Action enhances perception: Visually evoked neural responses are enhanced when engaging in a motor task

Jason J. Ki[1]**, Lucas C. Parra**[1]**, and Jacek P. Dmochowski**[1,✉]

[1]Department of Biomedical Engineering, City College of New York, New York NY 10031

While it is well known that vision guides movement, it is less appreciated that motor cortex also provides input to the visual system. Here we asked whether neural processing of visual stimuli is acutely modulated during motor activity, hypothesizing that visual evoked responses are enhanced when subjects engage in a motor task. To test this, we recorded neural activity from human participants during a car racing video game under 3 conditions: active play with manual control, passive viewing of game playback, and "sham play", where participants were under the false impression that their brain activity was controlling the game. This condition aimed to engage the motor system while avoiding evoked responses related to actual movement. In each case, we assessed the strength of evoked responses as the temporal correlation between the visual stimulus and the evoked electroencephalogram (EEG). We found reduced correlation during passive viewing, but no difference between active and sham play. Moreover, participants that were successfully deceived showed more correlated responses in the sham condition compared to those that were not deceived. Alpha band (8-12 Hz) activity was reduced over motor cortex during sham play, indicating recruitment of motor cortex despite the absence of overt movement. These findings are the first to demonstrate a link between visual evoked responses and motor cortex engagement.

visual perception | motor cortex | electroencephalography | stimulus-response correlation
Correspondence: *jdmochowski@ccny.cuny.edu*

Introduction

Visual processing in the brain has historically been delineated into two streams corresponding to the primary roles of vision: perception and action (1, 2). Whereas recognition of objects is believed to be processed in the ventral stream, signal transformations along the dorsal stream are believed to underlie action planning and spatial awareness, and occur largely in the parietal cortex (3, 4). Studies of visually-guided action have generally adopted a feedforward view, where the relevant information flows from the dorsal stream to the premotor and motor centers. On the other hand, much less attention has been devoted to potential influences of upstream regions, including the motor cortex itself, on visual processing.

Despite this, multiple lines of evidence indicate that the motor system exerts influence over visual processing. First, the visual and motor cortices have reciprocal anatomical connections in the primate brain (5–7). Moreover, numerous behavioral studies have demonstrated that learning of motor actions improves subsequent recognition of congruent visual stimuli (8–10), and that perceptual decisions may be primed by action (11–13). Human sensitivity to visual motion appears to be higher when that motion matches the observer's own movement patterns (14, 15). There is also evidence from neuroimaging studies that objects affording actions enhance early visual event-related potentials (ERPs) via a purported sensory gain mechanism (16–19). Neural recordings from visual extinction patients demonstrate that graspable objects bias visual perception in an unconscious manner (20). Based on these findings, we suspected that the presence of a motor task would *acutely* modulate visual processing. Specifically, we hypothesized that visual evoked responses are enhanced when accompanying engagement of the motor cortex. Testing this hypothesis in the human brain is not straightforward due to the fact that manual actions (e.g. button presses) introduce somatosensory and motor related signals into recordings of sensory activity, potentially confounding measures of neural visual responses, particularly because actions are often time-locked to changes in the stimulus.

Here we developed a "sham" motor task aimed at identifying the online effect of motor engagement on the dynamics of concurrent visual processing. We chose to record neural activity with the scalp electroencephalogram (EEG) to capture fast neural responses that could then be correlated with rapid stimulus fluctuations without requiring exogenous stimulus labels. Employing a within-subjects design, we asked subjects to engage in a video game under three distinct conditions: manually controlled gameplay ("active play"), viewing but under the false belief that brain activity was controlling game play – "sham play" – and passive viewing. We assessed the strength of visually evoked neural responses by measuring the correlation between a time-varying feature of the visual stimulus (optic flow) and the evoked EEG responses: the stimulus-response correlation (SRC) (21).

We found significantly elevated SRC during the active and sham play conditions compared to passive viewing. Importantly, no significant differences were observed between active and sham play. Subjects that reported being deceived exhibited significantly higher SRC than undeceived subjects during sham play. The enhancement of visual evoked responses during sham play was accompanied by a concomitant decrease in alpha band (8-12 Hz) power over the motor cortex, suggesting that the motor system was engaged despite the lack of overt movement. To our knowledge, this is the first

demonstration that motor cortex engagement is accompanied by enhanced visual evoked responses.

Results

We hypothesized that visual evoked responses are enhanced when engaging in a motor task. To test our hypothesis while ruling out activity associated with actual movement, we informed study participants that their brain activity will be controlling a car racing video game but instead presented them with playback of a previously recorded game ("sham play"). In other trials, subjects experienced the game under conventional manual control ("active play") or passively viewed game playback ("passive viewing"; Fig 1A). Our dependent measure was the temporal correlation between the time-varying optic flow of the video stream and the evoked brain response captured by the scalp EEG (Fig 1B). To account for the spatial diversity of the 96-channel EEG and varying response latencies, we captured multiple spatial components of the EEG and temporal components of the stimulus following Dmochowski et al. (21). We then summed the correlation measured in each component to arrive at the total stimulus-response correlation (SRC; Fig 1C). The regression approach outlined in Fig 1 implicitly models the EEG evoked responses as a sum of independent spatiotemporal components (21).

When applied to the present data, we obtained several neural response components evoked by optic-flow fluctuations (Fig 2). Notably, the strongest component was marked by a parietal topography centered at electrode CPz (centroparietal midline). The corresponding temporal response showed a positive peak at 200 ms (Fig 2A). The second strongest component exhibited negative poles over the medial frontal and medial occipital regions, and showed a late temporal response with a peak at 400 ms (Fig 2B). Components 3 and 4 showed mirror symmetric topographies with positive poles over occipital cortex and negative poles over frontocentral electrodes (Fig 2C-D). Thus, the optic flow stimulus evoked time-locked responses over broad regions of the cortex and a range of temporal delays, including late responses. Subsequent components exhibited weaker correlations with the stimulus and are not shown here, although the main outcome measure of the study, namely total SRC, encompassed all 11 components extracted from the data (see *Methods*).
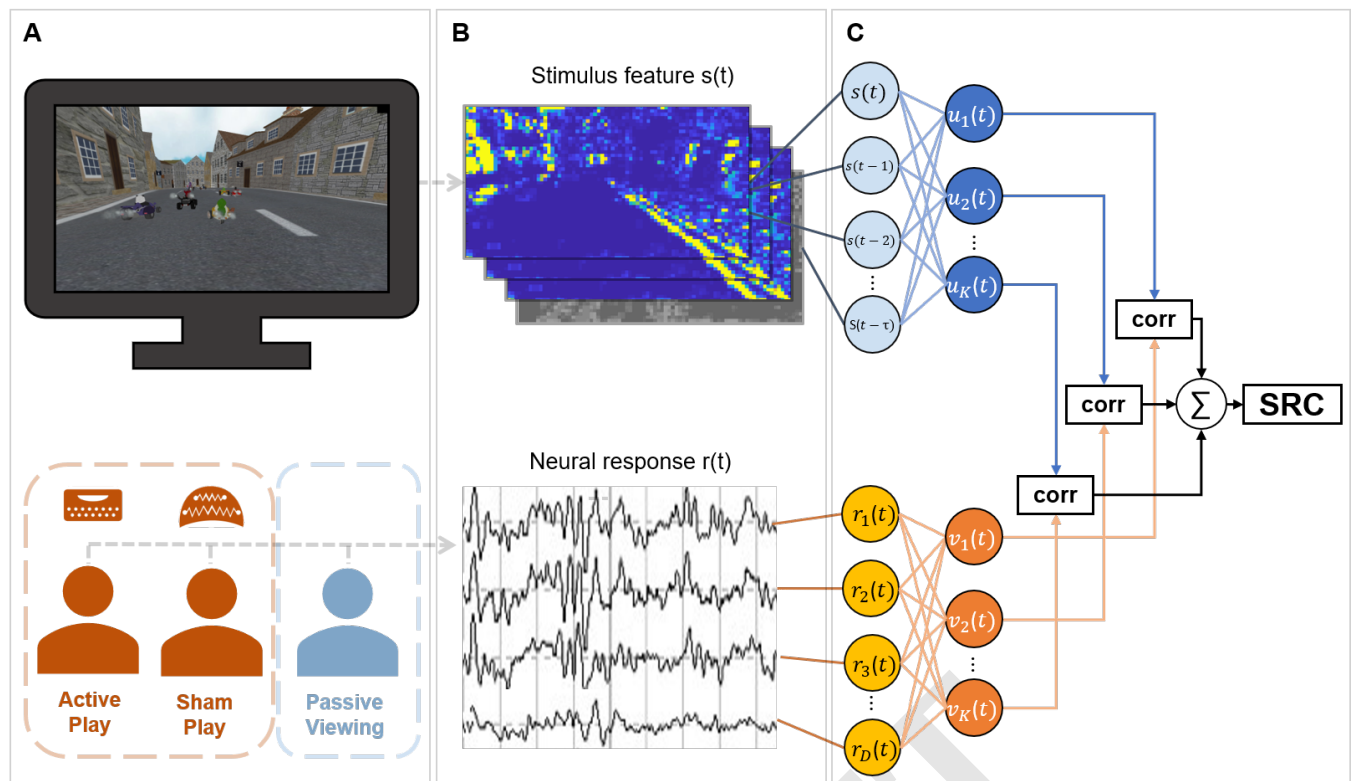
Enhanced visual evoked responses during active and sham play. We measured the total SRC separately for each experimental condition and found a significant reduction during passive viewing compared to both active play ($z = 2.31$, $p = 0.01$, paired, one-tailed Wilcoxon signed-rank test, $n = 18$ subjects; Fig 3A) and sham play ($z = 2.31$, $p = 0.01$; Fig 3A). No significant difference in total SRC was found between active and sham play ($z = -0.54$, $p = 0.58$, paired, two-tailed Wilcoxon signed-rank test, $n = 18$ subjects; Fig 3A). To compute the SRC, we employed the optic flow of the video stream because this particular feature drives the EEG stronger than other low-level visual or auditory features (21). However, similar results were obtained with temporal visual con-

trast (Fig S1), namely, a significant increase in SRC during sham play compared to passive viewing ($z = 2.61$, $p = 0.004$, paired one-tailed Wilcoxon signed rank test, $n = 18$), and a weaker effect of active play relative to passive viewing ($z = 1.48$, $p = 0.069$). We therefore continued our analysis with the optic flow feature.

