# Thèse de Doctorat en Sciences Cognitives

Université Paris Descartes

Ecole doctorale Cognition, Comportement, Conduites Humaines

## Une théorie de la preactivation pour la prédiction des effets de l'action

# A Preactivation theory of action effect prediction

**Cédric Roussel**

Directeur de thèse: Florian Waszak

**Membres du jury :**

| | | |
|---|---|---|
| Yann Coello | Université Charles de Gaulle, Lille 3 | Examinateur |
| Thérèse Collins | Université Paris Descartes | Examinatrice |
| Andrea Kiesel | Würzburg Universität | Rapporteur |
| Simone Schütz-Bosbach | Max Plank Institute | Examinatrice |
| Christopher Summerfield | Oxford University | Rapporteur |
| Florian Waszak | Université Paris Descartes | Examinateur |

**Date de soutenance :**

28 Novembre 2013

**Laboratoire de Psychologie de la perception**

45, rue des Saints Pères

75006 Paris, France

# Abstract

My PhD thesis aimed at contributing to the understanding of the neural implementation of action effect prediction. It has been suggested that the prediction of action effects is based on the preactivation of the sensory pathway involved in the treatment of the predicted effect (See Chapter I.H). Based on this suggestion we conceived a model of this preactivation hypothesis and derived a number of predictions about the perceptual processing of action consequences. In this thesis we tested the predictions made by the model (see Chapter II) as well as refined the model.

# Résumé

L'objectif du présent doctorat fut de contribuer à la compréhension des mécanismes de prédiction des effets de l'action en termes d'implémentation cérébral. Il a été suggéré que la prédiction des effets de l'action reposait sur la préactivation du réseau sensoriel impliqué dans le traitement de ces effets (voir Chapitre I.H). A partir de cette suggestion nous avons élaboré un model de cette hypothèse de preactivation nous permettant de dériver un certain nombre de prédictions quant au traitement perceptuel des conséquences de l'action. Au cours de cette thèse nous avons testé les prédictions faites par le model de la préactivation pour la prédiction des effets de l'action ainsi qu'en raffiner sa modélisation.

# Acknowledgements

First of all I would like to thank Florian Waszak who offered me the opportunity to write this Phd thesis. I also want to thank him for his kindness, his patience, his trust and his help at every step of my work, even in the most difficult time of my Phd studies. I'd like him to know that I am more than extremely grateful and that I will never forget these three years spend at his side.

I am also grateful to the member of my team, Andrea Desantis, Frederik Allenmark, Gethin Hughes, Yi-Fang Hsu, Alexander Jones and Karolina Moutsopoulou, for the interesting discussions, the help provided, and the very pleasant interactions we had, both on the professional and the social plan. I am especially grateful to Andrea and Frederik with whom I had sustained and very pleasant collaborations.

I am grateful to the Laboratoire de Psychologie de la Perception, which provided me a nice working environment, full of great researchers and great students, always available for discussions to get a better understanding of different issues in cognitive sciences or simply to share personal experiences.

I am grateful to my coworkers and friends, Andrea, Sergiu, Florian (both), Frederik, Marianne, Martin, Karolina, Pierre, Yi-Fang and others who made these three years a really pleasant time.

I am also grateful to my personal life friends and my family, who supported me during this challenge. I'd like to thank in particular Franck, who back in freshman year drove me to admit that knowledge was fun at that there was no shame to like studying.

Last but not least, a special thanks to my love, Marianne Fages who supported me the last few years, and who, I hope will keep sharing her life with me for many, many years.

# Summary

Table des matières

*Prologue*

> "*Give me a dozen healthy infants, well-formed, and my own specified world to bring them up in and I'll guarantee to take any one at random and train him to become any type of specialist I might select – doctor, lawyer, artist, merchant-chief and, yes, even beggar-man and thief, regardless of his talents, penchants, tendencies, abilities, vocations, and race of his ancestors. I am going beyond my facts and I admit it, but so have the advocates of the contrary and they have been doing it for many thousands of years.*" J.Watson

It is difficult to tell a story without also sketching its context. That is why I start with a citation from one of the fathers of cognitive psychology: J.Watson. I choose this citation for two reasons. First, because I see in it the expression of the author's desire for equality. In the author's mind equality between humans can be obtained through the plasticity of cognitive systems. Second, the citation brings us back to the core meaning of cognition. The latin verb 'cognoscere' refers to the abstract acquisition of knowledge, the getting to know, the ability to comprehend. According to J.Watson this ability to comprehend is shaped through the interaction between the agent and the environment. In his view, cognitive systems are designed by their constant interaction with the environment. These interactions and the mechanisms they involve are precisely the topic of the Cognitive Sciences. Perception, action, communication, memory, reasoning, and emotion are all part of this interaction.

