PRIFYSGOL BANGOR / BANGOR UNIVERSITY

Using guitar learning to probe the Action Observation Network's response to visuomotor familiarity

Gardner, Tom; Aglinskas, Aidas; Cross, Emily

Neuroimage

DOI:
10.1016/j.neuroimage.2017.04.060

Published: 01/08/2017

Publisher's PDF, also known as Version of record

Cyswllt i'r cyhoeddiad / Link to publication

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):
Gardner, T., Aglinskas, A., & Cross, E. (2017). Using guitar learning to probe the Action Observation Network's response to visuomotor familiarity. *Neuroimage*, *156*(August), 174-189. https://doi.org/10.1016/j.neuroimage.2017.04.060

Hawliau Cyffredinol / General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

12. Oct. 2024

NeuroImage 156 (2017) 174–189

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Using guitar learning to probe the Action Observation Network's response to visuomotor familiarity

Tom Gardner, Aidas Aglinskas, Emily S. Cross*

School of Psychology, Bangor University, Bangor, Gwynedd LL57 2AS, United Kingdom

ARTICLE INFO

Keywords:
Action Observation Network
Familiarity
Direct matching
Predictive coding
ROI regression
fMRI
Neural efficiency
Music learning
Guitar

ABSTRACT

Watching other people move elicits engagement of a collection of sensorimotor brain regions collectively termed the Action Observation Network (AON). An extensive literature documents more robust AON responses when observing or executing familiar compared to unfamiliar actions, as well as a positive correlation between amplitude of AON response and an observer's familiarity with an observed or executed movement. On the other hand, emerging evidence shows patterns of AON activity counter to these findings, whereby in some circumstances, unfamiliar actions lead to greater AON engagement than familiar actions. In an attempt to reconcile these conflicting findings, some have proposed that the relationship between AON response amplitude and action familiarity is nonlinear in nature. In the present study, we used an elaborate guitar training intervention to probe the relationship between movement familiarity and AON engagement during action execution and action observation tasks. Participants underwent fMRI scanning while executing one set of guitar sequences with a scanner-compatible bass guitar and observing a second set of sequences. Participants then acquired further physical practice or observational experience with half of these stimuli outside the scanner across 3 days. Participants then returned for an identical scanning session, wherein they executed and observed equal numbers of familiar (trained) and unfamiliar (untrained) guitar sequences. Via region of interest analyses, we extracted activity within AON regions engaged during both scanning sessions, and then fit linear, quadratic and cubic regression models to these data. The data best support the cubic regression models, suggesting that the response profile within key sensorimotor brain regions associated with the AON respond to action familiarity in a nonlinear manner. Moreover, by probing the subjective nature of the prediction error signal, we show results consistent with a predictive coding account of AON engagement during action observation and execution that also takes into account effects of changes in neural efficiency.

Introduction

Watching others in action provides important information about other people's goals, intentions, and desires. When we observe others moving around us, we can predict how their current and future actions might unfold, thus enabling us to respond appropriately to those we encounter in a social world (Blakemore and Frith, 2005). Action observation elicits activity in a network of sensorimotor brain regions collectively termed the Action Observation Network (AON; Cross et al., 2009; Grafton, 2009; Keysers and Gazzola, 2009; Caspers et al., 2010). The core brain regions that compose the AON include occipitotemporal regions associated with observing bodies in motion, as well as the premotor cortex and inferior parietal lobule. These latter two brain regions have been shown to contain so-called mirror neurons in the non-human primate brain (di Pellegrino et al., 1992; Gallese et al., 1996; Rizzolatti et al., 2001; Umiltà et al., 2001), and demonstrate a similar response profile during action observation and execution in the human brain (Gazzola and Keysers, 2009; for a review see Molenberghs et al., 2012). Previous literature demonstrates that the more familiar an action is, the stronger the response is within these core AON regions (Buccino et al., 2004; Calvo-Merino et al., 2005; Cross et al., 2006; Shimada, 2010). Moreover, we recently demonstrated that complex, whole body movements that participants rated as more familiar were associated with greater AON activity compared to movements rated as less familiar (Gardner et al., 2015). These magnitude-based approaches support experience-driven simulation accounts of action perception (Sinigaglia, 2013), which form the foundation of the direct matching hypothesis of action understanding (Rizzolatti et al., 2001; Gallese and Goldman, 1998; Wolpert et al., 2003; although see Csibra, 2005 and Kilner, 2011 for alternative accounts). In terms of familiarity, a linear relationship between magnitude of AON activity and familiarity would be consistent with this hypothesis: as familiarity increases, the

* Corresponding author.
E-mail address: e.cross@bangor.ac.uk (E.S. Cross).

http://dx.doi.org/10.1016/j.neuroimage.2017.04.060
Received 2 March 2016; Accepted 25 April 2017
Available online 04 May 2017
1053-8119/ © 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/BY/4.0/).

simulation of how an action might unfold over time becomes more accurate and resonance between an observer's motor system and an observed action is maximised. This relationship is illustrated in Fig. 1***A***.

On the other hand, an increasing number of studies report findings demonstrating that AON activity does not necessarily follow this linear trend of increasing engagement with increasing familiarity (Gazzola et al., 2007; Liew et al., 2013; Cross et al., 2012; Tipper et al., 2015). These studies demonstrate equivalent or greater AON activity when participants observe actions that are *unfamiliar* (compared to more familiar actions), a finding that appears at odds with a simulation-based account of AON function. The findings from these studies suggest that a linear relationship between AON activity and familiarity is likely too simplistic. In terms of the direct matching hypothesis, this theory would struggle to explain why an unfamiliar action that is *not* in the observer's repertoire would elicit greater AON activity compared to a familiar action. Aspects of predictive coding models of AON function (Keysers and Perrett, 2004; Kilner et al., 2007a, 2007b; Gazzola and Keysers, 2009; Schippers and Keysers, 2011; Tipper et al., 2015), predicated on the use of perceptuomotor maps to predict and interpret observed actions (Lamm et al., 2007; Schubotz, 2007; Urgesi et al., 2010) may help resolve these seemingly discrepant findings concerning the relationship between action familiarity and engagement of sensorimotor cortices. This framework proposes a Bayesian comparison of predicted and observed actions, creating a reciprocally modulated network comprising premotor cortex (including the inferior frontal gyrus), inferior parietal lobule, and posterior temporal cortices (middle and superior temporal gyri). Activity in this network serves to minimise differences between observed and predicted actions. When observing a less familiar action, predictions (feedback signals from frontal → parietal → temporal cortices) are lacking or are under informed, and thus do not match incoming information about the observed action (feedforward signals from temporal → parietal → frontal cortices), which equates to high prediction error. This could result in robust AON engagement for highly unfamiliar actions, as the influence of feedforward/perceptual activity is heavily relied upon. When viewing an action that is highly familiar, however, predictions generated by the network should be much more precise, thus minimising prediction error. The minimising of prediction error could also manifest as robust AON engagement, this time due to the strength of feedback signals projecting posteriorly (which were weaker when movements were unfamiliar and prediction error was higher; see also Cross et al., 2012). The reciprocal nature of exchanging prediction error signals between core AON nodes allows for the explanation of robust AON engagement for both familiar *or* unfamiliar actions, relative to actions of an intermediate level of familiarity (illustrated in Fig. 1***B*** [1]). It is important to note as well that while this Bayesian framework has been most fully developed in the realm of action observation, it also has been applied to action execution, formally known as active inference (Friston, 2005).

As several authors have now suggested, a predictive coding account of action familiarity and AON engagement could manifest as a quadratic, or U-shaped, function (Cross et al., 2012; Liew et al., 2013). However, as identified within the predictive coding literature (Kilner et al., 2007a, 2007b; Friston, 2005), a system that relies on Bayesian comparisons would need to continually update predicted movements in relation to actual movements. For example, when observing an expert guitarist playing a scale (a familiar sequence of actions for both novice and expert guitarists), the player may use all her fingers to achieve this goal. As such, the actions performed by an expert guitarist to play a scale might look much different to those performed by a novice guitarist to play the same scale (i.e., a novice might use fewer fingers and/or transition between notes more awkwardly), even though the musical outcome (playing a scale) remains the same. Therefore, the ability to predict others' actions is subject to continual evaluation, and, at times, reassessment of predictions (c.f. Shadmehr and Holcomb, 1997). A quadratic function may thus not fully capture the dynamic nature of learning, prediction, and experience-driven changes in AON engagement. When considering a quadratic framing of the AON engagement and familiarity relationship, a question remains concerning what happens to AON engagement during the trough of the curve. One possibility is that ongoing evaluation of predicted and actual actions manifests as local reductions in activity within a testing session due to practice, in line with Neural Efficiency (NE) effects (Babiloni et al., 2010; Kelly and Garavan, 2005; Wiestler and Diedrichsen, 2013). In keeping with this prior work on neural efficiency, we might expect that reduced activity within a testing session should recover during subsequent testing sessions, and then reduce again as familiarity and experience continue to accrue. This conceptualisation, combining the predictive coding theoretical account with notions of neural efficiency, would create a cubic shaped response of AON engagement. To our knowledge, these three framings of the relationship between familiarity and AON engagement (i.e., direct matching vs. predictive coding vs. predictive coding + neural efficiency) have not yet been directly compared with empirical evidence.

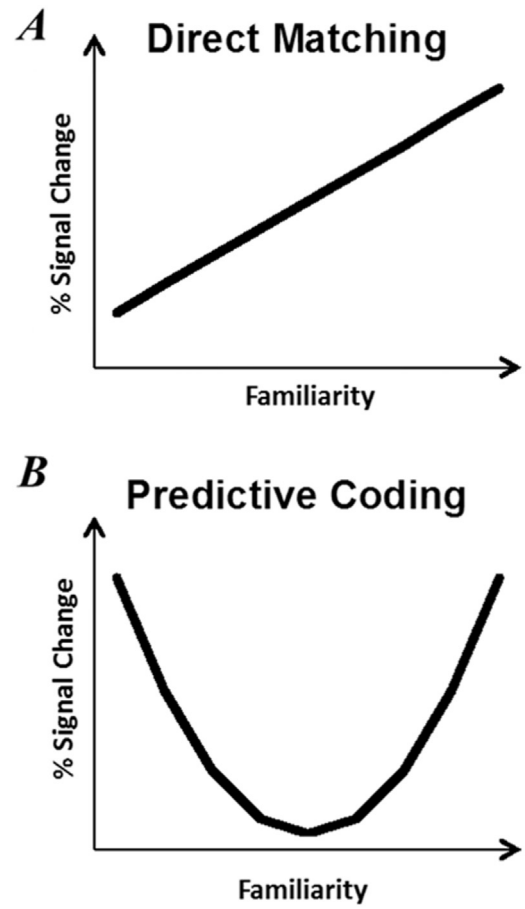

Fig. 1. Hypothesised relationships between familiarity and % signal change (BOLD signal) for both ***A*** direct matching (as proposed by Rizzolatti et al., 2001; Gallese and Goldman, 1998) and ***B*** predictive coding (as proposed by Cross et al., 2012; Liew et al., 2013).

[1] It should be clearly noted, however, that the predictive coding account is much broader in scope in terms of feed-forward and feedback exchange of information between and within networks engaged in action observation and action execution than the experimental approach and resolution of the current study can satisfactorily address (e.g., Keysers and Perrett, 2004; Keysers and Gazzola, 2009; Kilner, 2011). Ongoing work in our laboratory seeks to use effective connectivity measures to explore Hebbian learning and these broader predictive coding ideas in more depth, while the present study is focused on evaluating magnitude-based hypotheses AON engagement that have their origins in distinct theoretical accounts.