Following the experiment, participants were asked to rate their engagement with the game in each condition. Analogous to the SRC measure, subjects reported higher engagement scores for active play ($z = 3.44$, $p = 2.92 \cdot 10^{-4}$, paired, one-tailed Wilcoxon signed-rank test, $n = 18$ subjects) and sham play ($z = 2.13$, $p = 0.016$) relative to passive viewing (Fig 3B). No significant difference in self-reported engagement was observed between active and sham play ($z = 1.17$, $p = 0.24$, paired, two-tailed Wilcoxon signed-rank test, $n = 18$ subjects; Fig 3B).

Enhancement is limited to the strongest response components of the EEG. The total SRC was computed as a sum of correlations across a total of 11 response components (see *Methods*). As each EEG component captured a distinct spatiotemporal response to the stimulus (Fig 2), we examined which of these components showed the effect of increased SRC during active or sham play relative to passive viewing (Fig 3C). We found significantly increased SRC in the sham play condition relative to passive viewing in parietal component 1 ($z = 2.53$, $p = 0.0058$, paired Wilcoxon signed rank test, $n = 18$; Fig 3C). In fronto-occipital component 2, a significant SRC increase in the active play condition was observed relative to passive viewing ($z = 2.04$, $p = 0.02$; Fig 3C). We also found significantly higher correlation in the active play condition relative to sham play in component 3 ($z = 2.09$, $p = 0.018$; Fig 3C). No significant effects were found in any of the remaining components (i.e., components 4-11). Thus, the increased SRC during active and sham play was confined to only the strongest response components.
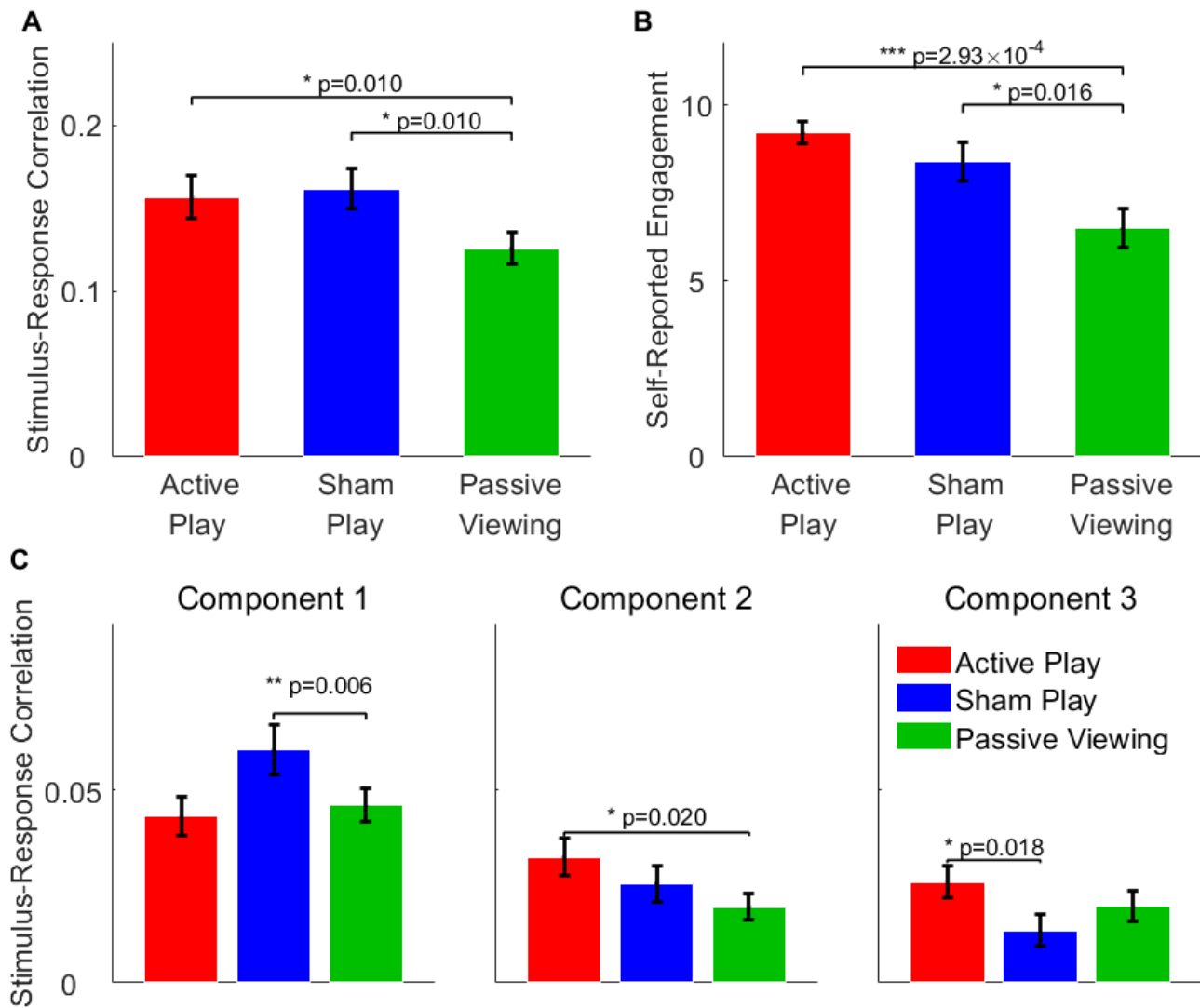
Enhanced visual evoked responses in deceived subjects. Based on post-experimental surveys, 13 of the 18 study participants reported being deceived by the sham play condition, with these subjects maintaining the perception of neural control throughout sham experimental trials. We therefore conducted a between-group analysis to test whether those deceived subjects exhibited significantly higher total SRC compared to the undeceived subgroup. Indeed, we observed a significant elevation of SRC in the deceived subgroup for the recordings made in the sham condition ($z = 1.97$, $p = 0.0243$, one-tailed Wilcoxon rank sum test, $n_{\mathrm{d}} = 13$, $n_{\mathrm{nd}} = 5$; Fig 4A). Notably, there were no significant group differences in total SRC during either active play ($z = 0.69$, $p = 0.25$; Fig 4A) or passive viewing ($z = 0.98$, $p = 0.46$; Fig 4A). Mirroring the SRC findings, deceived subjects reported significantly higher engagement with the video game during sham play ($z = 2.95$, $p = 0.002$, one-tailed Wilcoxon rank sum test $n_{\mathrm{d}} = 13$, $n_{\mathrm{nd}} = 5$; Fig 4B) but not during active play ($z = -2.02$, $p = 0.98$) or passive viewing ($z = 1.54$, $p = 0.06$).

