Title page

Title

The interaction between attention and motor prediction. An ERP study

Alexander Jones12, Gethin Hughes3, & Florian Waszak12

1Université Paris Descartes, Sorbonne Paris Cité, Paris, France
2CNRS (Laboratoire Psychologie de la Perception, UMR 8158), Paris, France
3Department of Psychology, University of Essex, United Kingdom

Corresponding and first author

Alexander Jones

Université Paris Descartes
Laboratoire Psychologie de la Perception
45 rue des Saints-Peres
75006 Paris
France

Email : john.alex.jones@gmail.com
Tel : +33(0)142862198
Fax : +33(0)142863322
Abstract

Performing a voluntary action involves the anticipation of the intended effect of that action. Interaction with the environment also requires the allocation of attention. However, the effects of attention upon motor predictive processes remain unclear. Here we use a novel paradigm to investigate attention and motor prediction orthogonally. In an acquisition phase, high and low tones were associated with left and right key presses. In the following test phase, tones were presented at random and participants attended to only one ear whilst ignoring tones presented in the unattended ear. In the test phase a tone could therefore be presented at the attended or unattended ear, as well as being congruent or incongruent with prior action-effect learning. We demonstrated early and late effects of attention as well as a later independent motor prediction effect with a larger P3a for incongruent tones. Interestingly, we demonstrated an intermediate interaction, showing an action-effect negativity (N_{AE}) for tones which were unattended, whilst no motor prediction effect was found for attended tones. This interaction pattern suggests attention and motor prediction are not opposing processes but can both operate to modulate prediction, providing valuable new insight into the relationship between attention and the effects of motor prediction.

Keywords: Action effect, Auditory, EEG/ERP, Motor prediction, Selective attention

Highlights:

1. New paradigm to orthogonally manipulate attention and motor prediction
2. Independent and interdependent effects of attention and motor prediction
3. Incongruent action effect elicits P3a
4. Larger action effect negativity (N_{AE}) for unattended tones only
1. Introduction

Most actions humans perform have an effect on the agent’s surroundings. However, not all these effects are equally important. Some effects are merely incidental, whereas others have been produced on purpose. Some effects might be important for the next action, whereas others are not. To coordinate the current agent-environment interaction, humans use two mechanisms. First, they predict the sensory consequences of their actions. Action effect prediction has been discussed extensively in the literature as a mechanism of voluntary action control, since it enables the agent to select actions that are appropriate for his goals (cf., ideomotor theory of action control; e.g., Greenwald, 1970; Herwig, et al., 2007; for reviews see Waszak, et al., 2012; Shin et al., 2010). Second, humans allocate attention to sensory feedback that is important in the current context. Sensory feedback that is considered to be irrelevant, on the other hand, is ignored. Humans use this mechanism in order to channel the sensory information that floods their senses at any given moment (e.g., Desimone and Duncan, 1995; Kastner and Ungerleider, 2001).

Notice that in the control of agent-environment interactions both these mechanisms need to work hand-in-hand, with action predictive mechanisms anticipating the particular up-coming sensory effect and attentional mechanisms anticipating whether this effect is relevant for further processing. Although motor prediction and attention have separately been investigated extensively, it remains largely unknown how the two mechanisms relate. We devised a new paradigm that orthogonally manipulates motor prediction and attention. Using EEG we were able to test for topographical and temporal differences/commonalities of the effect of motor prediction and attention. The aim of the present study was to investigate whether the two mechanisms are independent or interdependent.

Both motor prediction and attention have been demonstrated to influence stimulus processing. As concerns effect anticipation, it has been shown that sensory action effects are attenuated when they
match the prediction. Sensory attenuation as a result of one’s own action has been demonstrated as a perceptual phenomenon in the somatosensory (e.g., Blakemore et al., 1998), visual (Cardoso-Leite et al., 2010), and auditory modality (Sato, 2008), as well as a neurophysiological effect (for a review see Waszak et al., 2012). In the auditory domain, the N1 component showed reduced amplitude for an action-triggered compared to externally-triggered tone (e.g., Baess et al., 2011; Aliu et al., 2009; Baess et al., 2009; Baess et al., 2008; Ford et al., 2007; Martikainen et al., 2005). Initially this has been taken as an indication that this reflects sensory suppression, with the N1 being generated in primary and secondary auditory cortex. However, recently it was shown that the N1 suppression for self- vs. externally generated sounds was not generated in sensory areas, but affected later stages of the N1, the so called unspecific N1 (N1bU) (San Miguel et al., 2013; Timm et al, 2013). More importantly in the present context, since studies which have contrasted externally generated with action-driven stimuli differ in several other attributes, and not solely action prediction, other authors have used a paradigm that systematically varies the relationship between action and action-effects (see Hughes et al., 2013a, for a review on this issue). Cardoso-Leite et al. (2010) employed a paradigm that is similar to the manipulation of effect prediction used in the present study. In an acquisition phase, participants associated a left key press with a, say, left tilting Gabor patch and a right key press with a, say, right tilting stimulus. In the test phase the action-triggered visual stimulus was either congruent or incongruent with the previously learned action-effect association. Cardoso-Leite and colleagues observed reduced sensitivity for congruent compared to incongruent stimuli, providing evidence that sensory attenuation results from a match between predicted and observed sensory action-effects. Similarly, using ERPs and auditory stimuli Hughes et al. (2013b) found an attenuated N1 for congruent compared to incongruent tones. Taken together, there is evidence to suggest that correctly predicted sensory effects are attenuated. However, since relatively few studies have adopted this type of paradigm, with the majority comparing internal vs. externally generated stimuli, the generalizability of these effects remains to be established.
Action effect prediction has also been considered to be a basis of action error monitoring (Band et al., 2009; Waszak and Herwig, 2007), as a discrepancy between predicted and true effect indicates a problem in the on-going action episode. Band et al. (2009), for example, compared the processing of explicit, performance related feedback with the processing of feedback-unrelated action effects. They found negative performance feedback to yield a feedback-related negativity ($N_{FB}$; Miltner et al., 1997). More importantly, they also showed that unpredicted sensory action effects elicit a signal similar to the $N_{FB}$. That is, tones which were incongruent with an action showed larger negativity compared to congruent action effects around 200 ms post stimulus onset, and this action–effect negativity ($N_{AE}$) was, similarly to the $N_{FB}$, largest over posterior electrodes. The existence of this action–effect negativity ($N_{AE}$) suggests that a mismatch between predicted and actual action effect results in processes related to outcome evaluation, analogous to the $N_{FB}$.