In the current study, we address two distinct questions relating to the relationship between familiarity and AON engagement. First, we aimed to compare the direct matching hypothesis (linear model) with a predictive coding account (quadratic model) with a predictive coding plus neural efficiency account (cubic model), to determine which model of AON engagement best explains the impact of varying levels of familiarity on executed or observed actions. To test whether the response of the AON to varying levels of familiarity is best captured by a linear, quadratic, or cubic function, we combined an intensive training intervention, pre- and post-training fMRI scans, and a region of interest-led analytical approach (similar in methodology to that of Mattavelli et al., 2012). Our task involved two types of action-related task: action observation and action execution. In the observation condition, guitar-naïve participants observed an expert musician playing short musical sequences on a bass guitar, after which participants responded to an attentional control question. In the execution condition, the participants played a different set of short musical sequences during scanning on a custom-built scanner-compatible bass guitar. This particular training paradigm enabled us to establish a clear distinction between familiar and unfamiliar stimuli, and the use of execution and observation conditions facilitates closer comparison and scrutiny of how these two kinds of experience shape AON responses. A guitar-playing paradigm was used as it enabled the initial testing session to feature truly unfamiliar actions for all participants. The actions required to pluck strings and press frets are themselves accessible to those who have never before played guitar, but the specific timing and sequences of actions required to play along with the songs featured in the experiment created a challenging task that required considerable practice. This complex task allowed us to track the evolution of novel multimodal action representations as they transitioned from unfamiliar to familiar (without plateauing or reaching ceiling), whilst also maintaining participant engagement. Higuchi et al. (2012) used a similar guitar playing paradigm, where participants used the fingers of one hand to press the appropriate frets for different guitar chords. The results from this prior study showed evidence of NE for both observation and execution conditions across training days. Based on this precedent, we expect to broadly replicate Higuchi et al.'s (2012) finding, with the caveat that our stimuli and task were more complex as they required coordination of both hands. The regions of interest were identified from an action observation and action execution vs. implicit baseline contrast. For both conditions, these regions were taken from both days of scanning, for all blocks, to which linear, quadratic and cubic regression models were fitted for each region, for each block, for every participant. According to the direct matching hypothesis (Rizzolatti et al., 2001; Gallese and Goldman, 1998; Wolpert et al., 2003), when blocks are ordered by familiarity, the response in these regions should increase linearly in magnitude (Fig. 1*A*). A quadratic relationship between familiarity and BOLD signal would provide support for the predictive coding account (Fig. 1*B*). These first two models were designed to evaluate whether the relationship was linear or nonlinear (quadratic) in nature, the two most prominent accounts of AON engagement and familiarity. Furthermore, the inclusion of a cubic model (predictive coding account + neural efficiency) enabled us to examine whether a more sophisticated non-linear model that takes into account within-session changes in familiarity best explains the relationship between increasing familiarity and AON engagement.

Our second aim was to evaluate the internal consistency of our findings. We recently demonstrated that self-report ratings can provide a sensitive measure of action familiarity (Gardner et al., 2015). To test the correspondence between objective and subjective measures of familiarity, we collected individual ratings of familiarity by each participant and used these as the independent variable in a second set of regression models, as an alternative to simply the number of times participants were exposed to each musical sequence. This approach enabled us to establish the extent to which findings from our first approach (using objective measures of exposure/familiarity) are replicated via subjective ratings of familiarity reported by each participant for each musical sequence.

Materials and methods

Participants

Twenty-two healthy young adult volunteers recruited from the local University community took part in the experiment and received £20 in exchange for their time. Two volunteers were excluded from the final sample due to excessive head motion during scanning. The final sample comprised 20 volunteers (9 males, *M_age* = 20.60 years, *SD* = 1.73). All participants had normal or corrected to normal vision with no history of neurological illness. All participants were right-handed and required to play a right-handed guitar in the scanner. Participants were brought to the lab prior to scanning to ensure that they could play the instrument in the manner required inside the scanner bore, and also to ensure that all were guitar novices. The study was approved and conducted following the guidelines of the Ethics Committee of the School of Psychology at Bangor University and the Bangor Imaging Unit. All participants provided written informed consent prior to their participation.

Stimuli & apparatus

Sequences

We chose 16 sequences from the computer game Rocksmith® (Ubisoft, 2014), lasting an average of 15.8 s (*SD* = 2.37 s). These sequences were excerpts taken from songs initially chosen due to their lack of lyrical content. This restriction was selected so participants would not associate any particular action sequence with lyrics. The sequences were also matched on the difficulty level assigned by Rocksmith®. Rocksmith® assigns difficulty level via an algorithm that assesses song speed and the number of notes to be played within the time window (the difficulty level is visible at the top of Fig. 2). To ensure that difficulty levels were matched as closely as possible across all stimuli, the number of notes to be played and the length of the sequences were matched, as were beats per minute for the individual song excerpts. In addition, the inclusion criteria for these sequences required that notes fell within a fret range of 1–7 and string range of 1–3. This restriction ensured that participants would not have to move their heads to identify frets during scanning, while at the same time maintaining a level of difficulty that would challenge participants throughout the training period. Furthermore, we also excluded technical guitar playing movements such as "hammering" and "sliding" so that the actions required would be accessible to novices. Finally, we matched the sequences on mean amount of motion energy displayed in each video (see Bobick, 1997; Schippers et al., 2010; Cross et al., 2012), to ensure gross differences in the amount of motion displayed in individual sequence videos did not contribute to basic visual differences between stimuli or training conditions.

Rocksmith® Guitar Task

Shown in Fig. 2 is a screenshot of the gameplay from Rocksmith® for the sequences that were physically executed (Panel A) and those that were observed (Panel B). The horizontal coloured lines at the foot of the screen correspond to the strings of the guitar being played. We used a bass guitar for this study as it has fewer strings than a standard electrical guitar (4 vs. 6) and at a basic level is generally considered easier to learn to play than a traditional guitar. These coloured strings are illustrated from a first person view, as if one is looking through the back of the fretboard of the guitar (i.e., when holding the guitar, the red string corresponds to the top string, green the second string from the top, and so on).

The second aspect of the gameplay to be explained is the translucent blue "conveyor belt" of notes seen in the centre of the screen. The

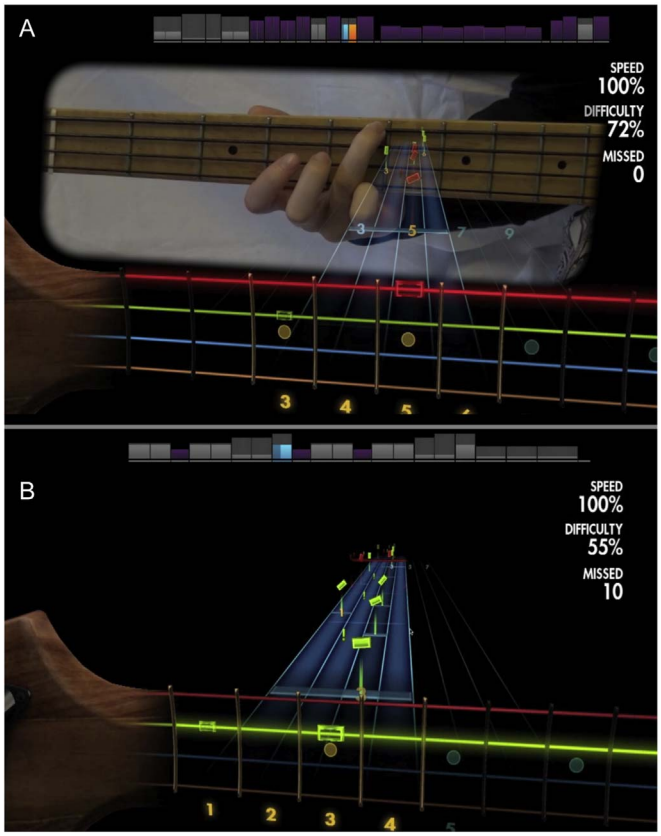

Fig. 2. A still video frame example of the stimuli used in the Observation condition (A) and execution condition (B). The hand and fretboard of the musician playing the guitar in the centre of (A) is superimposed on the actual game play. See Supplementary Materials for video examples of the stimuli.

numbers on this conveyor correspond to the fret number on the guitar itself. The coloured rectangles are critical for the participant to attend to in order to make the correct movement. The colour of the rectangle corresponds to the string to be played with the right hand, and the number fret it appears on corresponds to which fret should be pressed down with the left hand. As the rectangles and numbers on the conveyor move towards the fore of the screen, the rectangles on the conveyor rotate 90 degrees from vertical to horizontal. When they reach the horizontal position, they contact the strings at the bottom of the screen, and this is when the participant must play the appropriate note. If correctly executed, the rectangle illuminates slightly but if missed, the word "miss" appears on the string and fret that should have been played and the counter located at the upper right corner of the screen is adjusted accordingly. To obtain a perfect score, the correct fret on the correct string had to be plucked within a ± 250 ms window of the onset of the note.

Observation task

Fig. 3 shows a schematic of the observation task. Participants first viewed a brief fixation cross, followed by a video of an expert musician playing one song excerpt. The musician performs all of the songs without fault, thus providing a precise template for participants to observe.

In order to ensure that participants paid close attention to the actions performed by the expert musician in the Observation condition, an attentional control task was implemented whereby participants had to identify whether the musician palmed the strings during each stimulus video. A palming of the strings occurs when the musician removes her fingers from the fretboard, extends them vertically, and places the palm of her hand over the frets. This action was performed quickly as the musician then immediately continued to play the correct notes without stopping and without error (see Supplementary Videos 1 and 2 for examples of sequences with and without palming actions, respectively). At the end of each sequence, participants were asked

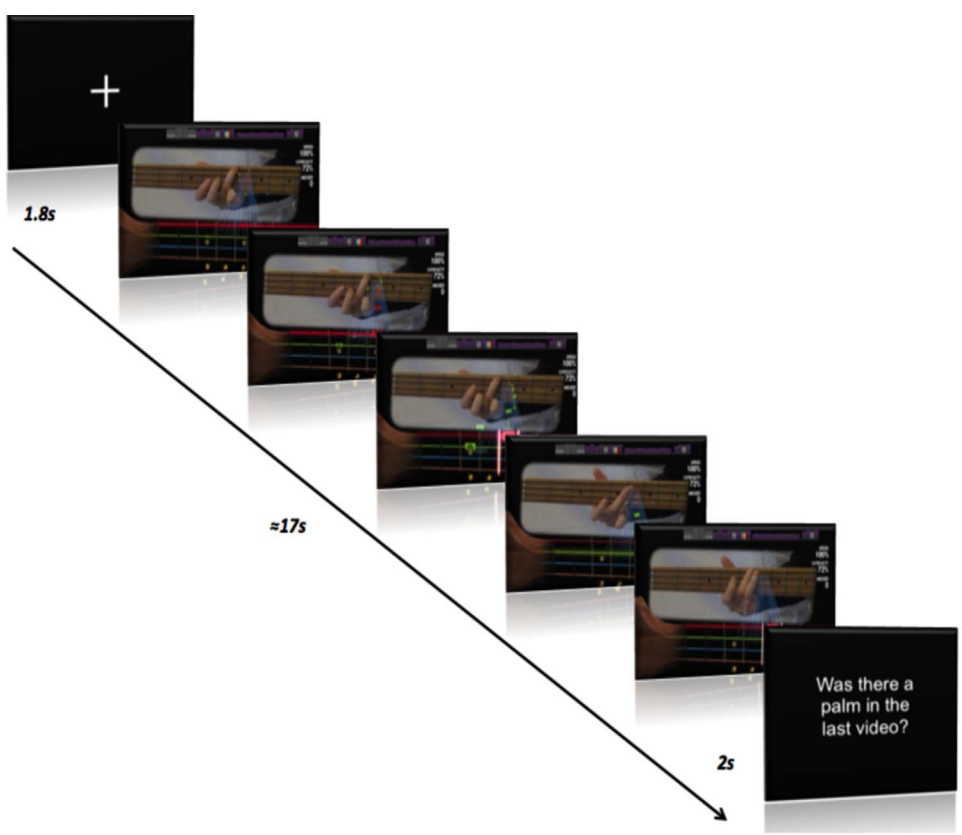

Fig. 3. A schematic of one observation trial. The task was to attend to the hand and respond to whether there was a palming of the strings.