Fig. 1. Measuring visual evoked responses with and without motor engagement. (**A**) Study participants experienced a car racing video game under 3 conditions: manual control ("active play"), viewing but under the false belief that brain activity was controlling game play ("sham play"), and knowingly viewing game playback ("passive viewing"). (**B**) Throughout the experiment, we recorded the video stream as well as the evoked scalp EEG. (**C**) Visual evoked responses were assessed by measuring the temporal correlation between the optic flow of the video stream and the time-locked neural response. To account for varying response latencies and multiple recording electrodes, we formed multiple spatial components of the EEG and temporal components of the stimulus. The sum of correlations across all components formed the dependent measure which we term here the stimulus-response correlation (SRC).
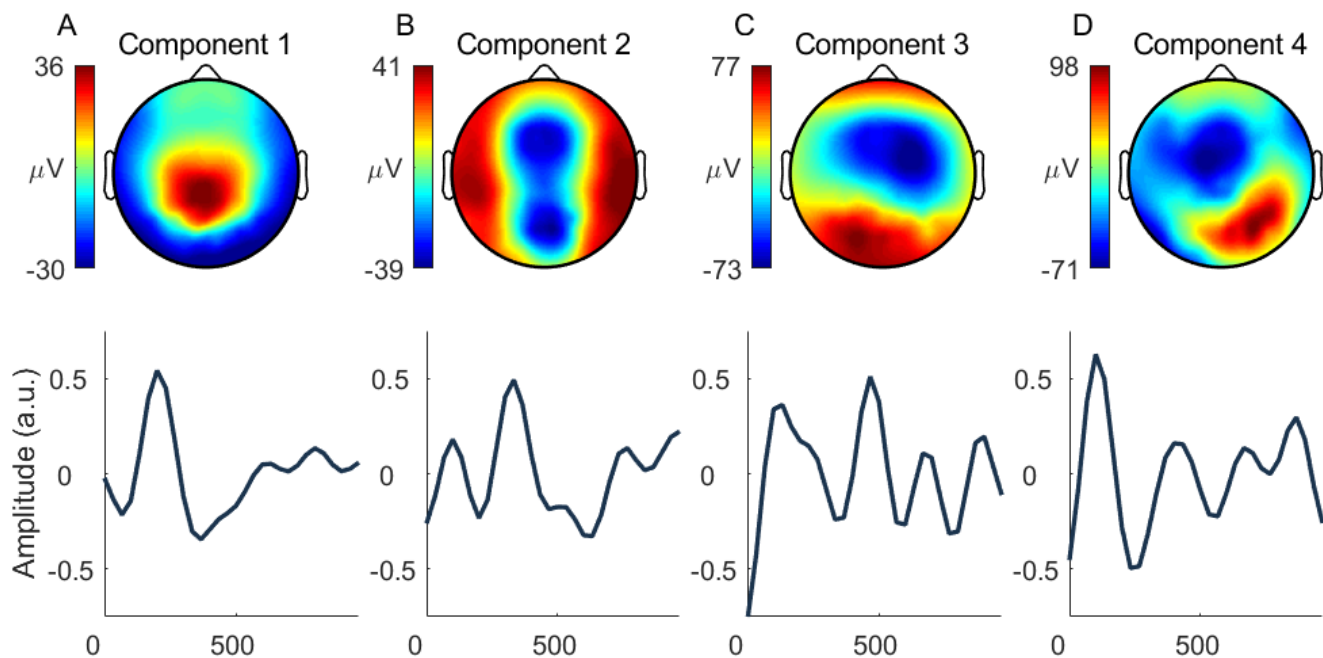
Fig. 2. Late and broad evoked responses to the visual stimulus. (**A**) The strongest neural response component was marked by a parietal topography centered at electrode CPz (centroparietal midline). The corresponding temporal response showed a positive peak at 200 ms. (**B**) The second strongest component exhibited negative poles over the medial frontal and occipital cortices, with a late temporal peak at 400 ms. (**C**) Component 3 was marked by poles over the right frontocentral and left occipital regions, and responded with a peak at 500 ms. (**D**) The topography of component 4 was mirror-symmetric to that of component 3, but the component showed an earlier temporal response. Collectively, the set of evoked response components indicate that the visual stimulus drove neural activity over broad scalp regions and a range of time delays, included late responses.
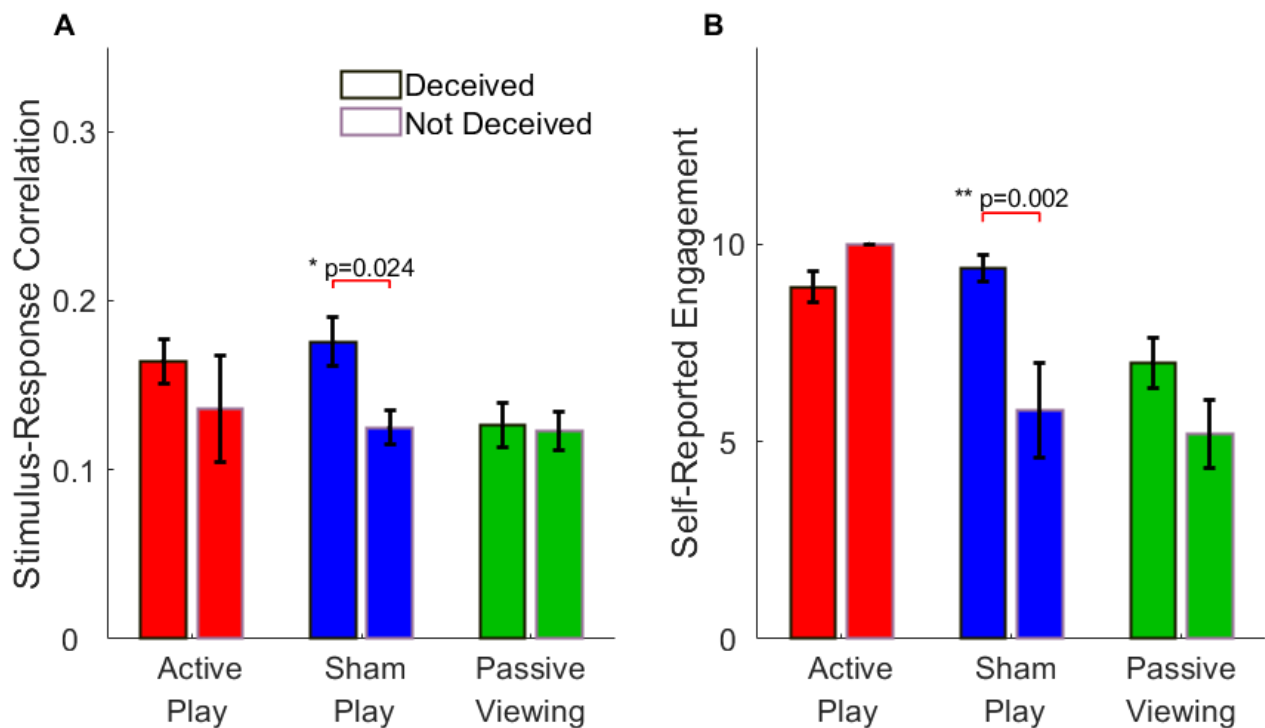
Fig. 3. Enhanced visual evoked responses during active and sham play. (**A**) The total SRC was measured separately for each condition (bars depict mean $\pm$ sem across $n = 18$ subjects). Passive viewing elicited significantly lower total SRC compared to active play ($p = 0.01$, $n = 18$, paired one-tailed Wilcoxon signed rank test) and compared to sham play ($p = 0.01$, $n = 18$, paired one-tailed Wilcoxon signed rank test). No significant difference was found between active and sham play ($p = 0.58$, $n = 18$, paired, two-tailed Wilcoxon signed rank test). (**B**) Subjects reported significantly higher engagement during active play ($p = 5 \times 10^{-4}$, $n = 18$, paired one-tailed Wilcoxon signed rank test) and sham play ($p = 0.031$, $n = 18$, paired one-tailed Wilcoxon signed rank test) relative to passive viewing. No significant difference in self-reported engagement was found between active and sham play ($p = 0.24$, $n = 18$, paired, two-tailed Wilcoxon signed rank test). (**C**) Differences in SRC between conditions were confined to the strongest response components: increased SRC during sham play relative to passive viewing was found in component 1 ($p = 0.0058$). Active play increased SRC compared to passive viewing at component 2 ($p = 0.02$). We also found an increase of SRC during active play relative to sham play in component 3 ($p = 0.018$). No significant effects were observed in subsequent components.

Alpha activity over motor cortex indicates motor engagement during sham play. By design, there were no overt differences in behavior between sham play and passive viewing – in both conditions, participants viewed the stimulus without manual actions. This prevents confounds due to motor or somatosensory evoked responses that could have been present during active play. To test whether the sham condition nevertheless engaged motor cortex, we measured the power of alpha-band (8-12 Hz) oscillations for each condition (Fig 5A-C). Reductions in alpha activity have long been observed over the motor cortex ("mu" rhythm) when subjects perform or visualize motor actions (22). Indeed, we found a significant reduction in alpha activity during both active and sham play relative to passive viewing, with the largest differences ob-

served over the left motor cortex (paired two-tailed Wilcoxon signed rank test, $n = 18$, corrected for multiple comparisons with FDR at 0.05; Fig 5D,E). Note that during active play, subjects controlled the game with their right hand. Both active and sham play led to alpha power reductions across broad scalp regions, with no significant differences between the two conditions (all $p > 0.05$, paired two-tailed Wilcoxon signed rank test, $n = 18$, corrected for multiple comparisons with FDR at 0.05; Fig 5F). This provides evidence that the motor system was indeed engaged during sham play.

Reduced power during active and sham play. Note that the SRC quantifies the strength of neural responses divided by their standard deviation, thus capturing the effect size of the stimulus on the EEG evoked response. The observed

Fig. 4. Enhanced visual evoked responses in deceived subjects. (**A**) Of the 18 study participants, $n_d = 13$ perceived neural control throughout the sham play trials while $n_{nd} = 5$ were not deceived. Computing SRC separately within each group, we found enhanced visual evoked responses in the deceived subjects during sham play ($p = 0.024$, $n_d = 13$, $n_{nd} = 5$, one-tailed Wilcoxon rank sum test). No significant difference was found between deceived and non-deceived subjects during active play or passive viewing (both $p > 0.05$, $n_d = 13$, $n_{nd} = 5$, one-tailed Wilcoxon rank sum test). (**B**) Deceived subjects reported significantly higher engagement during sham play ($p = 0.0026$, $n_d = 13$, $n_{nd} = 5$, one-tailed Wilcoxon rank sum test), but not during active play or passive viewing (both $p > 0.05$, $n_d = 13$, $n_{nd} = 5$, one-tailed Wilcoxon rank sum test).

increase of SRC could thus result from increased stimulus-related evoked responses, or alternatively, from a reduction in stimulus-unrelated neural activity, i.e., lower noise power. To disambiguate these two scenarios, we computed the total power spanned by all 11 EEG components. We found significantly *reduced* EEG power during both active and sham play relative to passive viewing (Fig 6A; active play: $z = 3.07, p = 0.0021$, sham play: $z = 2.55, p = 0.011$; paired two-tailed Wilcoxon signed rank test, $n = 18$) but no difference between active and sham play ($z = 0.55, p = 0.59$). The reduction in power during sham play was observed at each of the first three response components (Fig 6B-D).