Hence, correctly predicted and unpredicted action effects give rise to distinct stimulus-driven processing. However, in an agent-environment interaction stimulus processing is not only determined by motor prediction. It is also largely affected by whether or not a stimulus is expected to be relevant for further processing, that is, whether attention is allocated to the anticipated stimulus event. Attended events are typically enhanced rather than attenuated. Endogenous orienting of attention has been reported to increase the neural processing in sensory regions rather than decreasing it (e.g., Corbetta et al., 1990; Gandhi et al., 1999; Reynolds and Heeger, 2009). In auditory attention, endogenous selective attention has frequently demonstrated amplitude enhancement in the N1 time range (Hillyard et al., 1973; Näätänen et al., 1992; Woldorff et al., 1993; Talsma and Woldorff, 2005), as well as to affect later stages of processing effects (e.g., Hari et al., 1989; Näätänen et al, 1978, Näätänen et al, 1992, Ross et al., 2010). There has been a long, and still on-going debate whether endogenous attention enhances the sensory N1 per se (e.g., Hillyard et al., 1973; Giard et al., 1988; Woldorff and Hillyard, 1991), or whether N1 attention effects are due to superposition of sustained negativity, known as a ‘negative difference’ (Nd) between attended and
unattended stimuli (Näätänen, 1982; Näätänen et al., 1978; see Alho, 1992; Näätänen & Alho, 2004 for a reviews). The rate of stimulus presentation for one has shown to influence opinion standpoints, whereby rapid tone presentation with short stimulus onset asynchrony (SOA) suggest the attention effect may be seen as an enhanced N1, whilst long intervals between tones more likely results in the difference between attended and unattended on the descending slope on the N1, attributed to an Nd effect (Schwent et al. 1976; Teder et al., 1993).

Hence, when tested separately both motor prediction and attention show characteristic signatures in EEG components reflecting different aspects of prediction- and attention-related stimulus processing. The present research addresses the question of how the two anticipatory mechanisms relate. Several scenarios were possible: motor prediction and attention might be independent of each other, both intervening at different stages of processing. Some of these independent effects might also concern the same processing stage and be opposed to each other (attenuation vs. enhancement). Or, motor prediction and attention may be completely inseparable and show only dependent effects, with, for example, attention being a precondition for motor predictive effects, or, at least, enhancing them. We also might expect to find both independent and dependent components. It is this latter scenario that turned out to be correct: We found both independent components of attention and motor prediction as well as interactions between the two processes, with attention showing early and also late effects, and motor prediction showing late effects, and interactions between the two demonstrated in an intermediate time period.

2. Materials and Methods

2.1 Participants

19 paid participants took part in this study (2 were removed due to excessive artifacts). 9 male and 8 female participants were all right handed. They had a mean age of 25.6 years. All participants provided written informed consent.
2.2 Stimuli and apparatus

Experimental tones were either high (1000 Hz) or low frequency (500 Hz) and all 100 ms in duration, with a 5 ms rise and fall time. Catch tones were exactly the same as experimental but with a 20 ms dip in pitch in the middle (40-60 ms period of tone), creating the effect of a “broken tone”. The catch trials were of 5 different difficulties. An easy catch tone had a larger dip in pitch than a hard catch tone. The amplitude of 20 ms (40-60 ms) was reduced by -16.3dB, -11.3dB, -6.3dB, -3.8dB, and -1.3dB to create 5 catch tone difficulties, using a digital audio editor (Audacity v2.0). The difficulty was set individually for each participant (see discrimination (A)). Participants used a keyboard to produce tones, the ‘S’ and ‘L’ keys for left and right hand respectively. Tones were presented monaurally via headphones at 60dB SPL. Stimulus presentation was controlled using E-Prime (Psychology Software Tools, Inc.) on a computer which also sent markers to a second computer recording EEG, using PyCorder (Brain Products GmbH.).