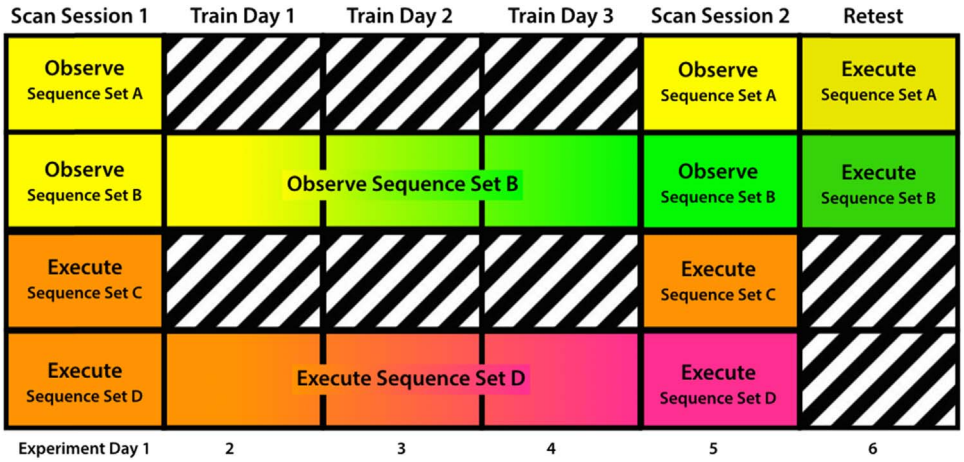

Fig. 4. The timeline of both training and test days for both conditions. The gradient of colour illustrates the change in familiarity due to training. The Retest (far right column) always occurred after Scan Session 2, although not the day immediately after (between 14 and 21 days post Scan Session 2).

whether a palming action was seen in the last sequence, and were required to make their response by plucking one of two strings with the right index finger. To ensure there was no confusion about what was required of participants during the observation task, the concept of palming the strings was explained and demonstrated before the experiment began.

Execution task

In the execution condition, Fig. 2 *(Panel B)*, participants were instructed to play along with the sequences to the best of their ability. Participants were also instructed to simply move on to the next note if they missed a note, to ensure they did not move excessively when trying to compensate for an error. After each sequence was played, the gameplay presented a count of how many notes were missed, providing a general measure of performance feedback.

Training procedure

The study began with all participants taking part in an fMRI scanning session (the first column of Fig. 4), wherein they observed eight sequences and executed eight different sequences.

Each sequence contained between 20–36 notes (M= 29.79; SD = 5.09). Sequences were matched on difficulty and note distribution across the entire sequence. Participants played or watched each of the eight sequences from the different training conditions twice per block, and completed two blocks of both kinds of training on each day of scanning. All the stimuli on the first day are labelled unfamiliar, as they were all novel to the participants. The order of the stimuli, which set was observed and executed, and the order of conditions were counterbalanced across participants.

Experiment days 2 through 4 were the training phase (green boxes in Fig. 4). During these days, participants were invited to the lab and asked to perform the same tasks they completed in the scanner, on precisely half of the execution sequences and half of the observation sequences. The four execution sequences and four observation sequences were performed or observed four times per training day, and the order of practice was counterbalanced across condition and sequences. The training set up in the laboratory was designed to match that of the scanner as closely as possible. Due to the fragile nature of the scanner safe guitar, a standard 4-stringed bass guitar was used in the training sessions to eliminate risk of damage to the equipment. Participants were required to lay on a table with the guitar placed over their midriff – similar to how the guitar was positioned in the scanner, and participants viewed the 24-inch iMac screen through prism glasses (this can be seen in Fig. 5). Stimulus presentation and response collection were performed using Psychophysics Toolbox (v3) via

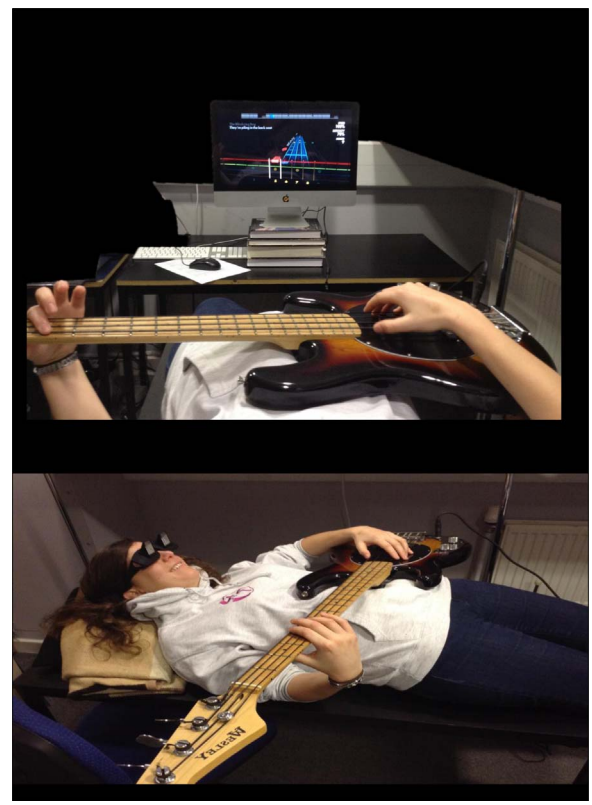

Fig. 5. The setup for the training period. This setup was designed so that it matched the fMRI setup as accurately as possible. *N.B.* The guitar was held like a double bass, with the neck of the guitar at 30 degrees, the guitar is held at 90 degrees above for illustrative purposes only.

MATLAB R2015a (MathWorks). The instruction was also given to keep movement to a minimum (for example, no tapping the guitar to the beat of the song), and the researchers monitored this. During the observation condition, participants were asked to rest their left hand over the frets so that there was no possibility that they could move their hand along with the musician's, ensuring that any learning was due to observation alone. During the execution condition, participants were asked to play the songs as accurately as they could.

On the fifth day of testing, participants returned to the scanner where they completed an identical scanning protocol to scanning session 1, performing or observing both trained and untrained sequences. Following the second scanning session, participants were

invited back to the laboratory for a sixth and final day of testing, wherein they were asked to perform all the songs from the observation condition (including the four observationally trained and the four untrained songs; Retest, Fig. 4). Due to scheduling complications, not all participants were able to attend this final testing session, and we thus report data from this final day of testing for the 15 participants who were able to return for this final session.

Familiarity rating

After the second scanning session, participants came out of the scanner and performed a rating task on the stimuli. Participants were asked to observe videos of the expert guitarist playing each sequence and to rate on a Likert scale of 1–9 on how familiar they were with each sequence (with anchors 1 = highly unfamiliar and 9 = highly familiar; the identical scale to that used by Gardner et al., 2015). Participants were asked to use the whole scale and to respond as quickly as they could.

Neuroimaging procedure & parameters

Each participant completed two identical fMRI sessions for scanning sessions 1 and 2 of the experiment that followed an event related design. Each scanning session featured two action execution blocks and two action observation blocks, presented in a counterbalanced order across participants and across scanning sessions. All eight excerpts (henceforth referred to as 'trials') were experienced twice per block. Each observation block lasted an average of 5 minutes (range = 5.07–5.70 min), and each action execution block lasted 11 min (range = 10.97–11.80 min). This difference occurred due to the buffering varying lengths of the different sequences and loading times, two factors that were not modelled within a trial nor used in matching sequence length. For the observation trials (shown in Fig. 3), at the start a fixation cross was shown for 1.8 s. This was followed by the video clip of the agent playing along with the excerpt (audio was included here and during the training period). After each clip, a black screen was presented with the question "did the musician seen in the video make a palming action over the strings?" The question screen was displayed for 2 s before moving on to the next trial. Participants were required to respond to the question within that 2 s window by plucking the appropriate string to denote their answer. During the action execution blocks, participants first saw a brief interval where the song was being loaded. This aspect was unavoidable as we wanted to gain actual response accuracy via the game, so had to load each song as if it were selected by the user (the transition between menu and sequence to play was automated via a MATLAB script, and the entire load time before each sequence ranged between 19.02 and 41.45 seconds. Once the appropriate sequence was selected, there was a buffer supplied via the game so that there was an adequate amount of time before the participant began performing the sequence, allowing for finger position adjustment before each execution sequence began. After playing along with each sequence, participants' accuracy scores were displayed for participants to see before returning to the menu screen to begin the next trial.

Stimulus presentation and response collection were performed using Psychophysics Toolbox (version 3) via MATLAB R2015a (MathWorks) run via a MacBook Pro laptop computer. The stimuli were presented on a 24" LCD BOLDScreen (Cambridge Research Systems), which was visible to the participant via a mirror mounted on the head coil. Participants listened to the song excerpts through Phillips MR-compatible headphones.

Participants were given a MR-compatible bass guitar to make their responses during the execution and observation blocks. The guitar was a full-length bass guitar, which presented some challenges for participants to manage whilst in the scanner. Participants were positioned into the scanner bore slowly and shown the best way to hold the bass guitar so as not to damage the guitar or the head coil. The guitar worked via a piezo-pickup embedded in the head of the guitar, under the strings (which were made of nylon). The guitar's tuning pegs were manufactured via 3D printing using a glass/plastic alloy. The output of the guitar was passed along a fibre optic cable from the scanner room to the control room where it was amplified and fed into the MacBook Pro running MATLAB. Offline, the responses for the observation condition were filtered to remove any RF interference created by the scanner. The gameplay applied filtering for the execution condition so that the note being played could be heard by the participant.

Data acquisition was conducted at the Bangor Imaging Unit at Bangor University, Wales. Functional images were acquired on a 3.0 T Phillips MRI scanner using a SENSE phased-array 32-channel head coil. Functional images were acquired covering the whole brain using an echo-planar imaging (EPI) sequence (35 axial slices, ascending slice acquisition, repetition time = 2000 ms, echo time = 30ms, 90° flip angle, matrix = 64 × 64, slice thickness: 3 × 3 × 3 mm, field of view (FOV): 224 mm). Before the functional run, 196 two-dimensional anatomical images (256 × 256 pixel matrix, T1-weighted) were obtained for normalization ROI selection and manipulation.

fMRI data analysis

The total number of functional scans collected for the observation blocks ranged between 156 and 178, and between 316 and 340 scans for each execution block. The number of scans for each subject was identical across scanning sessions. Data were analyzed using Statistical Parametric Mapping (SPM12: Wellcome Trust Centre for Neuroimaging, London; Friston, 2007), implemented using MATLAB R2015a (MathWorks). The data were first realigned and then slice-time corrected and preliminarily preorientated within standard stereotaxic space as defined by the MNI (Friston, 2007). This preorientation allowed for better spatial normalization to the MNI template. Participants' EPI images were then coregistered to their T1 anatomical scans, which were then spatially normalized to standard stereotaxic space. The spatially normalized EPI images were filtered using a Gaussian kernel of 8 mm full-width at half maximum in the x, y, and z axes. For the observational blocks, the design matrix was fitted for each subject with a single regressor for the familiar stimuli, a single regressor for the unfamiliar stimuli and a single regressor for the fixation and responses. For the execution blocks, this setup was the same with the inclusion of a loading period rather than the fixation and response regressor. The 4 blocks (2 Observation, 2 Execution) for both scanning sessions were placed into the same design matrix, giving one design matrix per participant. A whole brain analyses was performed for all analyses at the $p < 0.001$, $k = 10$ threshold. Only clusters that survived FWE correction were considered for further ROI analysis. All brain regions that emerged from analyses were identified via the Anatomy Toolbox (Eickhoff et al., 2007).