We found the same result when limiting the power measurement to the alpha band: total alpha power was significantly higher in the passive condition (active play vs passive viewing: $z = 3.2$, $p = 0.0014$; sham play vs passive viewing: $z = 2.98$, $p = 0.0029$; Fig 6E). Each of the first three response components showed the effect (Fig 6F-H). These results, computed for the EEG components that were correlated with the stimulus, are consistent with what was observed in the native electrode space (Fig 5). It is nevertheless surprising that passive viewing had more and not less fluctuations in the ongoing EEG activity.

Differences in evoked responses are broadly distributed and include late responses. We argued that the increased SRC during sham play may have resulted from reduced variance of the stimulus-unrelated neural activity. To determine if the stimulus-evoked responses also differed between conditions,

we compared their spatial topographies and time courses, focusing on the three strongest components. To allow the spatial comparisons, we regressed the EEG of each condition onto the first three stimulus components (see *Spatial and temporal differences in neural response* in *Methods* for details). Similarly, we regressed the stimulus of each condition onto the first three response components to probe temporal differences across conditions. Due to the presence of motor actions during active play, we focused on differences between sham play and passive viewing (Fig 7). The spatial and temporal responses of all 3 conditions are shown in Fig S2.

The spatial and temporal patterns of evoked responses were largely consistent for sham play and passive viewing, with the exception of the spatial pattern of the second component (Fig 7A-C,G-I). Despite the stability in the pattern of responses, we observed significant differences in the *magnitude* of both spatial and temporal responses. In particular, the spatial response was stronger during sham play relative to passive viewing for the first component, with a more negative response over parietal cortex (Fig 7D). Moreover, the evoked responses measured during sham play were larger between 400 and 800 ms in components 1 and 2 (Fig 7G, H). This suggests that motor engagement may amplify late visual evoked responses that were driven by the dynamic visual stimulus.

Discussion

Our study provides human evidence suggesting that engagement of the motor cortex is associated with an online en-
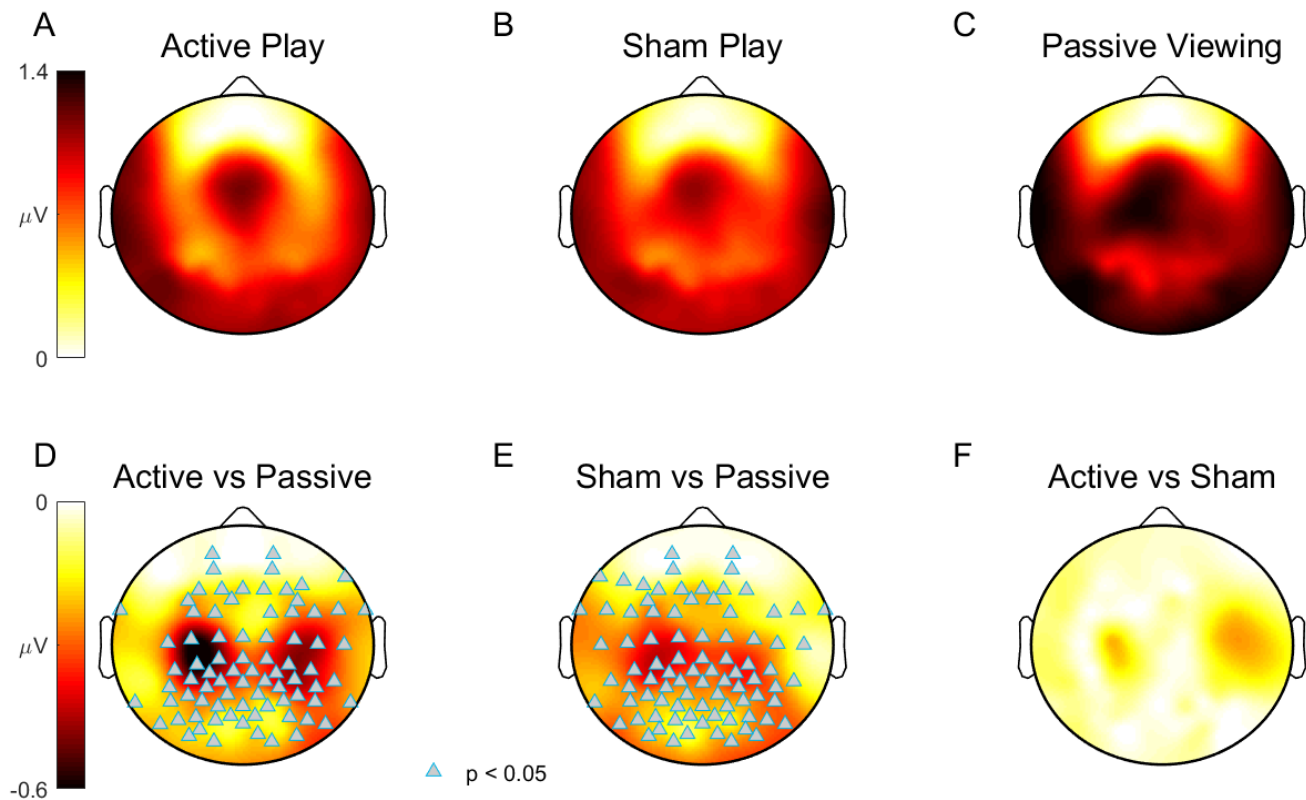
Fig. 5. Alpha activity over motor cortex is reduced during active and sham play. The power of the EEG in the alpha band (8-12 Hz) across the scalp, shown separately for (**A**) active play, (**B**) sham play, and (**C**) passive viewing. (**D**) A significant reduction in alpha power was found across broad regions of the scalp, most notably the left central electrodes, during active play compared to passive viewing ($p < 0.05$, one-tailed, paired Wilcoxon signed rank test, $n = 18$). (**E**) Similarly, sham play led to decreased alpha activity over much of the scalp and especially the left central region. (**F**) No significant differences in alpha were detected between active and sham play. The finding of reduced alpha activity over the left motor cortex during sham play indicates that, despite the lack of overt motor actions, the motor cortex was engaged when subjects believed that they were controlling game play.

hancement of visual evoked responses. This finding was facilitated by a sham in which participants believed that their brain activity was controlling a video game when in fact they were viewing a recording. By mitigating potential confounds from movement and somatosensation, we showed that visual evoked responses were significantly more correlated with the stimulus during this "sham play" condition compared to passive viewing. The increased fidelity was observed despite a lower overall EEG power during sham play, and was partly due to an increased magnitude of the neural response evoked by the visual stimulus. A reduction in alpha-band activity over left central electrodes was observed during sham play, indicating that the motor cortex was indeed engaged despite the lack of overt actions. Thus, we concluded that motor engagement is associated with an enhancement of visually evoked responses.

The SRC approach taken here allowed us to measure continuous visually evoked responses during a sensorimotor task that more closely mimics real-world settings than conventional event-related designs that employ discrete stimuli. Moreover, we were able to capture several components of the neural response to the optic flow stimulus (Fig 2), with the components expressed differently among the experimental conditions (Fig 3C). Note that in our framework, the analogues of the classical visual ERP are the temporal responses shown in

the bottom panels of Fig 2. These temporal responses indicate the brain's response to an impulse of optic flow. Spatially, these responses are obtained by linearly combining the activity of multiple electrodes as reflected in the top panel of Fig 2. Note also that while optic flow is a low-level feature of the visual stimulus, the neural response that it evokes may be modulated by complex brain states such as anticipation, surprise, fear, or arousal. Thus, the neural activity that was measured here captured potentially more than conventional visual ERPs. For example, the effects of an engaged motor cortex were to enhance late responses over central and parietal cortex. While not "visual" in the conventional sense, these evoked responses were nonetheless driven by the dynamics of the visual stimulus.

Future studies will be needed to identify the mechanism producing the enhancement of visually evoked responses found here during motor activity. It is possible that either top-down or bottom-up attentional modulation may have contributed to the enhancement. The introduction of a manual or sham task may have modified the brain state of the subjects prior to each trial – such an effect may manifest as lower overall noise power, including alpha activity, relative to passive viewing (Fig 6). On the other hand, a top-down influence is less likely to account for the observed increases in evoked response magnitude (Fig 7).

