and could never recover a response profile that resembles the GFP response. The “correction method” we have come to favor, is repeating experiments using GFP (i.e. what we have done here).

> *(5) Several correction approaches could be considered: for instance, the strong correlation between GFP responses and blood vessel diameter (as shown in Figure 4) could potentially be leveraged to predict and compensate for the activity-independent signals. Alternatively, expressing an activity-independent marker alongside the activity sensor in orthogonal spectral channels could enable simultaneous monitoring and correction of activity-independent signals. Finally, computational procedure to remove common fluctuations, measured from background or 'neuropil' regions (see, e.g., Kerlin et al., 2010, Neuron; Giovannucci et al., 2019, eLife), may help reduce the contamination in cellular ROIs. The authors could try some or all of these methods, and benchmark their effectiveness by assessing, e.g., the number of GFP responsive neurons after correction.*

Over the years we have tried many of these approaches. A correction using a second fluorophore of a different color likely fails because blood absorption is strongly wavelength dependent, making it challenging to calibrate the correction factor. Neuropil “correction” on GCaMP data, even with the best implementations, is just a common mode subtraction. The signal in the neuropil – as the name implies is just an average of many axons and dendrites in the vicinity – most of these processes are from nearby neurons making a neuropil response simply an average response of the neurons in some neighborhood. Adding the problem of hemodynamic responses (which on small scales will also influence nearby neurons and neuropil similarly) makes disentangling the two effects impossible (i.e. neuropil subtraction makes the problem worse, not better). However, just because we fail in implementing all of these methods, does not necessarily mean the method is faulty. Hence we have chosen not to comment on any such method, and simply provide the only mitigation strategy that works in our hands – record GFP responses.

> *(6) Given the potential usefulness of the GFP imaging data, I encourage the authors to share these data in a public repository to facilitate the development of correction methods.*

Certainly – all of our data are always published. In the early years of the lab on an FMI repository here https://data.fmi.ch/ - more recently now on Zenodo.

> *(7) As noted in the public review, several methodology details are missing. Most importantly, I could not find the description in the Methods section explaining how fluorescence signals from individual neurons were extracted from two-photon imaging data. The existing section on 'Extraction of neuronal activity' appears to cover only the wide-field analysis, with details about two-photon analysis seemingly absent.*

Please excuse the omission – this has all been added to the methods. In brief, to answer your questions:

> *Were regions of interest (ROIs) for individual cells identified manually or automatically?*

We use a mixture of manual and automatic methods for our two-photon data. Based on a median filtered (spatially) version of the mean fluorescence image, we used a threshold based selection of ROIs. This was then visually inspected and manually corrected where necessary such that ROIs were at least 250 pixels and only labelled clearly identifiable neurons.

> *Was fluorescence within each ROI calculated by averaging signals across pixels, or were signal de-mixing algorithms (e.g., PCA, ICA, or NMF) applied?*

We use the average fluorescence across pixels without any de-mixing algorithms here and in all our two-photon experiments. De-mixing algorithms can introduce a variety of artefacts.

> *Additionally, did the authors account for and correct the contribution of surrounding neuropil?*

No neuropil correction was applied. It would also be difficult to see how this would help. If the model of hemodynamic occlusion is correct, one would expect occlusion effects to change on the length scale of blood vessels (i.e. tens to hundreds of microns). Thus, the effect of occlusion on neuropil and cells should be the similar. Neuropil "correction" is always based on the idea of removing signals that are common to both neuropil and somata, thereby complicating the interpretation of the resulting signal even further.