The main neuroimaging analyses were designed to achieve two distinct objectives:

Imaging Objective 1: Evaluate shape of regression function within ROIs based on objective measure of familiarity

The first imaging objective was designed to compare the direct matching hypothesis with the predictive coding account, with the predictive coding + neural efficiency account, in terms of which model encapsulates varying levels of familiarity best for our task. The steps in this process involved first identifying ROIs and then fitting the regression models.

Identification of ROIs. ROIs were identified and extracted from the final observation and execution blocks from the post-training scan session. The contrast used to identify these ROIs was observed *and* executed sequences > implicit baseline (i.e., whatever is not included in the model, in this case, rest/intertrial fixations). This enabled us to

identify regions that were active when viewing both familiar and unfamiliar actions (as the final blocks contained trained and untrained sequences for both the observation and execution conditions), as well as active for both action observation and execution. The clusters were identified with a threshold of $p < 0.001$ (uncorrected) and $k = 10$ voxels at the group level. The same threshold was then used to identify regions for each participant for each condition, for each block of the remaining three blocks (two blocks from the pre-training session, and the first block from the post-training session; split by familiarity, creating six block). The final block (used to identify the ROIs) was excluded from the subsequent regression analysis to avoid circularity of analyses. Next, the MARSBAR toolbox (Brett et al., 2002) was used to extract the time series for each region, for each subject, which was then transformed into percent signal change, averaged over trials. For each block, a single statistic of the relative activity for the given region, compared to the mean activity of the brain was given. The resulting percent signal change values for each block and condition (for each region) were then used to address the main hypotheses in the study. ROI analyses for the observation condition and execution condition were performed independently.

Fitting regression models across varying levels of familiarity. The first aim of this study was to address whether the response to varying levels of familiarity is best captured by a linear (direct matching), quadratic (predictive coding) or cubic (predictive coding plus neural efficiency) regression model. To address this, we fitted linear, quadratic and cubic regression models to the percent signal change within each ROI, following the procedures reported by Mattavelli et al. (2012).

Appropriate weighting was applied to the stimuli so that the differences between them were comparable to that of the amount of training (to be) undertaken. For example, the difference between Session2_Un_Block1 and Session2_Fam_Block1 was 9 as the former was only viewed three times over the course of the experiment whereas the latter was viewed 12 times. This weighting allowed for a better approximation of the level of familiarity of the blocks, thus facilitating a more accurate fit of the regression equations. For each participant, for each training condition, a linear regression, a quadratic regression and a cubic regression curve were fitted to each region. For each region and training condition, three R-squared values were taken: one indicating the fit of the linear regression model, the second indicating the fit of the quadratic regression model, and the third indicating the fit of a cubic regression model. We then ran a 3×10 repeated measures ANOVA for the variables model (linear, quadratic, cubic) and area (10 ROIs), for both conditions. This allowed us to test two questions: first, whether a main effect of area was present, which would indicate whether any of the models fit the response profile of a particular region better than any of the other regions (manifest as a higher R-squared value). Second, and most crucially, we could also test whether a main effect of model was present. An interaction of area x model was also tested, and all post hoc analysis were Bonferroni-corrected for by multiple comparisons.

Imaging Objective 2: Evaluate the shape of regression function within ROIs based on subjective measure of familiarity

The aim of the second set of imaging analyses was to evaluate the internal consistency of our findings concerning objective and subjective measures of familiarity. These analyses were performed on the first block of the second scan session only. The rationale behind this approach was that the subjective ratings should only be meaningful from the second scanning session, once participants have undergone the training period. The ROIs were extracted as stated above and participants' individual subjective ratings for each sequence were used as the independent variable in the regression models, rather than exposure. For example, if a subject gave the ratings of 1,3,5,7,9 to the stimuli, these values were used in place of the objective number of exposures independent variable from the previous section. In this example, a linear model would fit the data if the percent signal change were lower for the videos given the lower ratings than those given higher ratings. Sequences were pooled across ratings. The ratings of each subject were standardised to the mean, with the number of different ratings used by each participant ranging between 4 and 6 (out of 9 possible values). One participant was excluded from the execution condition due to only having used three rating values out of the 1 – 9 rating scale (see Section: Familiarity rating).

Results

Behavioural results

For the observation condition, the accuracy score related percent correct in terms of whether there was a palming of the strings or not (coded by the experimenter). For the execution condition, Rocksmith® provided a count of how many notes were missed. This was then transformed into a percentage of notes correctly played and used for analysis. The same protocol was used for the training and scanning sessions.

Observation condition

The observation responses for the training sessions (see Fig. 6***A***) showed that for all training days, the responses were significantly greater than chance (50%): training day 1 ($M = 82.29$, $SE = 3.20$); $t(19) = 10.07$, $p < 0.001$; training day 2 ($M = 81.77$, $SE = 2.31$); $t(19) = 13.75$, $p < 0.001$; and training day 3 ($M = 85.21$, $SE = 2.29$); $t(19) = 15.36$, $p < 0.001$). These results suggest that on each day of training, participants could accurately identify whether there was palming of the strings, confirming close attention was paid to the hand actions performed in the observation sequences. A repeated-measures

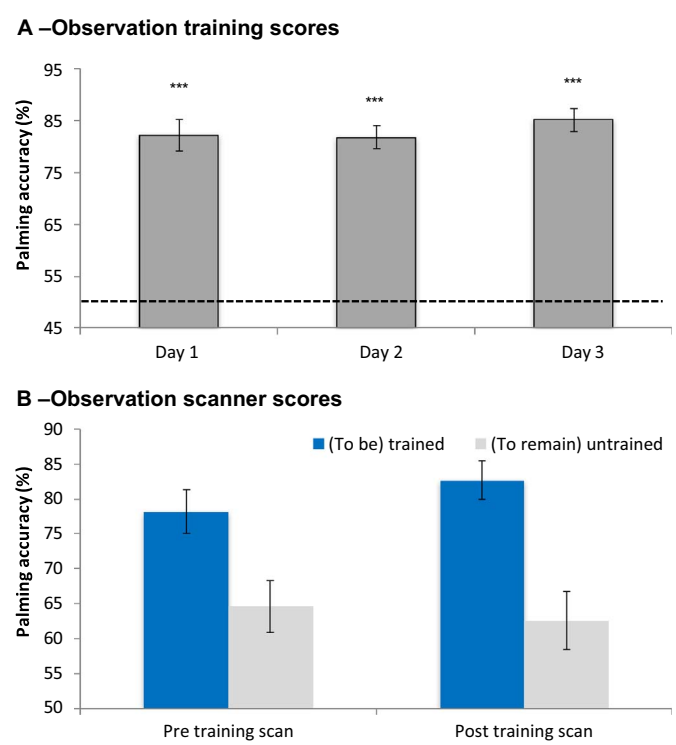

Fig. 6. Palming accuracy scores (%) for the Observation condition for ***A***, the training period and ***B***, the scanning sessions. The dotted line denotes chance (50%). *** $p < 0.001$. Bonferroni correction for multiple comparisons applied; error bars represent the standard error of the mean. *n.b.* Due to a programming error, it was not possible to parse observation scores in the scanner by scanning run (as done with execution scanner scores in Fig. 7b).

ANOVA on these accuracy scores reveal no reliable differences between performance across training days, $\boldsymbol{F}(2, 38) = 0.91$, $p = 0.413$, suggesting that participants performed this rather simple detection task at a consistently high level of accuracy throughout the training period.

To assess the responses made to the observation task during the scanning sessions (see Fig. 6*B*), a 2×2 ANOVA was run with factors Scan Session (Session 1/pre-training vs. Session 2/post-training) and Familiarity (familiar vs. unfamiliar). A main effect of familiarity emerged, $F(1, 16) = 90.699$, $p < 0.001$, such that participants' average accuracy for detecting the palming action on familiar stimuli ($M = 80.94$, $SE = 3.03$) was greater than the unfamiliar stimuli ($M = 63.06$, $SE = 3.98$). No main effect of scanning session was found ($p = 0.681$), nor did an interaction emerge between scanning session and familiarity ($p = 0.266$). While Fig. 6B clearly shows a difference in observation task accuracy between to be trained and to remain untrained observation sequences during the pre-training scan, as all sequences were equally unfamiliar at this stage, so any pre-training differences are likely to reflect noise rather than actual differences.

A subsample of 15 participants returned to the laboratory after completing both fMRI sessions and the training sessions to perform those stimuli which were only observed, never executed, in order to get an objective measure of how much participants actually learned via observation (see Supplementary Materials). This follow-up test showed participants could play the observed sequences significantly better than those that remained untrained; $t(14) = 5.782$, $p < 0.001$. This result suggests that participants did indeed learn to some degree how to perform those sequences that were passively observed during the training period.

Execution Condition

For the execution condition, a repeated measures ANOVA revealed a significant difference in participants' ability to play the guitar riffs across training days; $F(2,42) = 40.00$, $p < 0.001$ (Fig. 7***A***). Further analysis revealed that a significant increase in accuracy between Day 1 (M = 48.01) and Day 2 ($M = 62.05$; $p < 0.001$), Day 1 and Day 3 ($M = 69.58$; $p < 0.001$), and between Day 2 and Day 3 ($p = 0.022$) These differences indicate that accuracy significantly improved across training days, thus demonstrating clear learning induced by the RockSmith® guitar playing task.

The execution scores achieved during scanning are illustrated in Fig. 7***B***. These scores were subjected to a 2x2×2 ANOVA with factors Scan Session (Session 1/pre-training vs. Session 2/post-training), Block (Block 1 vs. Block 2), and Familiarity (familiar vs. unfamiliar) A main effect of scan session emerged, $F(1, 21) = 68.78$, $p < 0.001$, such that the average performance accuracy at scan session 2 ($M = 55.96$, $SE = 2.81$) was greater than that at scan session 1 ($M = 33.12$, $SE = 2.27$). There was also a main effect of Block, $F(1, 21) = 20.02$, $p < 0.001$, indicating that accuracy in the second Block of each scan session ($M = 47.37$, $SE = 2.33$) was better than the first block ($M = 41.70$, $SE = 2.14$). A main effect of familiarity also emerged, $F(1, 21) = 30.41$, $p < 0.001$, indicating that accuracy was better for familiar (to-be-trained/trained; $M = 48.25$, $SE = 1.99$) compared to unfamiliar excerpts (to-remain-untrained/untrained; $M = 40.82$, $SE = 2.49$).

In terms of interactions, a significant interaction emerged between familiarity and scan session, $F(1, 21) = 22.79$, $p < 0.001$, indicating that differences in performance accuracy for unfamiliar compared to familiar excerpts increased as a function of training. A significant interaction between familiarity and block was also present, $F(1, 21) = 5.02$, $p = 0.036$, suggesting that some learning occurred across the blocks (regardless of scanning session), driven by the unfamiliar excerpts showing more marked improvements in performance in the second block compared to the first block of each scanning session. No significant interaction between session and block was found, nor was a 3-way interaction between session, block, and familiarity; all p values > 0.05.