2.3 Design and procedure

The study consisted of different parts. First, a discrimination (A) and a practice block (B). The main experiment started after the practice block. It consisted of two types of acquisition block (action-effect acquisition blocks (C) and tone sequence acquisition blocks (D)). These acquisition blocks were interleaved with experimental blocks (E) in the following order: CCCCDCCEDEEDCEE. Each of these blocks will be described below. The EEG cap was mounted after the practice block (B). In short, in the main experiment, each testing phase, where the participant attended to one ear (i.e., experimental blocks, E), was preceded by an acquisition phase where the participant built up action effect associations (acquisition phase, C/D).

Prior to the experiment, the participants performed the discrimination task (A) meant to introduce participants to the difference between experimental and catch tones, and find a suitable difficulty
level to perform the catch detection task. They were instructed to fixate on a central cross. Each trial started with the increase in size of the fixation cross (50 ms in duration) to highlight the start of a trial, after which participants were required to press either the left or right key. They were instructed to make a decision on each trial whether to press left or right key. 200 ms after the onset of the key press a tone appeared (200 ms SOA). If the tone was a catch trial then they responded by pressing one of the keys again. If no key press was made within 1400 ms then the trial terminated. If it was an experimental tone, then no response was required, and the next trial started after a random inter-trial interval (ITI) of 2000-2500 ms (see Figure 1 for a schematic view of trial events). The discrimination of catch and experimental tones became increasingly more difficult, with five levels. Each level consisted of 8 catch and 8 experimental tones, presented in a random order. The accuracy was calculated after each difficulty level and when the participant scored below 75%, then the previous difficulty level was adopted in the experiment. Ten participants ran the experiment (and practise block) with discrimination difficulty of 4 (1 being easiest and 5 hardest to discriminate between catch and experimental tones), five had difficulty 3, and two had difficulty 2.

The practise block (B) consisted of 48 trials divided into two short-blocks of 24 trials (16 experimental tones and 8 catch tones). The practise block was simply a shortened version of the experimental block (E). The trial sequence was the same as the discrimination task (A) with the addition that prior to the start of each short block ‘Attend left’ and ‘Attend right’ was presented on the monitor. This required participants to attend to only one ear and only respond to catch trials presented in that ear, whilst ignoring all tones, including catch tones, in the other ear. Simply put, the participant pressed a key which produced a high or low frequency tone. The tone was either presented to the attended or unattended ear and they were to respond to catch tones in the attended ear only (see Figure 1).

---

1 The actual discrimination performance in the experiment was near ceiling. This increased accuracy was most likely due to practise (see results section 3.1).
Each action-effect acquisition block (C) consisted of 32 trials, including only experimental tones. A left press always produced a low tone, in the left ear on 8 trials and in the right ear on 8 trials. A right hand press always produced a high tone, 8 to the left ear and 8 to the right ear. This assignment of tone frequency to left/right button presses was counterbalanced between participants. The same trial sequence was used in the acquisition block as described above in discrimination task (A), apart from that no response was required (no catch tones presented) but there were only experimental tones.

The tone sequence acquisition blocks (D) consisted of a series of 4-7 high and low tones with a random inter stimulus interval (ISI) of 500-700 ms. The participant first listened to the tones and then reproduced the sequence of high and low tones by pressing the left and right keys (regardless of which ear the tones were presented in). For each tone sequence acquisition phase, the participant reproduced 8 sequences, adding up to 45 key presses in total. After each sequence had been reproduced an accuracy score was presented.

![Figure 1](image)

**Figure 1.** Schematic representation of events in a trial in the test phase. At the start of a short block of 19-21 trials participants were instructed to which ear to attend (left or right). There were 4 short blocks within a full experimental block (E), there was a short break in between experimental blocks. The enlargement of the fixation cross indicated that the participant was to press either the left or right key. This action produced a tone and if it was a catch tone then a second press was required (using either hand). In a random order, all tones were equiprobable to appear in the left or right ear, and of high or low frequency. SOA = Stimulus onset asynchrony, ITI = Inter-trial interval.