Ki *et al.* | Action enhances perception
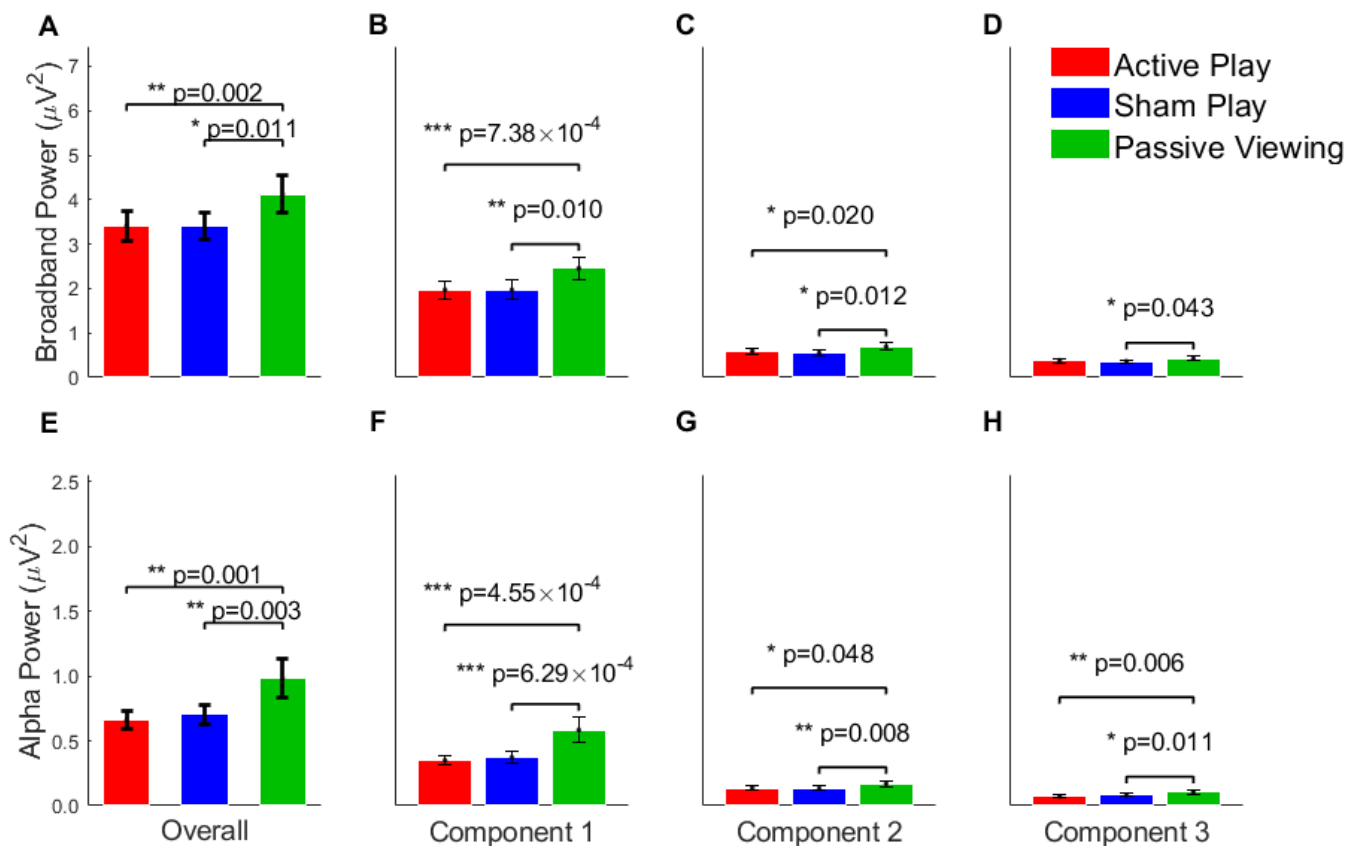
Fig. 6. Reduced EEG power during active and sham play. (**A**) Despite increased correlation with the stimulus, the total power of the measured EEG was significantly lower during active and sham play relative to passive viewing (active versus passive: $p = 0.0021$, sham versus passive: $p = 0.011$, $n = 18$, paired two-tailed Wilcoxon signed rank test). (**B-D**) The decrease in EEG power during sham (active) play was resolved in each of the first three (two) components. (**E**) The reduction in EEG power during active and sham play was also found when limiting the comparison to the alpha band (active versus passive: $p = 0.0014$, sham versus passive; $p = 0.0029$, $n = 18$, paired two-tailed Wilcoxon signed rank test). (**F-H**) Reduced alpha was found in each of the first three components for both active and sham play relative to passive viewing.

Visual stimuli containing objects that afford actions have been shown to increase visual spatial attention and amplify evoked responses, but only when the premotor and prefrontal cortices are activated (16, 23). This implies connectivity between premotor and prefrontal regions and the visual cortex, which has been shown anatomically in the primate brain (5–7). Here, the presence of the race car on the screen may have similarly amplified the evoked response to the optic flow stimulus. Note, however, that the modulation of visual responses required an active motor plan, in that the same actionable stimulus did not enhance visual responses during passive viewing.

Observing motor actions has been shown to generate imitative motor plans in the observer (24), but the role of these motor plans has been debated (25). One account is that the function of this motor activation is to generate a prediction of future perceptual input, thus bypassing the delays of sensory processing (26). During active and sham game play, study participants may have formed a prediction of the evolving optic flow stimulus, consistent with increased stimulus-driven activity over the central cortex (Fig 7A). This interpretation of the results is consistent with the theory that perceived events and planned actions share a common representational domain (27).

In general, active and sham play may have exhibited

stimulus-driven neural activity along a broader portion of the brain. For example, it is possible that the optic flow stimulus evoked correlated activity in dorsal regions downstream from striate visual cortex, such as the parietal or premotor cortices. The strongest modulation of the evoked response, as well as alpha power, was seen over the parietal and central cortices. The first component was expressed over these areas (Fig 2) and showed significantly higher SRC when participants believed that they were controlling the game (Fig 3). The posterior parietal cortex (PPC) has been shown to code motor intentions in the macaque (28), and it is tempting to speculate that a PPC-like component tracked the visual stimulus more faithfully in the sham play condition compared to the passive state. However, a limitation of our study is the poor spatial resolution of the scalp EEG, and the associated difficulties in recovering cortical sources from observed scalp topographies. The ill-posed nature of the EEG inverse problem is exacerbated when averaging scalp topographies over multiple subjects, as was implicitly done here. A natural extension of this work is thus to replicate the experiment with functional magnetic resonance imaging (fMRI) to glean insight into the brain areas driving the enhancement of visual evoked responses. However, note that the high temporal resolution of the EEG allowed us to measure fast evoked responses to the dynamic stimulus, which may not be feasible with fMRI
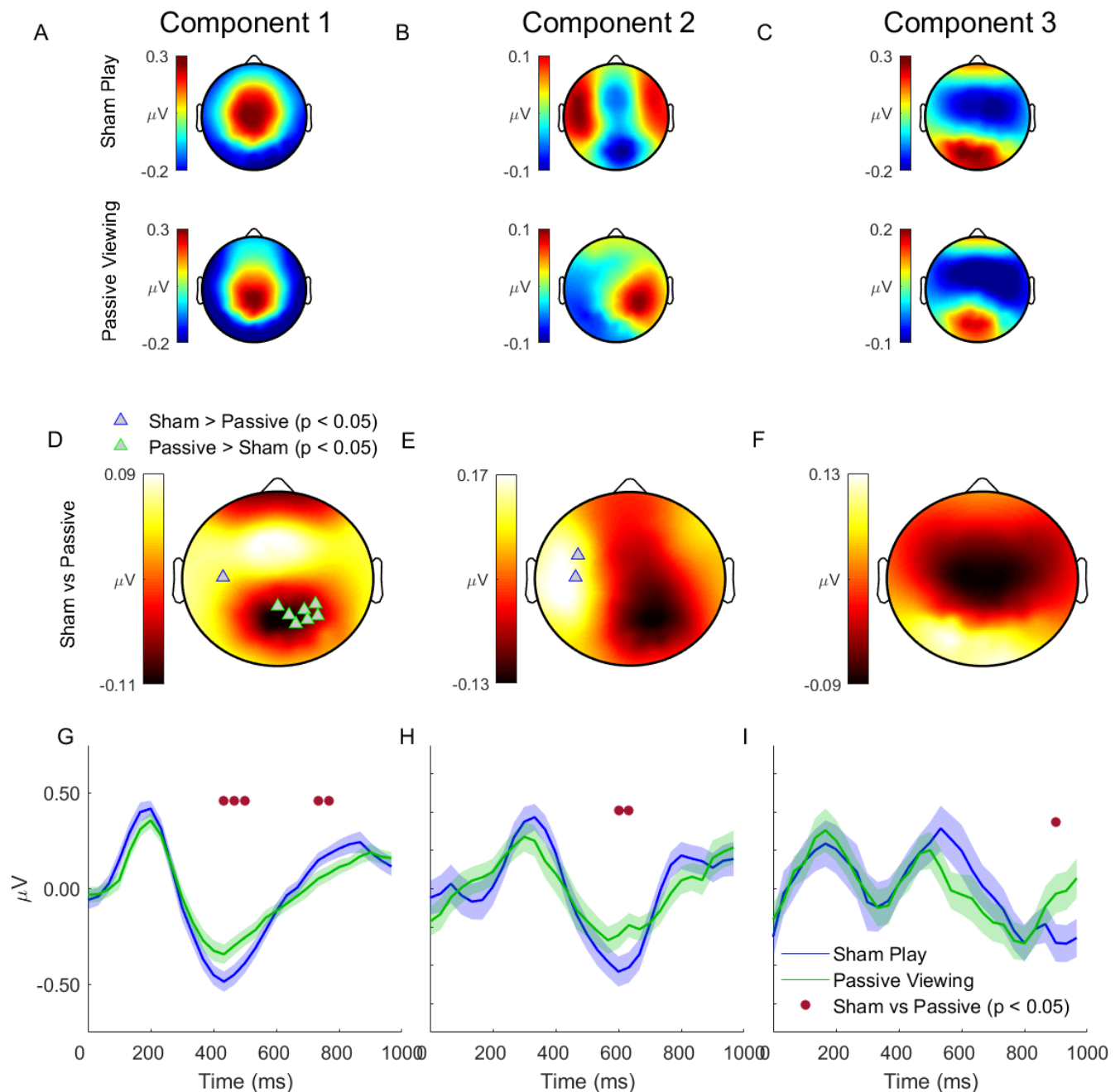
Fig. 7. Evoked responses differ spatiotemporally between conditions. (**A**-**C**) To test whether sham play and passive viewing produced different response patterns on the scalp, we computed separate topographies for the first three response components of each condition. The spatial pattern of evoked responses was consistent across the two conditions, with the exception of component 2. (**D**-**F**) However, we found significant magnitude differences, with stronger responses differences over parietal cortex during sham play in component 1. Sham play also produced stronger activation over left central cortex in component 2. (**G**-**I**) Separate temporal responses for sham play and passive viewing were also computed. The temporal response of component 1 exhibited significantly larger magnitude between 400 ms and 800 ms during sham play compared to passive viewing. Sham play also produced larger temporal responses in the second component at 600 ms. These results indicate that sham play elicited stronger evoked responses compared to passive viewing.