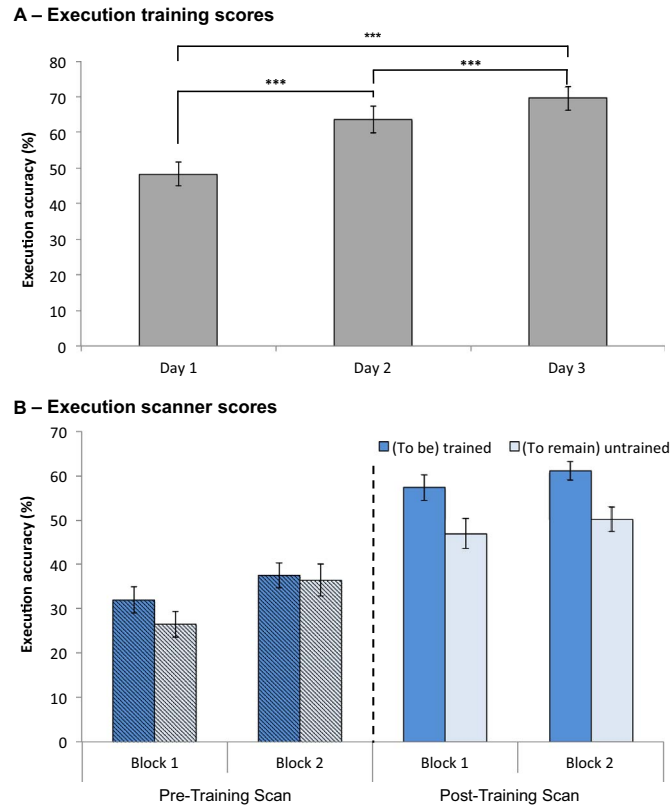

Fig. 7. Execution accuracy scores (%) for the Execution condition for ***A***, the training period and ***B***, the scanning sessions. Gradient bars indicate parentheses in legend and the dashed line represents where the training period occurred *** $p < 0.001$, ** $p < 0.01$. Bonferroni correction for multiple comparisons applied; error bars represent the standard errors of the mean.

fMRI results

ROI Identification

In order to identify regions of interest with the action observation network, we report a single random effects contrast from the second block of the post-training scan session: familiar and unfamiliar execution and observation vs. implicit baseline.

The ROI contrast of Execution and Observation vs. implicit baseline revealed widespread engagement of regions associated with action observation and execution (see Table 1 for a full list of regions). Only cluster corrected regions (denoted in bold in the table) were considered for further analysis. This analysis performed initially at the $p < 0.001$ (uncorrected) threshold revealed four cluster-corrected regions: right precentral gyrus (PrCG), right inferior frontal gyrus (IFG), left IFG and the right putamen (Fig. 8). Further examination revealed that the R PrCG cluster extended over 5000 voxels, therefore the threshold was elevated to $p < 0.05$ *(FWE*-corrected) to enable investigation of the discrete cluster-corrected regions within this larger cluster. From the $p < 0.05$ *(FWE*-corrected) thresholded analysis, right PrCG, left PrCG, right middle temporal gyrus (MTG), left middle occipital gyrus (MOG), right superior temporal gyrus (STG), right precuneus and right supplementary motor area (SMA) were identified. In addition to these seven regions, the R IFG, L IFG and right Putamen regions from the initial analysis were also considered for all subsequent ROI analyses, thus enabling consideration of a broad set of brain regions reliably engaged by action observation and execution.

Response profile to varying levels of familiarity: Testing the linear, quadratic and cubic accounts

Fig. 9 illustrates the mean percent signal change for each region of interest, for both the Observation and Execution conditions. At the

Table 1
Regions associated with Observation & Execution activity.

Anatomical region	BA	MNI Coordinates			Putative functional name	T value	Cluster size	P_{FWE}-corrected statistic
		X	Y	Z				
R Precentral Gyrus	**4**	**27**	**-13**	**64**	**PrCG**	**12.62**	**673**	**< 0.001**
R Precentral Gyrus	4/5	36	-40	58	PrCG	10.66		
R Precentral Gyrus	4/3	39	-31	43	PrCG	9.86		
L Precentral Gyrus	**4**	**-36**	**-16**	**58**	**PrCG**	**11.97**	**728**	**< 0.001**
L Precentral Gyrus	4/3	-30	-25	55	PrCG	10.26		
L Post Central Gyrus	7/4	-39	-43	46	PoCG	9.81		
R Middle Temporal Gyrus	**39**	**51**	**-70**	**1**	**MTG**	**10.70**	**260**	**< 0.001**
R Middle Occipital Gyrus	18	33	-85	1	MOG	9.09		
R Middle Occipital Gyrus	18	33	-85	16	MOG	8.11		
L Middle Occipital Gyrus	**18**	**-42**	**-73**	**4**	**MOG**	**9.32**	**134**	**< 0.001**
L Inferior Occipital Gyrus	18	-36	-82	4	MT+	8.83		
R Superior Temporal Gyrus	**42**	**63**	**-31**	**16**	**STG**	**8.73**	**33**	**< 0.001**
R Inferior Parietal Lobule	40	54	-37	19	IPL	7.75		
R Inferior Parietal Lobule	40/42	63	-43	10	IPL	7.06		
R Precuneus	**7**	**12**	**-70**	**46**	**PreCu**	**8.24**	**34**	**0.002**
R Supplementary Motor Area	**6**	**6**	**-1**	**55**	**SMA**	**7.9**	**43**	**0.004**
R Supplementary Motor Area		6	-7	64	SMA	7.15		

Bold indicates cluster corrected regions. MOG = Middle Occipital Gyrus, MTG = Middle Temporal Gyrus, STG = Superior Temporal Gyrus, PrCG = Precentral Gyrus, PoCG = Postcentral Gyrus, IPL = Inferior Parietal Lobule, PreCu = Precuneus, SMA = Supplementary Motor Area.

group level, cubic patterns of response amplitude in response to familiarity emerge across the regions. Further analyses exploring differences between Area, Scanning session and Block can be found in Supplementary material.

The main aim of this study was to systematically evaluate whether the relationship between familiarity and activity within brain regions engaged during action observation and execution is linear, quadratic or cubic. To directly assess this, linear, quadratic and cubic regression functions were fitted to each region for each participant. The average *R*-squared value was then calculated for all forms of regression and compared via repeated measures ANOVA (results shown in Fig. 10).

As shown in Fig. 10 **A**, the cubic regression fitted to the percent signal change within each individuals' seven ROIs explained more variance than the linear and the quadratic regression on the same individuals and regions for the observation of actions. To examine these differences, a 10 × 3 repeated measures ANOVA was performed on the R-squared values for each region with the factors Area (10 ROIs) and Model (linear, quadratic and cubic). For the Observation condition, the repeated measures ANOVA revealed that there was a main effect of Area, $F(9, 171) = 2.631$, $p=0.007$, indicating a difference in overall fit of the models across ROIs (shown in Fig. 11 ***A***). Post hoc analyses (Bonferonni corrected) revealed significant differences between RSTG ($M=0.446$, $SE=0.050$) and LMOG ($M=0.652$, $SE=0.045$; $p=0.004$), and between RSTG and RMTG ($M=0.656$, $SE=0.054$; $p=0.023$). This indicates that the regression models fit the response within the RSTG ROI less well than the two ROIs with the best fitting models. Also present was a main effect of Model, $F(2, 38) = 103.034$, $p < 0.001$, suggesting a difference in goodness of fit (in terms of R-squared) between the models tested. Post hoc analysis revealed that the quadratic model ($M=0.605$, $SE=0.037$) fit significantly better than the linear model ($M=0.359$, $SE=0.030$; $p < 0.001$), and that the cubic model ($M=0.785$, $SE=0.016$) fit better than the quadratic model ($p < 0.001$). Overall, this pattern suggests that the cubic model, which combines aspects of predictive coding with neural efficiency, best explains these data. No interactions emerged between the variables (all p values > 0.10).

For the Execution condition (illustrated in Fig. 10 **B**), here again the cubic function provided the best fit of the data. While the repeated measures ANOVA revealed a marginal main effect of Area, $F(9, 171) = 1.970$, $p=0.046$, post hoc analyses did not reveal any findings that survived correction for multiple comparisons. A main effect of Model was present, $F(2, 38) = 56.425$, $p < 0.001$, suggesting differences in goodness of fit between the three different models. Similar to the observation condition, posthoc analyses revealed that the quadratic

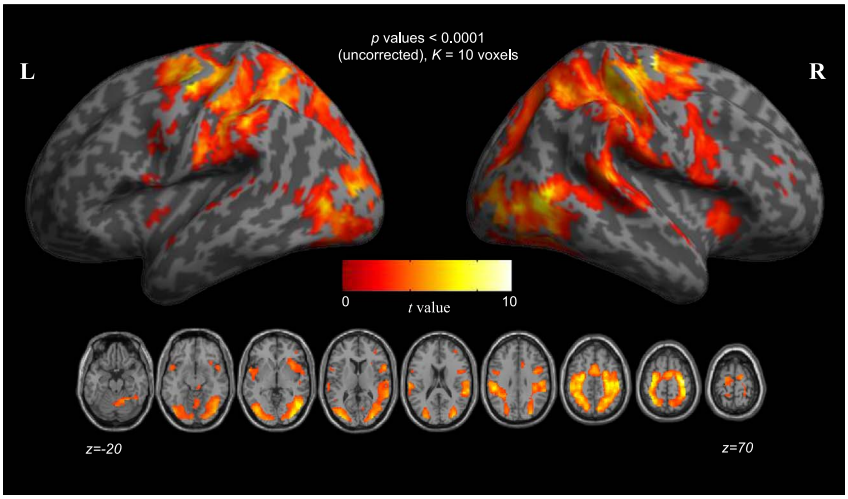

Fig. 8. Whole brain group analysis Session 2 Block 2 familiar and unfamiliar execution and observation vs. implicit baseline. All p values < 0.0001 (uncorrected), $K = 10$ voxels.

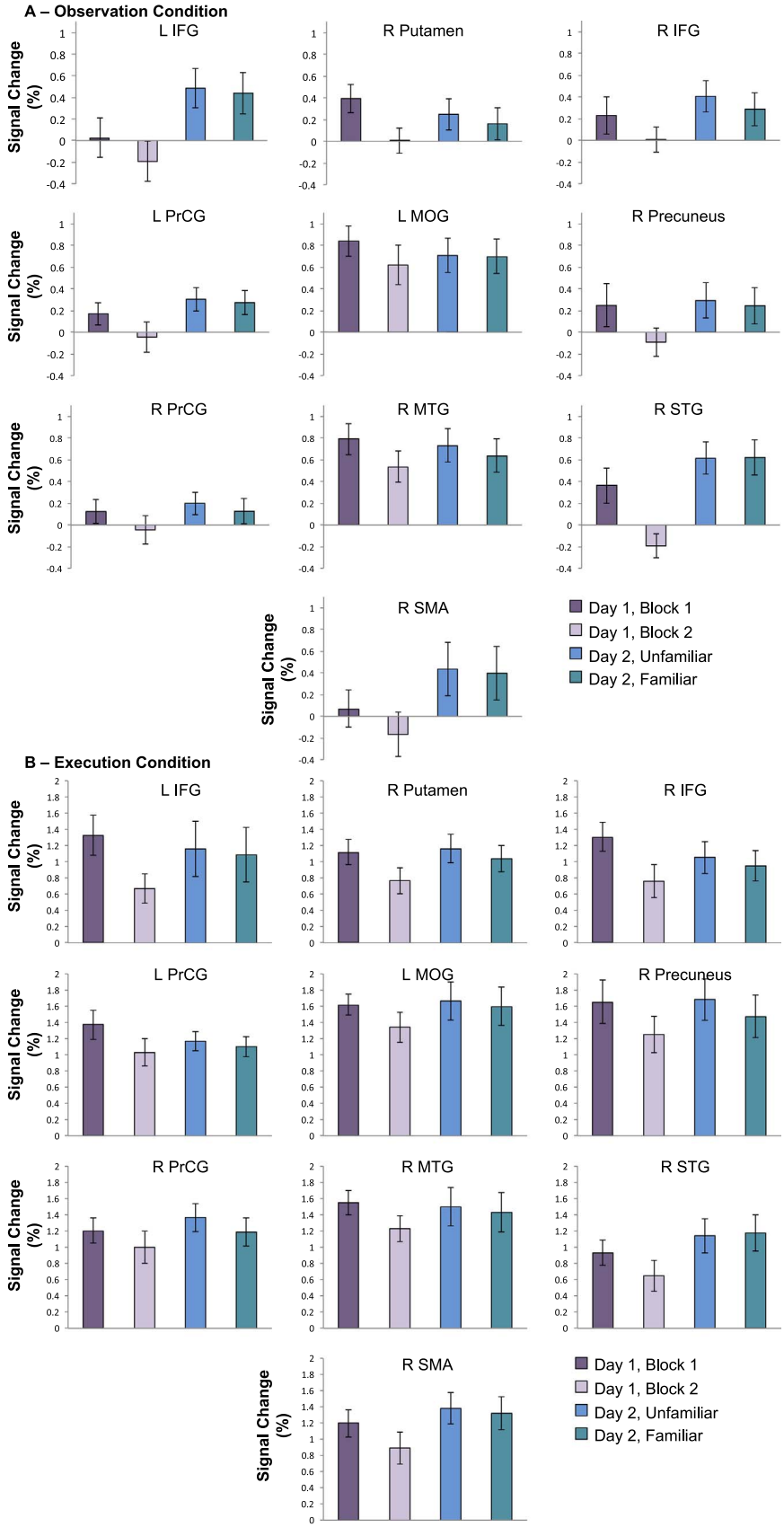

Fig. 9. Response profiles within each ROI, expressed as percent signal change over familiarity. **A** illustrates the results of the observation condition and **B** the execution; error bars represent the standard errors of the mean. MOG = Middle Occipital Gyrus, MTG = Middle Temporal Gyrus, STG = Superior Temporal Gyrus, PrCG = Precentral Gyrus, PreCu = IFG = Inferior Frontal Gyrus, SMA = Supplementary Motor Area.