> *Without these methodological details, it is difficult to accurately interpret the two-photon signals reported in the manuscript.*
>
> *(8) The rationale for using the average fluorescence of a ROI within the blood vessel as a proxy for blood vessel diameter is not entirely clear to me. The authors should provide a clearer justification for this approach in their revision.*

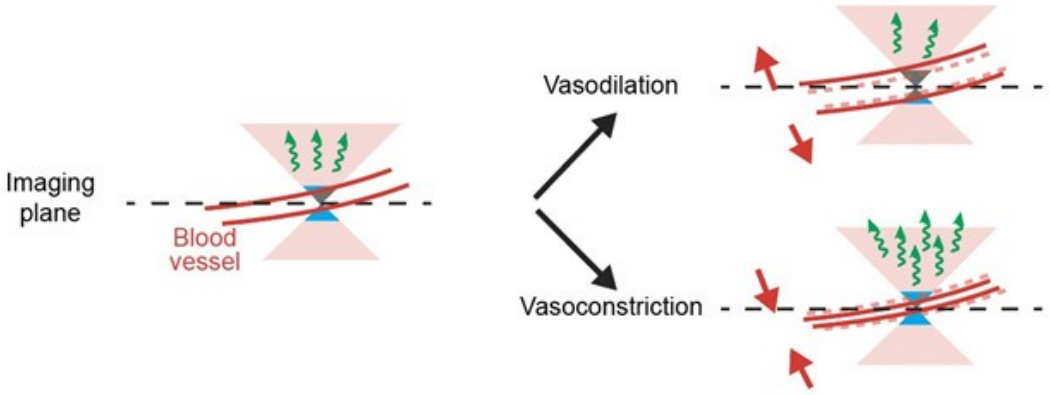
Consider a ROI placed within a blood vessel at the focus of the illumination cone (Author response image 3). Given the axial point-spread-function of two-photon imaging is in the range of 0.5 μm laterally and 3 μm axially (indicated by the bicone), emitted photons from the fluorescent tissue outside of the blood vessel but within the two-photon volume will contribute to change in fluorescence in the ROI. A change in the blood vessel volume, say an increase on dilation, would decrease the amount of emission photons reaching the objective by, one, pushing more of the fluorescent tissue outside of the two-photon volume, and two, by presenting greater hemodynamic occlusion to the photons emitted by the fluorescent tissue immediately below the vessel. Conversely, on vasoconstriction there are more emission photons at the objective.

In line with this argument, as shown in Figure 4A1-A2, B1-B2 and C1-C2, we do find that the change in fluorescence of blood vessel ROI varies inversely with the area of the blood vessel. Of course, change in blood vessel ROI fluorescence is only a proxy for vessel size. Extracting blood vessel boundaries from individual two-photon frames was noisy and proved unreliable in the absence of specific dyes to label the vessel walls. We thus resorted to using blood vessel ROI fluorescence as a proxy for hemodynamic occlusion, and tested how much of the variance in GFP responses is explained by the change in blood vessel ROI response.

We have added an explanation to the manuscript, as suggested.

Author response image 3.

Average response of ROIs placed within blood vessels co-vary with hemodynamic occlusion.



> *(9) I find that the Shen et al., 2012, Nature Methods paper has gone quite far to demonstrate the effect of hemodynamic occlusion in two photon imaging. Therefore, I suggest the authors describe and cite this work not only in the discussion but also in the introduction, where they can highlight the key questions left unanswered by that study and explain how their manuscript aims to address them.*

We have added the reference and point to the work in the introduction as suggested.

> ***Reviewer #3 (Recommendations for the authors):***
>
> *I appreciate very much that the study is presented in a very clear manner.*
>
> *A few comments that could clarify it even further:*
>
> *(1) Fig. 1: make clear on legend if it is an average of full FOVs.*

The traces shown are the average over ROIs (neurons) – we have clarified in the figure legend as suggested.

> *(2) Give a more complete definition of hemodynamic occlusion to understand the hypothesis in the relationship between blood vessel dilation and GFP fluorescence (116-119). Maybe, move the phrase from conclusion "Since blood absorbs light, hemodynamic occlusion can affect fluorescence intensity measurements" (219-220).*

Very good point – we expanded on the definition in the introduction.

> *(3) For clarity, mention in the main text the method used to assess how a parameter explains the variance (126-129).*

Is implemented.

> *(4) Discuss the possible relationship of the signals to neuronal activity.*

We have added this to the discussion.

> *(5) Discuss if the measurements could provide any functional insights, whether they could be used to learn something about the brain.*

We have added this to the discussion.

https://doi.org/10.7554/eLife.104914.2.sa0