due to the slowness of the hemodynamic response to neural activation.

Regardless of the neural mechanism underlying the enhancement of visual processing during game play, our results provide an avenue for decoding active versus passive vision from non-invasive measurements of neural activity. By measuring the correlation between an evoked neural response and a time-varying visual stimulus, one can extract an estimate of how active the viewer is. While here we measured SRC at the scale of a 3-minute trial, it can also be computed in finer time increments and tracked continuously. We speculate that there is a continuum between passive viewing and active control, and that the SRC can place the subjects on this continuum on a moment-to-moment basis. In the future, wearable devices may be equipped with various sensors for capturing environmental stimuli in real-time (e.g. microphones and video cameras). Given the development of unobtrusive techniques for non-invasive sensing of neural activity (29), such as that from inside the ear canal (30), the SRC represents a natural technique for gleaning information about individual brain state in real-time. For example, it may be possible to decode spatial attention (31) by computing the SRC separately for multiple areas of the visual field or directions of incoming sound. There is already evidence that SRC can be used to determine speech comprehension in the context of hearing aids (32). An advantage of the SRC approach is that it is unsupervised, in that no learning procedure is required to, for example, learn patterns of neural activity that distinguish active from passive viewing.

Finally, an interesting facet of this work is that we were able to deceive a substantial majority of our participants. 13 of 18 participants completed the experiment with the belief that their brain activity was controlling game play during trials in which they actually viewed prerecorded stimuli. It is likely that the car racing video game employed in our study elicited robust stereotyped manual (and imagined) responses across subjects, thus contributing to the efficacy of deception. It is also notable that the sham play condition elicited strong neural activity over the parietal cortex (33), a region associated with visually guided movement planning and control. This suggests that that such visuomotor pathways may be activated with only the perception of control. Aside of being an interesting psychological finding, this opens up new experimental paradigms for probing the brain under active scenarios.

Methods and Materials

Participants. All participants provided written informed consent in accordance with procedures approved by the Institutional Review Board of the City University of New York. 18 healthy human subjects (9 females) aged 20 ± 1.56 participated in the study.

Video game stimulus. We employed the open-source car racing game *SuperTuxKart* (http://supertuxkart.net), in which participants navigate vehicles around a track against simulated opponents. All experimental trials were

conducted on the default course and spanned three laps in "easy" mode. The average trial had a duration of 175.9 ± 5.51 seconds. To simplify the task, we removed several graphical items from the stimulus (e.g. time display) such that the video stimulus consisted of only the race car, track, and opponents. To generate the stimuli employed during the sham play and passive viewing conditions, we recorded several races for subsequent playback during the experiments. A non-participant played 4 races, with 2 serving as stimuli during the sham play condition and the other 2 employed during the passive viewing condition. With the exception of active play, which produces unique stimuli during each trial, all subjects experienced the same stimuli.

The stimulus was presented on a high-definition Dell 24-inch UltraSharp Monitor (1920-by-1080 pixels) at a frame rate of 60 Hz. Subjects viewed the stimulus in a dark room at a viewing distance of 60 cm. The game's sound was muted during the experiment. The video frame sequence of each race was captured with the open-source *Open Broadcaster Software* (https://obsproject.com/) at the native resolution and frame rate. In order to subsequently synchronize the video frame sequence with the recordings of the EEG, a 30-by-30 pixel was flashed in the top right corner of the display throughout each trial. An electrical pulse produced by a photodiode placed over the top right corner was transmitted with low latency to the EEG recorder.

Experimental procedures. Study participants experienced two trials of the video game stimulus in each of 3 conditions: "active play", "sham play", and "passive viewing". The ordering of the conditions was randomized and counterbalanced across subjects. Subjects were permitted one practice trial of the video game prior to commencing the experiment. During active play, subjects controlled the game via keyboard presses made with the right hand: the left and right keys controlled steering, while the up and down keys produced acceleration and braking, respectively. Prior to the first sham play trial, subjects were falsely told that their brain activity will be controlling the video game, and that they should imagine the intended command. Moreover, we primed the participants by implementing a mock calibration of a brain computer interface" prior to the playback of the sham play races during which subjects were asked to imagine game controls (accelerate, brake, steer left, steer right). During passive viewing trials, subjects were instructed to freely view playback of a previously recorded game. Upon completion of the experiment, participants filled out a survey reporting their experienced "engagement" during each condition. Scores ranged from 1 ("not engaged") to 10 ("fully engaged"). Following the survey, subjects were informed of the deception task, and were asked whether they had become aware of the fact that their brain activity was not controlling game play. Of the 18 study participants, 13 reported being deceived for the entirety of the experiment.

EEG acquisition and preprocessing The scalp electroencephalogram (EEG) was acquired with a 96-electrode cap (custom montage with dense coverage of the occipital region)

housing active electrodes connected to a Brain Products ActiChamp system and Brain Products DC Amplifier (Brain Vision GmbH, Munich, Germany). The EEG was sampled at 500 Hz and digitized with 24 bits per sample. The EEG was transmitted to a recording computer via the Lab Streaming Layer software (34) which ensured precise temporal alignment between the EEG and video frame sequence.

EEG data was imported into the Matlab software (Mathworks, Natick, MA) and analyzed with custom scripts. Data was downsampled to 30 Hz in accordance with the Nyquist rate afforded by the 60 Hz frame rate, followed by high-pass filtering at 1 Hz to remove slow drifts. To remove gross artifacts from the data, we employed the robust PCA technique (35), which provides a low-rank approximation to the data and thereby removes sparse noise from the recordings. Note that due to volume conduction, sparse EEG components are generally artifacts. We employed the robust PCA implementation of (36) with the default hyperparameter of $\lambda = 0.5$. To reduce the contamination of EEG from eye movements, we linearly regressed out the activity of four "virtual" electrodes constructed via summation or subtraction of appropriately selected frontal electrodes. These virtual electrodes were formed to strongly capture the activity produced by eye blinks and saccades. To further denoise the EEG, we rejected electrodes whose mean power exceeded the mean of all channel powers by four standard deviations. Within each channel, we also rejected time samples (and its adjacent sample) whose amplitude exceeded the mean sample amplitude by four standard deviations. We repeated the channel and sample rejection procedures over three iterations.

Stimulus feature extraction Video frames were downsampled to a resolution of 320-by-180 pixels to reduce data size, and then converted to grayscale images. Optical flow was computed with the Horn-Schunk method as implemented in the MATLAB Computer Vision System Toolbox (37). For each frame, we computed the mean (across pixels) of the magnitude of the optical flow vector. Temporal contrast was constructed by taking the mean (across pixels) of the framewise derivative of the video sequence (21). The resulting time series were z-scored prior to SRC analysis.