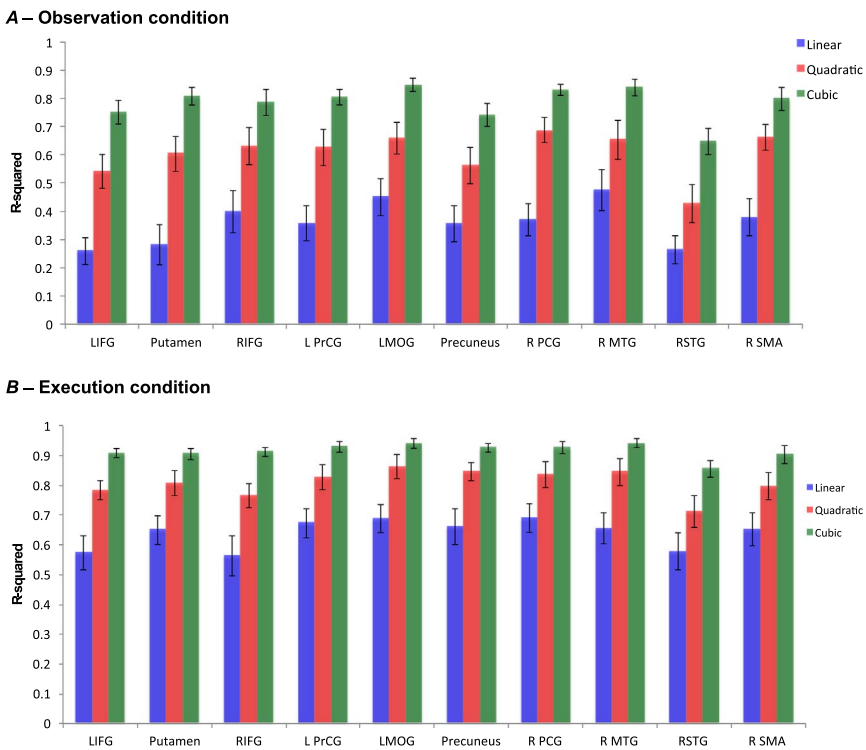

Fig. 10. *R*-squared values for the linear, quadratic and cubic regressions for each ROI using objective familiarity scores, for both Observation **A**, and Execution **B**. Error bars represent the standard errors of the mean. MOG = Middle Occipital Gyrus, MTG = Middle Temporal Gyrus, STG = Superior Temporal Gyrus, PrCG = Precentral Gyrus, SMA = Supplementary Motor Area, IFG = Inferior Frontal Gyrus.

model (M=0.809, SE=0.040) fit the data significantly better than the linear model (M=0.639, SE=0.040; $p < 0.001$), and that the cubic model (M=0.915, SE=0.011) provided a better fit than the quadratic model ($p < 0.001$). No interactions emerged between any of the variables (all p values > 0.10).

Internal Consistency

Thus far, we have shown that the response profile of all ROIs is most appropriately captured by a cubic regression model when using number of exposures as the independent variable. Our next aim was to test the extent to which this *objective* measure of action familiarity (i.e., number of times any given sequence has been practiced or observed) is consistent with a *subjective* measure of familiarity (i.e., participants' ratings). To achieve this, we applied the same analytical approach described above, and this time used participants' own subjective ratings as the independent variable for establishing the regression models. This approach enables us to evaluate whether our original, objective measure of familiarity is replicated by a subject, subject-specific (and potentially more sensitive) measure of familiarity. Participant ratings were taken from the 1–9 Likert scale collected post scanning. Participants' individual ratings were assigned to each stimulus, which became the independent variable in the regression model (no weighting was applied as there was equal distances between the rating points (e.g. the differences between ratings of 1 and 2 was equivalent to ratings of 6 and 7). For visualisation purposes, the ratings were grouped into unfamiliar (ratings of 1), unfamiliar to neutral (ratings of 2–3), neutral (ratings of 4–6), neutral to familiar (ratings of 7–8) and familiar (ratings of 9). The signal change for each bracket was averaged across for each region (shown in Fig. 11).

Similar to the objective familiarity ratings, Fig. 11 illustrates that the relationship between subjective familiarity ratings and AON engagement is also nonlinear. The overall picture to emerge from these findings is a subtle cubic function based on participant-reported increases in familiarity, for both observed and executed guitar sequences.

As can be seen in Fig. 12 A, the response within the same ten ROIs shows the linear model to provide the poorest fit of the data, a quadratic model to provide an intermediate level of fit, and a cubic model best capturing the relationship between engagement of key AON regions and increases in subjective ratings of familiarity. A repeated measures ANOVA on these data revealed no main effect of Brain Region, $F(9, 171) = 0.538$, p=0.845, nor an interaction between Brain Region and Model (p=0.269). A main effect of Model was present, $F(2, 38) = 161.826$, $p < 0.001$. Just as when the objective number of exposures was used as the independent variable (see Fig. 11A), here again we see a similar pattern in how the different models fit the data when familiarity is based on subjective report, with the quadratic model (M=0.518, SE=0.033) fitting significantly better than the linear model (M=0.118, SE=0.033; $p < 0.001$); and the cubic model (M=0.651, SE=0.024) fitting the data significantly better than the quadratic model ($p < 0.001$). When these results are considered alongside those reported the previous section (Fig. 10*A),* clear similarities emerge for both kinds of familiarity assessments, in that the data are best described by a cubic response profile (as opposed to linear or quadratic profiles) within all ROIs as a function of familiarity. However, we also see that when using subjective rating of familiarity, the differences between quadratic and cubic models are weaker, as we explore in more depth below.

The use of subjective ratings as the independent variable within the regression models for the execution condition yielded a similar finding to the objective number of exposures (shown in Fig. 10 B). The repeated measures ANOVA revealed no main effect of Brain Region, $F(9, 171) = 0.305$, p=0.972, or an interaction between Brain Region and Model (p=0.343). A main effect of Model emerged, $F(2, 38) = 178.826$, $p < 0.001$, with the cubic model accounting for the most variance (just as when the objective measure of familiarity was used, and as also seen in the observation condition). Here as well, we also find that the quadratic model (M=0.493, SE=0.028) fit the data significantly better than the linear model (M=0.118, SE=0.027; $p <$ 0.001); and that the cubic model (M=0.637, SE=0.026) fit the data

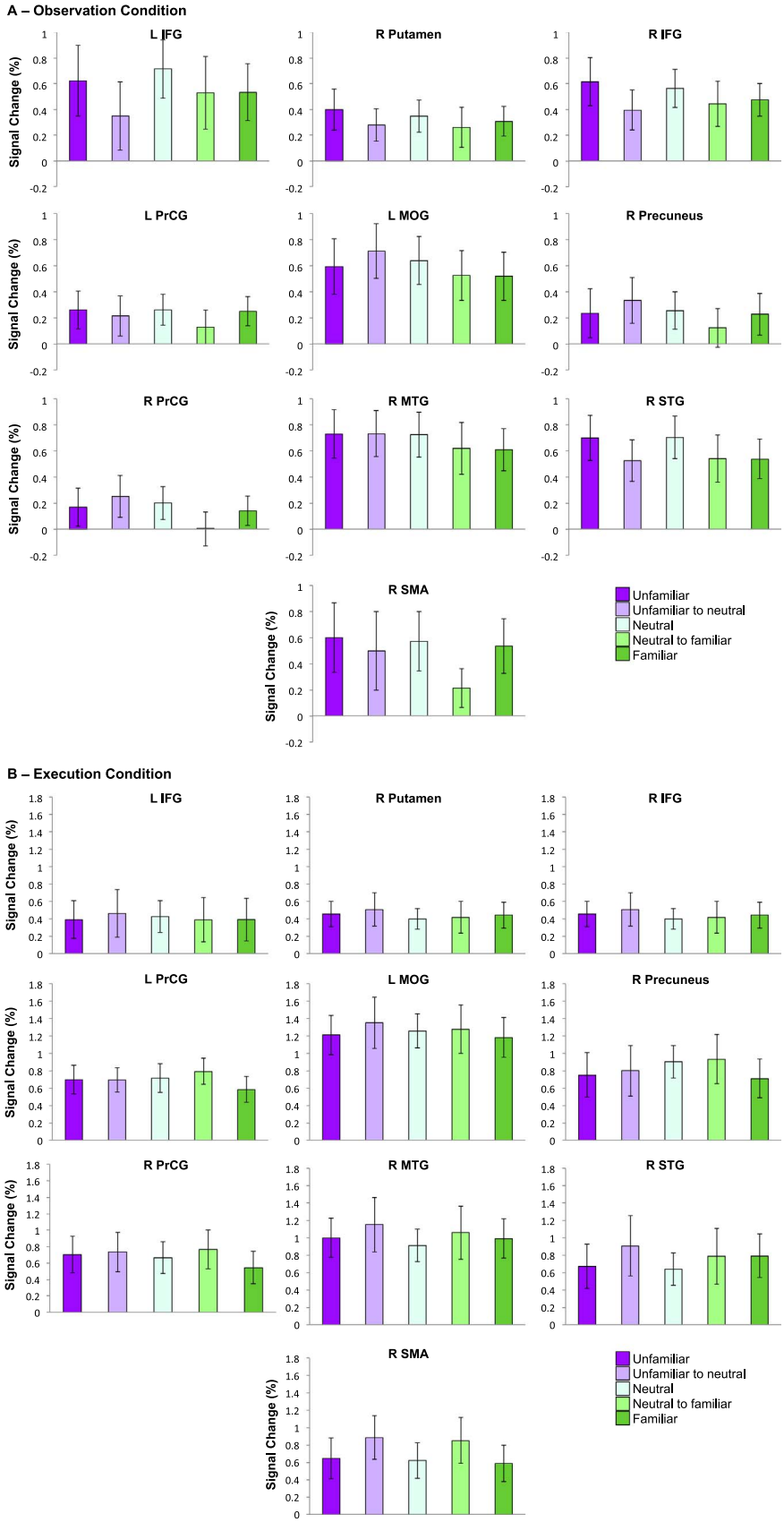

Fig. 11. Response profiles within each ROI, expressed as percent signal change over participants' subjective familiarity ratings. **A** illustrates the results of the observation condition and **B** the execution; error bars represent the standard errors of the mean. MOG = Middle Occipital Gyrus, MTG = Middle Temporal Gyrus, STG = Superior Temporal Gyrus, PrCG = Precentral Gyrus, PreCu = IFG = Inferior Frontal Gyrus, SMA = Supplementary Motor Area.