Each experimental block (E) consisted of 80 trials, 64 experimental (p=.80) and 16 catch trials (p=.20). High and low, and left and right ear, tone presentation was equally counterbalanced in each
experimental block. The trial sequence was the same as in the discrimination phase described above (see also Figure 1). The 80 trials were divided into shorter blocks of 19-21 trials each (catch trials varied between 3-5 in each short block to avoid anticipation of the number of catch trials in a block). Prior to the start of each short -block of approximately 20 trials ‘Attend left’ or ‘Attend right’ was presented on the monitor. This required participants to attend to only one ear and only respond to catch tones presented in that ear, whilst ignoring all tones, including catch tones, in the other ear. In the experimental block the presentation of predicted congruent and incongruent tones was equally weighted, as were attended and unattended tones. The motor prediction was based upon what action-tone association had been learned in the acquisition phase. Participants were told, before the first experimental block, that tones would be presented at random. It should be noted that equal weighting of the motor congruent and incongruent tones in the test phase may contribute to unlearning of the previously formed action-effect contingencies. For this reason, after every two experimental blocks re-acquisition was performed. In total each participant completed 6 experimental blocks which included 96 trials for each condition; attended/congruent, attended/incongruent, unattended/congruent, and unattended/incongruent.

2.4 Electrophysiological recording and analyses

Electroencephalogram (EEG) was recorded from 64 scalp locations (ActiCAP, BrainProducts GmbH). Amplifier bandpass was 0.1-200 Hz and digitisation rate was 500 Hz, online referenced to FCz. EEG analysis was conducted using Brain Vision Analyzer (BrainProducts GmbH). After recording, the EEG was digitally re-referenced to a common average. The offline filter was set at a bandpass of 0.1-40 Hz. Ocular artifact correction was conducted using independent component analysis. For each participant, separate eye-movement and blink components were extracted and corrected. EEG was epoched into 700 ms periods starting 200 ms prior and 500 ms post to stimulus onset. Baseline correction was performed on 200 ms prior to stimulus onset. Trials with artifacts (voltage exceeding ±80μV relative to baseline, at any electrode) were removed prior to EEG averaging. Epochs were
averaged separately for Attention (attended, unattended), and Motor prediction (congruent, incongruent). Only experimental trials were included in any ERP analysis and trials were participants responded were not analysed (catch trials).

N1 analyses

Effects on the N1 were analysed over the vertex (Cz, Picton et al., 1974; Teder et al., 1993). Repeated measures ANOVAs were conducted with the factors Attention (attended, unattended), and Motor prediction (congruent, incongruent). The first N1 time window (86-126 ms) was centred on the peak over Cz (106 ms, ±20 ms), averaged across all conditions. This time window was also used to analyse the polarity inverted N1 present at the scalp area below the Sylvian fissure, over the mastoids (Alho et al., 1986). The factors for mastoid analyses were Attention (attended, unattended), Motor prediction (congruent, incongruent), Hemisphere (ipsilateral and contralateral, to the ear receiving the tone), and Electrode site (T7/8, and TP7/8). The N1 has been proposed to be made up of several sub-components (Näätänen & Picton, 1987) and effects of attention have been suggested to be particularly present on the descending slope of the N1, rather than the N1 itself (e.g. Näätänen et al., 1978; Näätänen, 1982). Thus, effects of attention and motor prediction were also analysed on the downward slope of the N1 in this 50 ms time interval (126-176 ms), at Cz.

Analyses of N_{AE}, P3a, and later processing stages

To investigate later effects of action congruency and attention we used a cluster of Anterior, Central and Posterior electrodes. The time window following the N1 (176-300 ms) was chosen because it was meant to assess N_{AE} (Band et al, 2009) and also fronto-central P3 (P3a peak was 246 ms at Fz, averaged across all conditions). Two final time windows were used to investigate later sustained stages of processing (300-400 and 400-500 ms). The factors were; Cluster (Anterior, Central,
Posterior), Attention (attended, unattended), Motor prediction (congruent, incongruent). As there were not equal numbers of electrodes in each cluster, the mean activity across all electrodes within a cluster was analysed at this stage. Interactions with the factor ‘Cluster’ were followed up for each topographical area separately and Bonferroni correction applied, uncorrected probability values are reported but significance criterion was set at $\alpha=.017$. Moreover, to include whether any effects of motor prediction and attention were on particular electrodes within a cluster, any subsequent analysis of a single cluster also included Electrode as a factor (F1, Fz, F2, FC1, FCz, FC2 (Anterior); C1, Cz, C2 (Central); CP1, CPz, CP2, P1, Pz, P2 (Posterior)).

The Greenhouse-Geisser correction was applied to ANOVAs where the assumption of sphericity was violated and corrected probability levels, original degrees of freedom and epsilon values are reported. Partial eta squared ($\eta^2_p$) is reported as a measure of effect size.

3. Results

3.1 Behavioural results

In the unattended stream false alarms accounted for 6.7% of all unattended trials, that is, they successfully ignored 93.3% of all catch tones presented to the unattended ear. Participants missed only 2.3% of catch tones in the attended ear so were at ceiling when it came to detecting catch tones (97.7%). Moreover, their false alarm rate to non-catch tones was only 0.5%. These figures clearly indicate that participants performed the attention task accurately. The participants voluntarily selected which button to press on each trial (out of 480 trials per participant). There was no significant difference between left (average 243) and right presses (average 237) ($t(16)=1.3$, $p=.21$).