Stimulus-response correlation. To measure the correlation between the time-varying stimulus feature $s(t)$ and the D dimensional evoked neural response $r_i(t), i \in 1, 2, \ldots, D$, we employed the multidimensional SRC technique developed in (21). The approach consists of temporally filtering the stimulus:

$$u(t) = \sum_{\tau} h(\tau)s(t - \tau) \qquad (1)$$

and spatially filtering the neural response:

$$v(t) = \sum_{i} w_i r_i(t) \qquad (2)$$

to produce stimulus component $u(t)$ and response component $v(t)$ that exhibit maximal mutual correlation:

$$\mathbf{h}^*, \mathbf{w}^* = \arg\max_{\mathbf{h}, \mathbf{w}} \rho_{uv} \qquad (3)$$

where $\mathbf{h}^* = \begin{bmatrix} h(1) & \ldots & h(L) \end{bmatrix}^T$ are the optimal temporal filter coefficients of the L-length filter and $\mathbf{w}^* = \begin{bmatrix} w_1 & \ldots & w_D \end{bmatrix}^T$ are the optimal spatial filter coefficients, and where ρ_{uv} is the Pearson correlation coefficient between $u(t)$ and $v(t)$. The solution to (3) is given by Canonical Correlation Analysis (38) and consists of pairs of projection vectors $\{\mathbf{h_j^*}, \mathbf{w_j^*}\}_{j=1}^K$ that yield a set of maximally correlated components $u_j(t)$ and $v_j(t)$ with corresponding correlation coefficients that decrease in magnitude $\rho_{uv}^1 \geq \rho_{uv}^2 \geq \ldots \geq \rho_{uv}^K$. Note here that we regularized the CCA solution by truncating the eigenvalue spectrum of the EEG covariance matrix to $K = 11$ dimensions, as this value explained over 99% of the variance in the data. This regularization was implemented within the computation of the CCA filters via custom Matlab code. Encompassing all components, the total correlation between the stimulus and response is given by:

$$\mathrm{SRC} = \sum_{j=1}^K \rho_{uv}^j. \qquad (4)$$

When computing SRC, we employed leave one out cross-validation along the subject dimension. In other words, the SRC of a given subject was measured after performing CCA on all of the data excluding that subject, and then applying the resulting filters to the data of the subject. When displaying learned filters, however, we show the result of learning on all of the data as this is not dependent on which subject was left out (Figures 2, 7).

Alpha power analysis. To test for differences in alpha power between conditions (Fig 5), we temporally filtered the EEG response of each electrode $r_i(t), i = 1, \ldots, D$, to the alpha band (8-13 Hz) using a fourth order Butterworth filter. We then measured the alpha power at each electrode by computing the temporal mean square of the filter output. This procedure was repeated on the EEG components $v_j(t), j = 1, \ldots, K$, to test for alpha activity changes at the level of components (Fig 6E-H).

Statistical testing. We tested for conditional differences in SRC, self-reported engagement, and alpha power by conducting paired Wilcoxon signed-rank tests on sets of $n = 18$ samples in each condition, with each sample corresponding to a subject. When comparing SRC and self-reported engagement between active play (or sham play) and passive viewing, we performed right-tailed tests due to our hypothesis of increased SRC and engagement during active vision. Differences between active and sham play were probed with two-tailed tests due to a lack of any prior expectation. When testing for differences between deceived and non-deceived subjects, we performed right-tailed Wilcoxon rank sum tests due to the hypothesis of higher SRC and engagement for the deceived subjects.

Spatial and temporal differences in neural response. To assess conditional differences in the scalp topographies of the response components (Fig 7A-C, Fig S2), we spatially regressed the EEG of each condition onto the first three (filtered) stimulus components $u_j(t)$. Note that these filtered

stimulus components (Fig 2, bottom) were learned by pooling over the data of all three conditions. The regression produced a set of filter weights w_i that best mapped the EEG of each condition onto the previously learned stimulus components using ordinary least squares. Similarly, we examined temporal differences of the evoked responses across conditions by temporally regressing the stimulus of each condition onto the first three response components (Fig 2, top). This produced, for each condition, a temporal response $h(\tau)$ that indexes the neural response to an impulse of the stimulus (Fig 7G-I, Fig S2).

Forward models of response components. To display the scalp topographies of the response components, we transformed the spatial filter weights w_i, $i = 1, .., D$, to their corresponding "forward model" following $A = RW(W^T RW)^{-1}$, where R is the covariance matrix of the EEG, W is a matrix of spatial filter weights such that the element at row i and column j is the weight assigned to the jth electrode of component i, and the columns of matrix A denote the forward models of the corresponding spatial filters (39, 40). This representation depicts the scalp projection of each extracted component and takes into account the spatial correlation of the EEG. As a result, these topographies allow for visualizing where on the scalp the extracted activity is expressed.

Bibliography

1. Mortimer Mishkin, Leslie G Ungerleider, and Kathleen A Macko. Object vision and spatial vision: two cortical pathways. *Trends in neurosciences*, 6:414–417, 1983.
2. AD Milner and MA Goodale. Two visual systems re-viewed. *Neuropsychologia*, 2008.
3. Melvyn A Goodale and A David Milner. Separate visual pathways for perception and action. *Trends in neurosciences*, 15(1):20–25, 1992.
4. David Milner and Mel Goodale. *The visual brain in action*. Oxford University Press, 2006.
5. Malcolm P Young. The organization of neural systems in the primate cerebral cortex. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 252(1333):13–18, 1993.
6. Steven P Wise, Driss Boussaoud, Paul B Johnson, and Roberto Caminiti. Premotor and parietal cortex: corticocortical connectivity and combinatorial computations. *Annual review of neuroscience*, 20(1):25–42, 1997.
7. David C Van Essen. Corticocortical and thalamocortical information flow in the primate visual system. *Progress in brain research*, 149:173–185, 2005.
8. Antonino Casile and Martin A Giese. Nonvisual motor training influences biological motion perception. *Current Biology*, 16(1):69–74, 2006.
9. Heiko Hecht, Stefan Vogt, and Wolfgang Prinz. Motor learning enhances perceptual judgment: A case for action-perception transfer. *Psychological research*, 65(1):3–14, 2001.
10. Annerose Engel, Michael Burke, Katja Fiehler, Siegfried Bien, and Frank Rösler. Motor learning affects visual movement perception. *European Journal of Neuroscience*, 27(9):2294–2302, 2008.
11. Andreas Wohlschläger. Visual motion priming by invisible actions. *Vision research*, 40(8):925–930, 2000.
12. Liana E Brown, Elizabeth T Wilson, Melvyn A Goodale, and Paul L Gribble. Motor force field learning influences visual processing of target motion. *Journal of Neuroscience*, 27(37):9975–9983, 2007.
13. Hannah Barbara Helbig, Jasmin Steinwender, Markus Graf, and Markus Kiefer. Action observation can prime visual object recognition. *Experimental brain research*, 200(3-4):251–258, 2010.
14. Fani Loula, Sapna Prasad, Kent Harber, and Maggie Shiffrar. Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance*, 31(1):210, 2005.
15. Günther Knoblich and Rüdiger Flach. Predicting the effects of actions: Interactions of perception and action. *Psychological science*, 12(6):467–472, 2001.
16. Todd C Handy, Scott T Grafton, Neha M Shroff, Sarah Ketay, and Michael S Gazzaniga. Graspable objects grab attention when the potential for action is recognized. *Nature neuroscience*, 6(4):421, 2003.
17. Maha Adamo and Susanne Ferber. A picture says more than a thousand words: Behavioural and erp evidence for attentional enhancements due to action affordances. *Neuropsychologia*, 47(6):1600–1608, 2009.
18. Agnieszka Wykowska and Anna Schubö. Action intentions modulate allocation of visual attention: electrophysiological evidence. *Frontiers in Psychology*, 3:379, 2012.
19. Heath Matheson, Aaron J Newman, Jason Satel, and Patricia McMullen. Handles of manipulable objects attract covert visual attention: Erp evidence. *Brain and cognition*, 86:17–23, 2014.
20. Giuseppe di Pellegrino, Robert Rafal, and Steven P Tipper. Implicitly evoked actions modulate visual selection: Evidence from parietal extinction. *Current Biology*, 15(16):1469–1472, 2005.
21. Jacek P Dmochowski, Jason J Ki, Paul DeGuzman, Paul Sajda, and Lucas C Parra. Extracting multidimensional stimulus-response correlations using hybrid encoding-decoding of neural activity. *NeuroImage*, 2017.
22. Jaime A Pineda. The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing". *Brain research reviews*, 50(1):57–68, 2005.
23. Glyn W Humphreys, Eun Young Yoon, Sanjay Kumar, Vaia Lestou, Keiko Kitadono, Katherine L Roberts, and M Jane Riddoch. The interaction of attention and action: From seeing action to acting on perception. *British Journal of Psychology*, 101(2):185–206, 2010.
24. Giacomo Rizzolatti and Laila Craighero. The mirror-neuron system. *Annu. Rev. Neurosci.*, 27:169–192, 2004.
25. Giacomo Rizzolatti, Luciano Fadiga, Vittorio Gallese, and Leonardo Fogassi. Premotor cortex and the recognition of motor actions. *Cognitive brain research*, 3(2):131–141, 1996.
26. Margaret Wilson and Günther Knoblich. The case for motor involvement in perceiving conspecifics. *Psychological bulletin*, 131(3):460, 2005.
27. Wolfgang Prinz. Perception and action planning. *European journal of cognitive psychology*, 9(2):129–154, 1997.
28. James W Gnadt and Richard A Andersen. Memory related motor planning activity in posterior parietal cortex of macaque. *Experimental brain research*, 70(1):216–220, 1988.
29. Alexander J Casson, David C Yates, Shelagh JM Smith, John S Duncan, and Esther Rodriguez-Villegas. Wearable electroencephalography. *IEEE engineering in medicine and biology magazine*, 29(3):44–56, 2010.
30. David Looney, Preben Kidmose, Cheolsoo Park, Michael Ungstrup, Mike Lind Rank, Karin Rosenkranz, and Danilo P Mandic. The in-the-ear recording concept: User-centered and wearable brain monitoring. *IEEE pulse*, 3(6):32–42, 2012.
31. Gi-Yeul Bae and Steven J Luck. Dissociable decoding of spatial attention and working memory from eeg oscillations and sustained potentials. *Journal of Neuroscience*, 38(2):409–422, 2018.
32. Ivan Vladimirov Iotzov and Lucas C Parra. Eeg can predict speech intelligibility. *Journal of Neural Engineering*, 2019.
33. Christopher A Buneo and Richard A Andersen. The posterior parietal cortex: sensorimotor interface for the planning and online control of visually guided movements. *Neuropsychologia*, 44(13):2594–2606, 2006.
34. C Kothe. Lab streaming layer (lsl). *https://github.com/sccn/labstreaminglayer. Accessed on October*, 26:2015, 2014.
35. Emmanuel J Candès, Xiaodong Li, Yi Ma, and John Wright. Robust principal component analysis? *Journal of the ACM (JACM)*, 58(3):11, 2011.
36. Zhouchen Lin, Minming Chen, and Yi Ma. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. *arXiv preprint arXiv:1009.5055*, 2010.
37. Berthold KP Horn and Brian G Schunck. Determining optical flow. *Artificial intelligence*, 17(1-3):185–203, 1981.
38. Harold Hotelling. Relations between two sets of variates. *Biometrika*, 28(3/4):321–377, 1936.
39. Stefan Haufe, Frank Meinecke, Kai Görgen, Sven Dähne, John-Dylan Haynes, Benjamin Blankertz, and Felix Bießmann. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage*, 87:96–110, 2014.
40. Lucas C Parra, Clay D Spence, Adam D Gerson, and Paul Sajda. Recipes for the linear analysis of eeg. *Neuroimage*, 28(2):326–341, 2005.
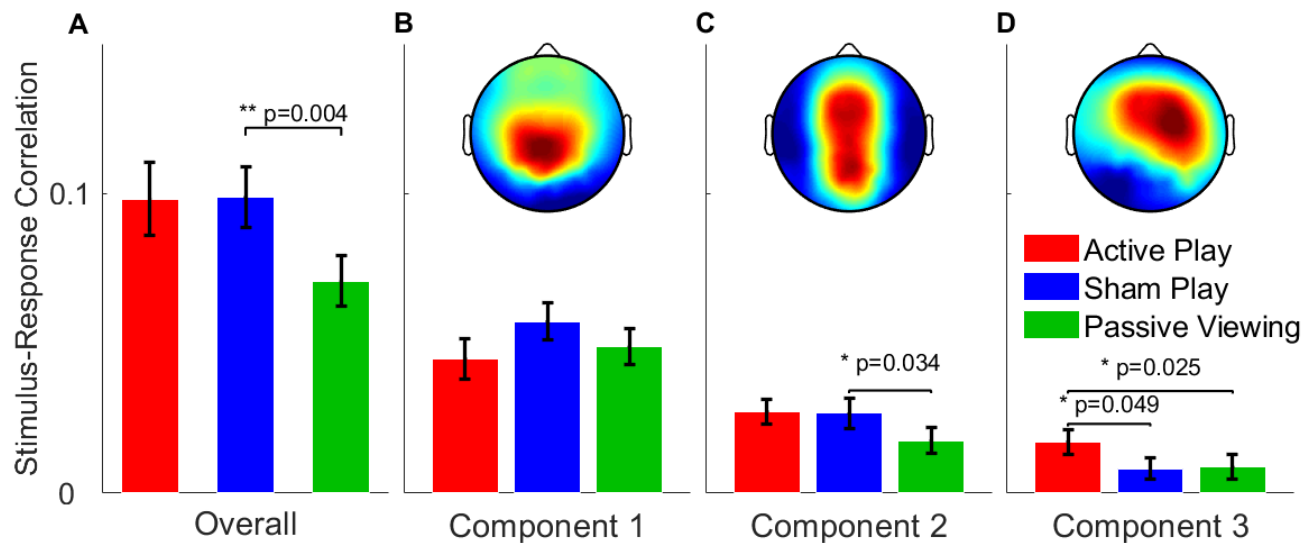
Fig. S1. Reproducibility of effect with visual contrast. We tested whether the effect of motor engagement on visual evoked responses would be reproduced when regressing the EEG onto visual contrast as opposed to the optic flow used in the main analysis. (**A**) Sham play elicited significantly higher total SRC compared to passive viewing ($p = 0.004$, one-tailed, paired Wilcoxon signed rank test, $n = 18$). Active play also produced higher total SRC than passive viewing but the difference fell short of statistical significance ($p = 0.069$). (**B**-**D**) Computing SRC separately for each component (topographies shown in insets), we found higher SRC in component 2 for sham play relative to passive viewing ($p = 0.034$). In component 3, active play produced higher SRC relative to both sham play ($p = 0.049$) and passive viewing ($p = 0.025$).