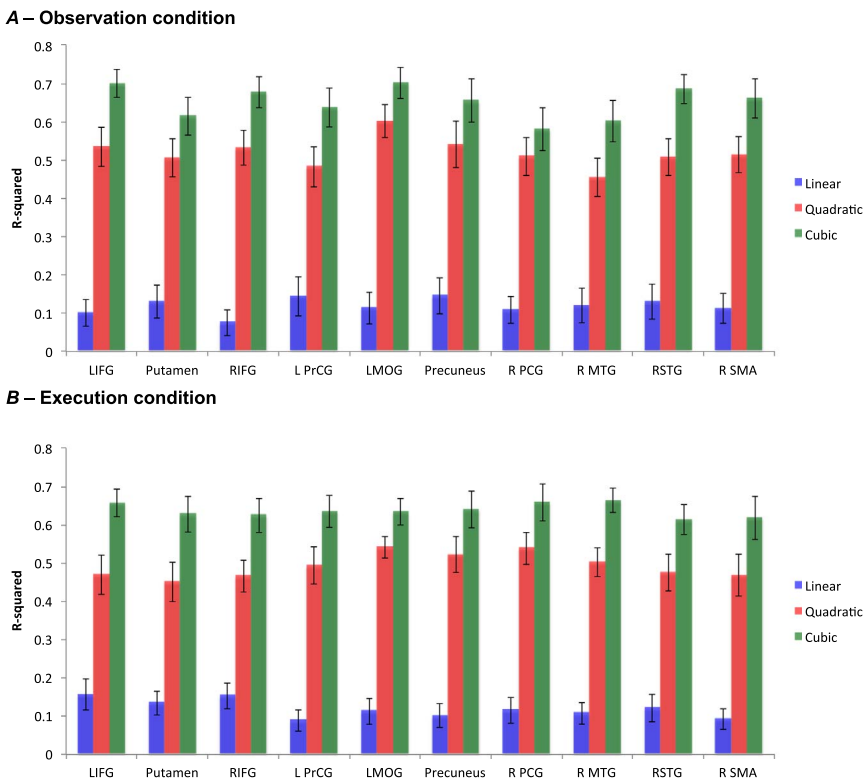

Fig. 12. *R*-squared values for the linear, quadratic and cubic regressions for each ROI using subjective ratings, for both Observation **A**, and Execution **B**. Error bars represent the standard errors of the mean. MOG = Middle Occipital Gyrus, MTG = Middle Temporal Gyrus, STG = Superior Temporal Gyrus, PrCG = Precentral Gyrus, SMA = Supplementary Motor Area, IFG = Inferior Frontal Gyrus.

significantly better than the quadratic model ($p < 0.001$).

Discussion

The aim of the present study was to investigate how varying familiarity levels modulate engagement of the action observation network (AON) during the observation or execution of guitar sequences. Specifically, we were interested in testing whether a direct matching account (Rizzolatti et al., 2001; Gallese and Goldman, 1998; Wolpert et al., 2003), a predictive coding-inspired account (Keysers and Perrett, 2004; Kilner et al., 2007a, 2007b; Gazzola and Keysers, 2009; Schippers and Keysers, 2011), or a predictive coding-inspired account that also takes into consideration effects of neural efficiency (Babiloni et al., 2010; Kelly and Garavan, 2005; Wiestler and Diedrichsen, 2013) would best explain the impact of increasing familiarity on AON engagement. To address this, we asked guitar-naïve participants to take part in identical fMRI sessions where they observed and executed different guitar playing sequences. In between the sessions, participants trained (via observational or physical practice) on half the sequences. By performing region of interest analyses on areas that exhibited activity during both action observation and execution, we addressed two distinct questions, specifically: (1) is the relationship between AON response amplitude and increasing familiarity better captured by a linear or nonlinear response profile?; and (2) how do subjective measures of familiarity compare to objective measures of familiarity in terms of influencing the response amplitude within core AON regions based on increasing familarity?

AON response to varying levels of familiarity – Evaluating the direct matching and predictive coding accounts

The first question we evaluated was whether the relationship between increasing familiarity and AON response magnitude was best captured by a linear response profile (more in keeping with a direct matching account; Rizzolatti et al., 2001; Gallese and Goldman, 1998; Wolpert et al., 2003), a quadratic response profile (which would be more consistent with a predictive coding account; Kilner et al., 2007a, 2007b; Gazzola and Keysers, 2009; Tipper et al., 2015; Cross et al., 2012; Liew et al., 2013) or a cubic response profile (in line with a more subtle conceptualization of the predictive coding account that takes into account effects of neural efficiency; Babiloni et al., 2010; Kelly and Garavan, 2005; Wiestler and Diedrichsen, 2013). Through use of pre- and post-training scanning sessions on either side of an intensive training intervention, we manipulated the degree of familiarity participants had with specific guitar riffs. After identifying regions of interest within the AON that responded to executing or observing guitar playing actions, we fitted linear, quadratic and cubic regression models to each region for each subject.

We found that for both observation and execution of complex guitar-playing actions, the response profile of all ROIs was best captured by a cubic regression model. This primarily suggests that the relationship between familiarity and AON engagement is not linear, and also that the theoretical framings that suggest a quadratic function of AON engagement based on increasing familiarity (Cross et al., 2012; Liew et al., 2013; Diersch et al., 2013) need adjustment to reflect the continual adaptations made by these regions in light of changes in action familiarity.

Generation of linear, quadratic and cubic regression models enables us to directly compare two dominant models of AON function: the direct matching framework (linear; Rizzolatti et al., 2001; Gallese and Goldman, 1998; Wolpert et al., 2003) and a framework inspired by a predictive coding model (nonlinear; Keysers and Perrett, 2004; Kilner et al., 2007a, 2007b; Gazzola and Keysers, 2009; Schippers and Keysers, 2011). Briefly, the predictive coding framework suggests reciprocal modulation between the nodes of the AON, which provides feedback predictions, and feedforward updates with the aim to minimise error signals. A nonlinear function of familiarity would therefore fit this assumption, as at the highly familiar end of this scale, greater

familiarity with an observed or executed action should result in greater/more accurate prediction, while at the highly unfamiliar end of the scale, less familiar actions should be associated with ongoing updating, and both of these scenarios could result in robust AON activity. In contrast, according to a direct matching framework (Rizzolatti et al., 2001; Gallese and Goldman, 1998; Wolpert et al., 2003) as familiarity increases, AON activity should track familiarity gains in a linear fashion. To our knowledge, the present study is the first to directly test these two accounts by examining the shape of the response profile within key AON ROIs, based on varying levels of familiarity, whilst also including an alternative non-linear hypothesis to examine whether the relationship was cubic in nature. Our data suggest that a nonlinear model fits the data better than a linear model, and the cubic nature of our best-fit model highlights the potential impact of neural efficiency on experience-dependent plasticity within the AON. This finding is concordant with findings reported by Higuchi et al. (2012), who used a similar guitar paradigm that required participants to finger different chord patterns with one hand only. These authors showed that for both observation and execution of guitar chords, reliable neural efficiency effects emerged across training days. Here we provide further support for this finding by demonstrating this relationship to be consistent across participants in terms of models based on objective *and* subjective measures of familiarity, and we further advance the findings reported by Higuchi et al. (2012) by showing a similar pattern with a more complex, bimanual guitar performance and observation task.

In addition, it is of interest that the mean *R*-squared values are remarkably similar between the observation and execution conditions. This result corroborates one of the main findings reported by neurophysiological investigations of mirror neurons in non-human primates: namely, that the response profile of some cells within parietal and premotor cortices during observation is comparable to those same cells' response profile during action execution (di Pellegrino et al., 1992; Gallese et al., 1996; Rizzolatti et al., 2001; Umiltà et al., 2001). Naturally, we cannot conclude that the ROIs examined in the present study are actually coding information in the same way during observation and execution. Nonetheless, the fact that cubic functions best capture the response profiles of these regions during *both* execution and observation supports the notion that not only do these core sensorimotor cortical regions play a critical role in action observation and execution, but also that they respond to changes in familiarity in a similar manner as well.

One important addition or alternative to the neural efficiency explanation of the current pattern of findings is worth noting. Revisiting Fig. 9, the pattern observed shows a greater discrepancy/reduction in response amplitude between blocks 1 and 2 on the first, pre-training day of the scanning, compared to between the unfamiliar and familiar trials observed on the second, post-training day of scanning. Alternative explanations that could drive this pattern of findings include the possibility that participants invested less effort during the second block of trials during the first day of scanning (and this reduction of effort was not seen in the post-training scan session), and/or some manner of general repetition suppression is at play during the early stages of experience or learning that fundamentally changes as experience builds (see, for example, Gotts, 2016; Wymbs and Grafton, 2015; Majdandzic et al., 2009). While the current study was designed to evaluate three relatively simple theoretical models of experience-dependent changes in AON function, an important challenge for future work will be to construct and evaluate ever more sophisticated models that take into account such issues as fluctuations in effort/engagement and changes in neural adaptation throughout the learning process as well.

Consistency between objective and subjective familiarity ratings

By evaluating a separate analysis wherein we used subjective ratings of familiarity as an independent variable within our regression model, we could investigate whether objective and subjective measures of familiarity align or differ within core AON regions. Participants' subjective familiarity ratings were obtained after the last scanning session and were assigned to the corresponding sequence within the general linear model. The fact that both approaches reliably supported a cubic relationship between AON engagement and familiarity suggests that subjective ratings offer a sensitive and subject-specific measure of experience. Such individual ratings can add value when used in conjunction with time-consuming training interventions (Cross et al., 2006; Kirsch et al., 2015; Casile and Giese, 2006; Läppchen et al., 2015), or when used in isolation when taking physical measures of performance are not feasible or possible (e.g., Gardner et al., 2015; Cross et al., 2011; Press and Kilner, 2013; Kawabata and Zeki, 2004).

Limitations and future directions

One potential limitation of the current study concerns the use of a weighted regression model. By using the number of times exposed to the stimuli as the objective familiarity parameter, the independent variable in the regression model was consequently not evenly distributed, potentially skewing the regression model. However, this limitation is less of a concern due to the fact that we by and large replicated the finding of a cubic relationship when using subjective ratings of familiarity as the independent variable in a separate set of regression models. This suggests that any potential skew introduced to the objective regression models by the uneven distribution of the independent variable is unlikely to be (solely) driving the effects reported here. A valuable extension to the current work could be to improve the regression model by examining more (and more evenly distributed) time points during training, which would provide a clearer representation of the curve (c.f. Braams et al., 2015).

Another potential limitation that warrants consideration is the use of different training scenarios for the observation and execution conditions. Specifically, in the execution condition, participants received on-line, real-time feedback about their performance, whereas no such feedback was possible for the observation task. In addition, the task participants performed for the observation training condition (an attentional control task that involved detecting an infrequent palming of the strings) was decidedly less challenging than the task participants performed in the execution condition, and also not as conducive to learning. Moreover, our methods did not enable us to measure with absolute certainty the extent to which participants were fully engaged throughout each observation video, as the palming question occurred only after each sequence. It should be noted, however, that limitations inherent to the observation condition would be of greater concern if we were interested in drawing direct comparisons between the effects of both kinds of training on performance. As our aims in the current study focused on how physical and observational experience impact AON engagement independently (but within the same group of participants), the differences in training experience were both necessary and warranted in this case. Nonetheless, we would recommend that follow-up investigations employ additional measures to ensure task engagement and compliance throughout the observation condition, such as eye tracking or a different and more continuous measure of visual engagement.