3.2 ERP results

12
3.2.1 Attention

There were two effects of attention in the N1 time window. First, there was a significant main effect of attention on the descending slope of the N1 with larger negative amplitude for attended over unattended tones ($F(1,16)=6.02$, $p=.026$, $\eta^2_p=.27$). Second, analysis of the polarity inverted N1 showed an Attention*Hemisphere interaction ($F(1,16)=5.31$, $p=.029$, $\eta^2_p=.26$) and follow-up analyses showed enhanced positivity for attended over unattended stimuli at ipsilateral hemisphere only ($F(1,16)=5.43$, $p=.033$, $\eta^2_p=.25$) (see Figure 2). Hence, the polarity inverted N1 over the mastoids showed a larger positivity for attended compared to unattended tones, for tones presented to the ipsilateral ear. The N1 peak analysis revealed no effect of Attention ($p=.52$).

Analysis of later stages of processing, 400-500 ms, showed a Cluster*Attention interaction ($F(2,32)=5.29$, $p=.024$, $n=.25$) $^2$. In subsequent analyses of each area there were a main effects of attention over posterior ($F(1,16)=19.57$, $p<.001$, $\eta^2_p=.55$) and central areas ($F(1,16)=8.64$, $p=.010$, $\eta^2_p=.35$) only, with larger positive amplitude for attended over unattended stimuli (see Figure 2). There was no effect of attention over anterior electrodes ($p=.20$). Moreover, analysis of the NAE and P3a, showed no effect of Attention in the 176-300 ms time interval at anterior ($p=.91$), Central ($p=.34$), or posterior electrodes ($p=.74$).

---

$^2$ There was also a Cluster*Attention interaction at 300-400 ms ($F(2,32)=7.94$, $p=.005$, $n=.33$). However, subsequent analyses for each cluster separately did not yield any significant effects of Attention.
Figure 2. Grand average ERPs for attended (black lines) and unattended (grey lines) trials at frontal (Fz) central (Cz), and Posterior (Pz) electrodes. Vertical line represents stimulus onset; left of this is the 200 ms baseline and to the right the 500 ms post-stimulus period. Top shows grand averaged ERPs for the mastoids, separately for stimuli presented contralateral (TP7/8c) and ipsilateral (TP7/8i) to stimulated ear. The middle graph shows a significant attention effect over the vertex (Cz), here on the descending slope of the N1, and at later stages of processing (400-500 ms). Bottom: Grand averaged ERPs for Fz and Pz, these represent a cluster of electrodes analysed for anterior and posterior sites. Shaded grey areas highlight a significant difference between attended and unattended stimuli (*p≤.05, **p≤.01, ***p≤.001).
3.2.2 Motor prediction

At anterior electrodes there was a main effect of motor prediction from 176-400 ms. There was a Cluster*Motor prediction interaction in the 176-300 ms (F(2,32)=6.30, p=.015, ε=.63, η_p^2=.28) and 300-400 ms time interval (F(2,32)=4.46, p=.035, ε=.69, η_p^2=.22). The effect, of larger positivity for incongruent compared to congruent stimuli, was present over the Anterior area only and was consistent in the two analysed time intervals: 176-300 ms (F(1,16)=10.72, p=.005, η_p^2=.40), and 300-400 ms (F(1,16)=7.85, p=.013, η_p^2=.33) (see Figure 3). In the 176-300 ms and 300-400 ms time interval there was no effect of Motor prediction over central (p=.62, and p=.70) or posterior (p=.24 and p=.38) electrodes, respectively. These results suggested a larger P3a for incongruent compared to congruent stimuli. There was no Motor congruency effect on the N1 peak (p=.82), or the descending slope of the N1 (p=.58). Analysis of the polarity inverted N1 showed a marginally significant Motor Congruency*Hemisphere interaction (F(1,16)=4.54, p=.049, η_p^2=.22), but follow-up analyses showed no effect of Motor congruency over contralateral (p=.15) or ipsilateral (p=.12) hemisphere.
Figure 3 Grand average ERPs for motor congruent (black lines) and incongruent (grey lines) trials at frontal (Fz) central (Cz), and Posterior (Pz) electrodes. Vertical line represents tone onset; left of this line is the 200 ms baseline and to the right the 500 ms post-stimulus period. Shaded grey area highlights the significant difference between congruent and incongruent at frontal electrode cluster (*p<.05, **≤.01). Topographical map insert represents the motor prediction effect (congruent minus incongruent). The right hemisphere represents activity ipsilateral to stimuli side, and left hemisphere shows the effects contralateral to stimuli presentation ear (amplitude range between –0.4 and 0.4 µV).

3.2.3 Attention *Motor prediction
Analysis of the 176-300 ms interval showed a close to significant Cluster*Attention*Motor prediction interaction ($F(2,32)=3.28, \varepsilon=.65, p=.075, \eta^2_p=.17$). Over posterior electrodes there was an Attention*Motor prediction interaction in the 176-300 ms time interval ($F(1,16)=8.43, p=.01, \eta^2_p=.35$) Follow-up analysis showed an effect of motor prediction at the unattended side only ($F(1,16)=13.83, p=.002, \eta^2_p=.46$) (see Figure 4). Taken together, there was a difference between motor congruent and incongruent stimuli when the stimuli were presented to the unattended ear. There were no Attention*Motor prediction interactions over anterior ($p=.76$) or central electrode sites ($p=.50$).