Ki *et al.* | Action enhances perception
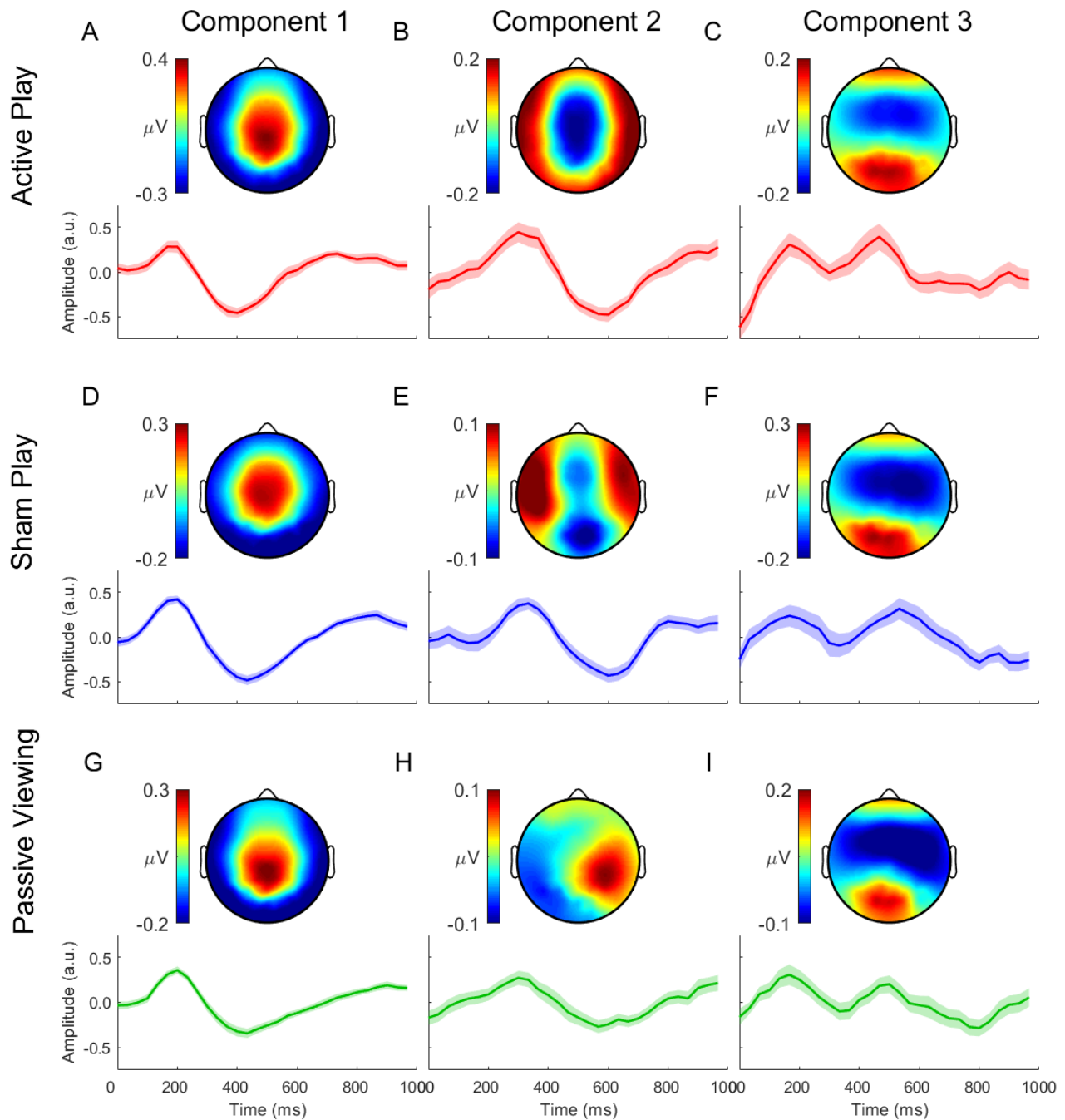
Fig. S2. Spatial and temporal responses for all experimental conditions: topographies and temporal responses are shown for the first three components for (**A**-**C**) active play, (**D**-**F**) sham play, and (**G**-**H**) passive viewing. Gross features are conserved for the three components, with the exception of the topography of component 2, which diverges across conditions. Note that the topographies of component 1 during active and sham play exhibit stronger responses over the medial frontocentral scalp relative to passive viewing.