A recent study sheds further light on using rich training interventions with multisensory training experience (Kirsch and Cross, 2015). In this study, the authors showed that by use of a multisensory training paradigm that layering auditory, visual and physical experience has a cumulative effect in shaping engagement of the premotor, posterior parietal and posterior temporal cortices during action observation (Kirsch and Cross, 2015). In the current study, we used visual plus auditory stimuli in the observation condition, and multimodal stimuli that involved motor, visual and auditory systems in the physical training condition. As such, one could argue that we have effectively

"stacked the deck" in our favour (in terms of maximally engaging the AON) by using rich, multisensory stimuli to investigate ROI responses. This was a deliberate decision, as past research also shows that auditory experience in addition to visual cues enables participants to reconcile the timing of the movements in accordance with auditory feedback (see Lotze, 2013 for insights on the importance of motor, somatosensory, auditory and visual aspects in musical imagery, with a particular focus on the audiomotor loop). It would be valuable for future work, however, to investigate in finer detail the impact of unimodal vs. multimodal experience as it relates to theoretical models of AON engagement and familiarity.

Conclusions

In conclusion, the present findings address a core question concerning how familiarity shapes action observation and execution-related processing within the AON. Via a region of interest analysis, we directly tested key theoretical models of the AON better accounted for the impact of varying levels of familiarity: the direct matching and predictive coding accounts. The findings indicate, for both objective and subjective familiarity, a cubic function linking familiarity and AON responses is more clearly aligned with a predictive coding and neural efficiency account of AON function than a direct matching account.

Acknowledgements

The authors gratefully acknowledge Zoe Oliver for her guitar playing expertise, Dave McKiernan and Alan Gardner for their assistance with engineering the fMRI compatible guitar, Paul Mullins and Andrew Fischer for guitar fMRI compatibility help, Bethan Howe and Luke Malam for testing assistance, James Kilner for fruitful discussion about the nature of the cubic response, and Paul Downing, Richard Ramsey and three anonymous peer reviewers for constructive comments on previous versions of this manuscript. This work was supported by the following funding to E.S.C.: a Netherlands Organisation for Scientific Research (NWO) Veni award (#451-11-002); an Economic and Social Research Council Future Research Leaders award (ES/K001892/1); and a Marie Curie Career Integration award (CIG11-2012-322256). The authors declare no competing financial interests.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.neuroimage.2017.04.060.

References

Blakemore, S.J., Frith, C., 2005. The role of motor contagion in the prediction of action. Neuropsychology 43, 260–267.
Bobick, A., 1997. Movement, activity and action: the role of knowledge in he perception of motion. Philos. Trans. R. Soc. Lond. B 352, 1257–1997.
Braams, B.R., van Duijvenvoorde, A.C.K., Peper, J.S., Crone, E.A., 2015. Longitudinal changes in adolescent risk-taking: a comprehensive study of neural response to rewards, pubertal development, and risk-taking behaviour. J. Neurosci. 35 (18), 7226–7238.
Babiloni, C., Marzano, N., Infarinato, F., Iacoboni, M., Rizza, G., Aschieri, P., Cibelli, G., Soricelli, A., Eusebi, F., Del Percio, C., 2010. "Neural efficiency" of experts' brain during judgement of actions: a high-resolution EEG study in elite and amateur karate athletes. Behav. Brain Res. 207 (2), 466–475.
Brett, M., Anton, J., Valabregue, R., Poline, J., 2002. . Regions of interest analysis using an SPM toolbox (abstract). Presented at the 8th International conference of Functional Mapping of the Human Brain, June 2–6. Sendai, Japan. Available in CD-ROM in Neuroimage, Vol 16, No 2.
Buccino, G., Vogt, S., Ritzl, A., Fink, G.R., Zilles, K., Freund, H.-J., Rizzolatti, G., 2004. Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. Neuron 42, 323–334.
Calvo-Merino, B., Glaser, D.E., Grèzes, J., Passingham, R.E., Haggard, P., 2005. Action observation and acquired motor skills: an fMRI study with expert dancers. Cereb. Cortex 15, 1243–1249.
Casile, A., Giese, M.A., 2006. Nonvisual motor training influences biological motion perception. Curr. Biol. 16 (1), 69–74.
Caspers, S., Zilles, K., Laird, A., Eickhoff, S., 2010. ALE meta-analysis of action observation and imitation in the human brain. Neuroimage 50, 1148–1167.
Cross, E.S., Hamilton, A.F., Grafton, S.T., 2006. Building a motor simulation de novo: observation of dance by dancers. NeuroImage 31, 1257–1267.
Cross, E.S., Kraemer, D.J., Hamilton, A.F., Kelley, W.M., Grafton, S.T., 2009. Sensitivity of the action observation network to physical and observational learning. Cereb. Cortex 19 (2), 315–326.
Cross, E.S., Kirsch, L., Ticini, L.F., Schütz-Bosbach, S., 2011. The impact of aesthetic evaluation and physical ability on dance perception. Frontiers in Human. Neuroscience 5 (102), 1–10.
Cross, E.S., Liepelt, R., Hamilton, A.FdeC., Parkinson, J., Ramsey, R., Stadler, W., Prinz, W., 2012. Robotic movement preferentially engages the action observation network. Hum. Brain Mapp. 33 (9), 2238–2254.
Csibra, G., 2005. . Mirror neurons and action observation. Is simulation involved? In ⟨www.interdisciplines.org/mirror⟩.
Diersch, N., Mueller, K., Cross, E.S., Stadler, W., Rieger, M., Schutz-Bosbach, S., 2013. Action prediction in younger versus older adults: neural correlates of motor familiarity. PLOS One 8, e64195. http://dx.doi.org/10.1371/journal.pone.0064195.
di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G., 1992. Understanding motor events: a neurophysiological study. Exp. Brain Res. 91 (1), 176–180.
Eickhoff, S.B., Paus, T., Caspers, S., Grosbars, M.H., Evans, A.C., Ziles, K., Amunts, K., 2007. Assignment of functional activations of probabilistic cytoarchitectonic areas revisited. Neuroimage 36 (3), 511–521.
Friston, K., 2005. A theory of cortical responses. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360, 815–836.
Friston, K.J., 2007. Statistical Parametric Mapping.. Academic Press, London.
Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G., 1996. Action recognition in the premotor cortex. Brain 119 (Pt 2), 593–609.
Gallese, V., Goldman, A.I., 1998. Mirror neurons and the simulation theory of mind reading. Trends Cogn. Sci. 2, 493–501.
Gardner, T., Goulden, N., Cross, E.S., 2015. Dynamic modulation of the action observation network by movement familiarity. J. Neurosci. 35 (4), 1561–1572.
Gazzola, V., Keysers, C., 2009. The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. Cereb. Cortex 19, 1239–1255.
Gazzola, V., Rizzolatti, G., Wicker, B., Keysers, C., 2007. The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. Neuroimage 35, 1674–1684.
Gotts, S.J., 2016. Incremental learning of perceptual and conceptual representations and the puzzle of neural repetition suppression. Psychon. Bull. Rev. 23 (4), 1055–1071.
Grafton, S.T., 2009. Embodied cognition and the simulation of action to understand others. Ann. N. Y. Acad. Sci. 1156, 97–117.
Higuchi, S., Holle, H., Roberts, N., Eickhoff, S.B., Vogt, S., 2012. Imitation and observational learning of hand actions: prefrontal involvement and connectivity. Neuroimage 59 (2), 1668–1683.
Kawabata, H., Zeki, S., 2004. Neural Correlates of Beauty. J. Neurophysiol. 91, 1699–1705.
Kelly, A.M., Garavan, H., 2005. Human functional neuroimaging of brain changes associated with practice. Cereb. Cortex 15 (8), 1089–1102.
Keysers, C., Gazzola, V., 2009. Expanding the mirror: vicarious activity for actions, emotions, and sensations. Curr. Opin. Neurobiol. 19, 666–671.
Keysers, C., Perrett, D.I., 2004. Demystifying social cognition: a Hebbian perspective. Trends Cogn. Sci. 8, 501–507.
Kilner, J.M., 2011. More than one pathway to action understanding. Trends Cogn. Sci. 15 (8), 352–357.
Kilner, J.M., Friston, K.J., Frith, C.D., 2007a. The mirror-neuron system: a Bayesian perspective. NeuroReport 18, 619–623.
Kilner, J.M., Friston, K.J., Frith, C.D., 2007b. Predictive coding: an account of the mirror neuron system. Cogn. Process 8, 159–166.
Kirsch, L.P., Cross, E.S., 2015. Additive routes to action learning: layering experience shapes engagement of the action observation network. Cereb. Cortex 25, 4799–4811.
Kirsch, L.P., Dawson, K., Cross, E.S., 2015. Dance experience sculpts aesthetic perception and related brain circuits. Ann. New Y. Acad. Sci. 1337, 130–139.
Läppchen, C.H., Ringer, T., Blessin, J., Schulz, K., Seidel, G., Lange, R., Hamzei, F., 2015. Daily iTBS worsens hand motor training–a combined TMS, fMRI and mirror training study. Neuroimage 107, 257–265.
Lamm, C., Batson, C.D., Decety, J., 2007. The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. J. Cogn. Neurosci. 19, 42–58.
Liew, S.-L., Sheng, T., Margetis, J., Aziz-Zadeh, L., 2013. Both novelty and expertise increase action observation network activity. Front. Human. Neurosci. 7, 541.
Lotze, M., 2013. Kinesthetic imagery of musical performance. Front. Hum. Neurosci. 7, 28.
Majdandzic, J., Bekkering, H., van Schie, H.T., Toni, I., 2009. Movement-specific repetition suppression in ventral and dorsal premotor cortex during action observation. Cereb. Cortex 19 (11), 2736–2745.
Mattavelli, G., Andrews, T.J., Asghar, A.U.R., Towler, J.R., Young, A.W., 2012. Response of face-selective brain regions to trustworthiness and gender of faces. Neuropsychologia 50, 2205–2211.
Molenberghs, P., Cunnington, R., Mattingley, J.B., 2012. Brain regions with mirror properties: a meta-analysis of 125 human fMRI studies. Neurosci. Biobehav. Rev. 36, 341–349.
Press, C.M., Kilner, J.M., 2013. The time course of eye movements during action observation reflects sequence learning. NeuroReport 24 (14), 822–826.
Rizzolatti, G., Fogassi, L., Gallese, V., 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. Nat. Rev. Neurosci. 2, 661–670.

Shadmehr, R., Holcomb, H.H., 1997. Neural correlates of motor memory consolidation. Science 277 (5327), 821–825.
Schippers, M.B., Keysers, C., 2011. Mapping the flow of information within the putative mirror neuron system during gesture observation. NeuroImage 57 (1), 37–44.
Schubotz, R.I., 2007. Prediction of external events with our motor system: towards a new framework. Trends Cogn. Sci. 11, 211–218.
Shimada, S., 2010. Deactivation in the sensorimotor area during observation of a human agent performing robotic actions. Brain Cogn. 72, 394–399.
Sinigaglia, C., 2013. What type of action understanding is subserved by mirror neurons? Neurosci. Lett. 540, 59–61.
Tipper, C.M., Signorini, G., Grafton, S.T., 2015. Body language in the brain: constructing meaning from expressive movement. Front. Hum. Neurosci. 9, 450.
Umiltà, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., et al., 2001. I know what you are doing. a neurophysiological study. Neuron 31, 155–165.
Urgesi, C., Maieron, M., Avenanti, A., Tidoni, E., Fabbro, F., Aglioti, S.M., 2010. Simulating the future of actions in the human corticospinal system. Cereb. Cortex 20 (11), 2511–2521.
Wiestler, T., Diedrichsen, J., 2013. Skill learning strengthens cortical representations of motor sequences. eLife 2, e00801.
Wolpert, D.M., Doya, K., Kawato, M., 2003. A unifying computational framework for motor control and social interaction. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 358, 593–602.
Wymbs, N.F., Grafton, S.T., 2015. The human motor system supports sequence-specific representations over multiple training-dependent timescales. Cereb. Cortex 25 (11), 4213–4225.