Figure 4. The top left graph shows grand average ERPs for all four conditions at CPz in the 500 ms time interval following stimuli onset. The shaded grey area corresponds to the time period showing a significant interaction (176-300 ms; **p<.01). The right two graphs represent the grand average waveforms for motor congruent and incongruent tones when the stimuli were attended (top) or unattended (bottom). Follow-up analyses indicated the interaction was driven by a significant motor prediction effect for tones presented at unattended side only, represented by grey shaded area, bottom right. The bars (A.) show the mean amplitude for each condition (with standard error bars), averaged over the 176-300 ms time interval and across all electrodes included in the analysed cluster of electrodes. The topographical map (B.) represents the significant effect in the 176-300 ms time interval, that is, the difference between motor predicted congruent and incongruent, for the unattended shaded area. The right hemisphere represents activity ipsilateral to stimuli side, and left hemisphere shows the effects contralateral to stimuli presentation ear (amplitude range between −0.5 and 0.5 µV).
4. Discussion

In this study we employed a new paradigm to manipulate motor prediction and attention orthogonally. We manipulated motor prediction by training participants with particular action-effect contingencies, which in the subsequent test phase were either validated (congruent) or violated (incongruent). At the same time, participants were required to respond to catch tones only when they were presented to a particular ear, allowing for the manipulation of spatial attention. This paradigm mimics genuine agent-environment interactions where action-triggered stimulus events are not only predicted or unpredicted, but also attended or unattended in the current context.

Our results demonstrate that the relationship of motor prediction and attention has both interdependent and independent aspects. We observed independent effects of attention in relation to the N1 as well as at later stages of processing (400-500 ms). We also observed differences related to motor prediction. A main effect of motor prediction was present at frontal electrodes (176-300 and 300-400 ms) with larger positivity for incongruent over congruent stimuli. Over and above these main effects reflecting independent attention- and prediction-related components, we also observed an attention and motor prediction interaction at posterior scalp regions around 176-300 ms after tone onset. We will discuss these effects in detail below.

The fact that we found prediction- and attention-related main effects, as well as an interaction is very telling. That we found an interaction of attention and motor prediction indicates that some aspects of stimulus-processing depend on both attentional and motor predictive processes. We will outline this interdependency below. However, the main effect of motor prediction at frontal electrodes (176-300 and 300-400 ms) clearly shows that the prediction as such is done independently of attentional processes. This is because without the prediction of the action effect there could not be a difference between predicted and unpredicted effects; the two types of action
effect are otherwise strictly identical. Since unattended tones show the same effect as attended
tones, it follows that action effects are predicted irrespective of whether or not they are attended.

**Effects of attention**

In our attention paradigm the participant fully attended to one ear and ignored all tones presented
to the unattended ear. Analysis of the behavioural errors suggested participants’ allocated attention
as instructed as the participants performed the task accurately; responding to 98% of attended
catch trials and successfully ignoring 93% of all unattended catch tones (see results section 3.1). Our
attention manipulation demonstrated two effects in relation to the N1. First, enhanced negativity for
attended over unattended tones of the descending slope of the N1 (see Figure 2), replicating well
established effects of attention (e.g., Näätänen, 1982; Näätänen et al., 1978). Second, an effect on
the polarity inverted N1 ipsilateral to the ear receiving the tones, with larger positivity for attended
and unattended stimuli.

The polarity inverted N1 is present at the scalp area below the Sylvian fissure, over the mastoids.
Although the polarity inverted N1 is less commonly investigated, attention effects have been
reported at these more temporal scalp regions in the N1 time window (Näätänen et al., 1978;
Woldorff and Hillyard, 1991; Teder et al., 1993). We did not find an effect of attention on the peak of
the N1, however, there was an attention effect on the descending slope of the N1. Whether
endogenous spatial attention enhances the amplitude of the N1 sensory component, or whether N1
attention effects are due to a superposition of sustained endogenous negativity (creating a negative
difference, Nd), is still debated and empirical evidence supports both views (see Näätänen and Alho,
2004 for a review). The latter view would suggest the attention effect is not on the N1 component
per se, but more likely to be present on the descending slope. These differences have previously
been attributed to the rate of stimulus presentation (Teder et al., 1993). If there is a rapid
presentation of tones with short stimulus onset asynchrony (SOA), then effects of attention may be
present on the sensory N1 component itself (e.g., Hillyard, 1973; Alcaini et al., 1994; Ozaki et al.,
2004). However, if there is a longer interval between tones (approximately 800 ms and longer) then the attention effect is not on the N1 component but starts on the descending slope instead (Schwent et al, 1976; Näätänen et al, 1978; Näätänen, 1982). The SOA between stimuli in the present study was, if at all comparable as an action was performed on every trial, very long as there was only one tone per trial. Thus, our results are in line with previous attention studies with the effect present on the descending slope on the N1 as SOA between tones was long. However, how including an action affects specific attention modulations and relates to the debate of N1/Nd effects of attention (e.g. Näätänen and Alho, 2004) would require further investigation. It should be noted that our effect did not continue in a sustained manner as common in pure attention studies using a continuous stream of tones (e.g. Talsma and Woldorff, 2005; Woldorff and Hillyard, 1991). We did however find a late sustained main effect of attention (400-500 ms) over central and posterior electrodes with larger negativity for unattended over attended tones. The typical effect would be sustained negativity for attended over unattended (as our N1 effect 126-176 ms), over more fronto-central electrodes. A recent paper by Gomes et al. (2012) demonstrated an enhanced negativity similar to the one we observed in the current experiment, namely for unattended stimuli at posterior areas (although not analysed) in the time window (approximately 300-500 ms). Taken together our effects of attention fit well with previous manipulations, thus further validating our behavioural observation that participants successfully followed instructions and attended to one side. Our findings also suggest that the effects of attention on sensory processing are at least partially independent from motor predictive processes (see also Timm et al., 2013, for independent effects of attention and sensory suppression using a paradigm contrasting externally vs. self-generated stimuli).

Effects of motor prediction

In line with a forward model of action control (e.g., Miall and Wolpert 1996), we demonstrated stimulus processing differences depending on whether or not the tones were predicted by the action, based on the bi-directional action-effect associations established in the acquisition phase.
The main effect of motor prediction was seen over a cluster of frontal electrodes in the period of 176-400 ms. Specifically, there was larger positivity for incongruent over congruent tones (see Figure 3). The topography and timing of this effect may indicate that incongruent tones generated a larger P3a compared to congruent tones. The P3a, with fronto-central scalp topography, has been suggested to reflect a switch of attention towards deviant stimuli (Näätänen, 1990; Friedman et al., 2001). In a study by Waszak and Herwig (2007) participants also learned action-effect associations in an acquisition phase and the ERPs to tones were compared in a test phase. They found deviant/incongruent tones to produce a larger P3a. However, their incongruent tones were also unexpected in terms of low probability, whilst the incongruent tones in the present study were, in the test phase, as likely to appear as the congruent tones. This demonstrates that a P3a to an event might not simply be related to deviance in terms of frequency as in for example, an oddball task, but simply elicited if incongruent with an action. As such, the P3a might reflect an aspect of prediction error, regardless of the specific source of this error.

Notice that, if the component we observed in the current experiment corresponds to a P3a and if it reflects allocation of attention to a deviant, unpredicted stimulus, this attention must be of a different type than the attention manipulated in this experiment. This is because the effect did not differ between attended and unattended trials. The type of attention triggered by the unexpected event is rather reactive in nature. It is allocated once an unforeseen event occurred. This type of attention is different from the one we manipulated in this study, namely proactive attention that is allocated in expectation of an upcoming event. The distinction between proactive and reactive processing has been suggested to involve different mechanisms and engage partly different neural networks (see Braver, 2012 for a recent review of the dual mechanisms framework).

However, we do not want to make a strong point concerning the function of the predictive component we observed in our study. More importantly, the fact that congruent vs. incongruent
trials showed the same effect, irrespective of whether they were attended or not, demonstrates that motor prediction takes place not only for attended action effects, but also for ignored action effects. One may consider this finding to be surprising. However, it is not only attended action effects that indicate whether the on-going agent–environment interaction is smooth or perturbed. Attended action effects might be as diagnostic as unattended effects. Thus, in order to monitor as much of the perceptual input as possible for action effects that do not meet their expectations humans should predict unattended action effects as well.

We observed no modulation of the auditory N1 component as a function of motor prediction. As highlighted in the introduction, previous studies reporting attenuation of N1 amplitude have typically compared self-generated with externally generated tones, which differ in more ways than action-prediction (e.g., temporal prediction, motor potentials; see Hughes et al., 2013a). Sensory attenuation of the N1 demonstrated in previous studies may have been amplified by confounding variables when contrasting internally with externally generated actions. However, some previous studies have observed prediction-specific N1 modulations in motor prediction paradigms (e.g. Hughes et al., 2013b), sensory prediction paradigms (Todorovic and de Lange, 2012), and speech production paradigms (e.g., Houde et al., 2002; Heinks-Maldonado et al., 2005). As such, the absence of any prediction-related N1 modulation in the present study is somewhat surprising. One possibility is that the prediction effects were smaller in the current study because tones were no longer predictable during the test blocks. However, the fact that we do observe later prediction-related differences, along with the fact that previous studies report behaviour effects of prediction using the same contingency manipulation (Cardoso-Leite et al., 2010; Roussel et al., 2013) strongly suggests that participants did make action predictions in the test blocks. One possibility is that the later prediction-related effects observed in the current study reflect mismatch processing rather than prediction per se. Future research should attempt to determine the extent to which early N1 modulations depend on the precise prediction contingencies and the degree to which early and late
ERP differences might reflect match and mismatch processes respectively (see Bendixen et al., 2012 for recent review on this topic).

Attention and motor prediction interaction

As discussed above, attention and motor prediction revealed dissociable ERP correlates. Interestingly, we also observed an interaction between these two processes at posterior electrodes around 200 ms after tone on-set (see Figure 4). This interaction consisted of a difference between predicted congruent and incongruent tones for unattended stimuli, whilst no difference was present for attended stimuli. This finding seems counterintuitive at a first glance. However, note that it is only the process reflected in the ERPs at posterior electrodes around 200 ms after stimulus onset that is present exclusively for unattended stimuli. This does not hold for the motor prediction as such. As discussed above, the main effect of prediction clearly shows that attended stimuli were predicted as well. We have two tentative explanations for the interaction at hand. First, the process reflected by the component in question does not need to take into effect if the stimulus is already attended. One possibility explaining the motor prediction effect at the unattended side could be in terms of a modulation of the action-effect negativity ($N_{AE}$). This component was described by Band, et al. (2009) in a probabilistic learning task and is thought to reflect a mismatch between expected and unexpected action-effects. Band et al. (2009) showed that this $N_{AE}$ was similar to a feedback-related negativity ($N_{FB}$; Miltner et al., 1997). That is, they proposed that the automatic processing of irrelevant action-effects were similar to intentional processing or feedback. Importantly, their $N_{AE}$ showed a similar latency and topography (maximal at Pz) to the interaction observed in the present study.

The experiment from Band and colleagues shows that a mismatch between anticipated and actual action effect results in an $N_{AE}$, which, analogous to the $N_{FB}$, probably reflects processes related to outcome evaluation and associative learning. In the experiment of Band et al. sensory action effects were not task-relevant, just as the unattended stimuli in our experiment. We suggest that attended
stimuli might undergo evaluative processing reflected in the $N_{AE}$ regardless of whether they are predicted or not, whereas unattended stimuli are only evaluated when they invalidate the prediction.

Our second tentative explanation refers to the pre-activation account of sensory attenuation (Waszak et al., 2012). This model suggests that motor prediction increases the mean level of activity in the network representing the expected effect to some pedestal level. Sensory attenuation is due to discrimination of the signal activation from the pedestal level being more difficult than discrimination of the signal activation from the baseline level (e.g., in the incongruent trials). Since the neural response saturates, the influence of pre-activation (i.e., the difference between predicted and unpredicted stimuli) decreases with increasing stimulus strength (Roussel et al., 2013). Given that attention has been demonstrated to increase sensory processing (Moran and Desimone, 1985; Chelazzi et al., 1993; Desimone and Duncan, 1995), it might be that the effect of prediction on sensory processing in our paradigm is smaller for attended compared to unattended stimuli. Although it is unlikely that the posterior component occurring about 200 ms after stimulus onset reflects sensory processing, it might reflect an upstream consequence of it.

To fully understand the detailed nature of how motor-prediction and attention interact requires further investigation. These two anticipatory mechanisms have independently shown opposite effects on neural and behavioural processing. This paradox is not exclusive to motor-related prediction and attention but is currently also discussed with non-action expectation and attention, an area also poorly understood (see Summerfield and Egner, 2009, for a review). Future research may wish to address the question of how motor-prediction and non-action expectation relate, and further, how attention influences these processes. Moreover, it is important to note that future research has to take into account that attention and prediction can be experimentally manipulated in a number of ways that might differ in important functional aspects. In the present study we
presented participants with two stimulus streams and asked them to respond to catch stimuli in only one of the streams. Hence, participants selected by means of attention which stimulus is task-relevant. However, catch stimuli were defined as broken tones, whereas the motor-sensory prediction concerned the frequency of the upcoming tone. That is, the predicted feature itself was not task-relevant. In other words, there was no overlap in attended and predicted features. We believe that the fact that we found interdependent components using the current design demonstrates the tight connection of these two factors, as feature overlap should further increase the interdependency. Future research can easily test this notion by varying task-relevant and predicted stimulus features.

In summary, we used a novel paradigm to investigate how attention and motor prediction relate, two mechanisms known to affect how we processes stimuli when interacting with our environment. We demonstrated independent effects of both attention and motor prediction suggesting these two mechanisms can separately affect stimulus processing, indicating attention and motor prediction are not simply two sides of the same coin and that even unattended stimuli are predicted. Moreover, we also demonstrated interdependence of attention and prediction in that certain motor-congruency effects occurred only when stimuli were unattended. This may indicate that when we selectively direct our attention towards upcoming stimuli, whether the stimulus is congruent or incongruent with an action is of less relevance. The interaction is in line with recent theories of action effect processing, such as the pre-activation account of sensory attenuation.
Authors’ note

The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement no. 263067. Moreover, we gratefully acknowledge that the Paris Descartes Platform for Sensorimotor Studies (Université Paris Descartes, CNRS, INSERM, Région Ile-de-France) supported the experimental work presented here.
References