The most essential components of a cognitive system are perception and action. Every cognitive system, even the most basic one, has the ability to interact with its environment by detecting changes in it and reacting to these changes. This reaction can be either a modification of the state of the organism (for example the depression of the immune system after a stressful stimulus) or of the environment (for example building a shelter when rain is coming), or both. As an extreme position, one could say that every other cognitive function 'merely' helps optimizing this basic adaptive function of the perception-action cycle. This prologue is meant to underscore how general the interest of the research outlined below is.

# I. Introduction

# A. Perception-Action Continuum

Already behaviorism considered action and perception to be un-dissociable. As a matter of fact, classical conditioning cannot be described without both perception and action. Classical conditioning, introduced by Pavlov (1927), refers to the fact that, after being conditioned, a neutral stimulus (NS) can elicit a conditional response (CR). This conditional response shall be a reflex response or an unconditional response (UR) to an unconditional stimulus (US). After the systematic temporal contiguity of NS and US, NS will become a Conditional Stimulus (CS), such that the basic behavioral relationship US $\xrightarrow{yields}$ UR is extended to CS $\xrightarrow{yields}$ CR, by introducing a phase during which NS$\xrightarrow{yields}$US. Note that at the end CS=NS and CR=UR.

In other words, in cognitive systems the natural response to a stimulus can be transferred to another stimulus. This demonstrates two things. First the plasticity of the cognitive system, and secondly, the fact that even a reflex action can be modulated by the central nervous system, or at least the central nervous system is able to extend a reflex to stimuli which were not initially supposed to elicit that reaction. It appear that through its associative nature the cognitive system is optimized to adapt its response to others stimuli.

However, the link between perception and action exists also from action to perception. Thorndike (1898), followed by Skinner (1937) and many others, introduced a paradigm in which action-effect relationships are crucial. Instrumental, Operant or Skinnerian conditioning refers to the fact that if an action outcome is rewarding for the organism then the behavior triggering the reward will increase in frequency over time. In this paradigm, the targeted behavior is not a reflex but a spontaneous action. That is to say, by manipulating the contingencies between a spontaneous behavior and a rewarding stimulus, it is possible to control the probability of the behavior's occurrence.

Behaviorism studied mainly the influence of action effects (in this case the reward/punishment) on the occurrence of the behavior. Thus the conditioned operant behavior appears to be goal-directed and share more characteristics with voluntary behavior since the behavior is maintained by its consequences, whereas a classical conditioned response shares more characteristics with reflex or automatic behavior since it is maintained only by the association established between the neutral and the unconditional stimulus (Domjan, 2003).

Of course it would be wrong to draw a strict line between both action types, especially so given the complexity of the human cognitive system. However, despite the difficulty to precisely and exhaustively define the concepts of "voluntary action" and "stimulus-driven action", cognitive neuroscience tried to shed light on the commonalities and differences between these two types of actions.

## B. Neurophysiological approach: TWO action networks

With the progress in imaging techniques scientists were able to investigate brain activity during voluntary and stimulus driven actions. A stimulus-driven action is a movement in response to an external event. These stimulus-driven actions will usually be responses issued as fast as possible to a stimulus. Hence, the event triggering the action is exogenous. Of course, the response (especially in human behavior) is actually done with the agents' consent, so it is actually rather misleading to call it "involuntary". However, usually stimulus-driven actions are highly trained actions. That is, they are automatic in that they need less processing than untrained actions.

In contrast, "voluntary" actions are induced by an endogenous cause. That is, they are triggered by some internal decision to act. Here, again, the term "voluntary" is actually misleading when used with respect to actions investigated in the laboratory, since the participants usually receive a number of instructions concerning which actions to perform, and when.

A number of paradigms demonstrated the existence of different but interrelated networks for Operant and Automatic action control (for a review see Haggard, 2008). The end point of both networks first appeared to be the primary motor cortex (M1) that projects directly to the muscles via the spinal cord. M1 seems to receive information from both the voluntary action and the stimulus driven path (Sherrington, 1906). By contrast, others areas involved in action control and selection tend to show preference for voluntary or stimulus-driven actions (Dum & Strick, 2002a). For example in case of voluntary action, M1 receives input from the pre-supplementary motor area (preSMA) and the supplementary motor area (SMA), which receives input from both basal ganglia and prefrontal cortex (Deiber et al., 1999; Jenkins, Jahanshahi, Jueptner, Passingham, Brooks, 2000; Picard & Strick, 1996). SMA and PreSMA have been demonstrated to show stronger activation for self-generated than for externally triggered actions. EEG studies have shown a long and increasing negativity beginning around 1 second or more before the onset of voluntary movements. This activity is thought to be caused by the activation of the preSMA/SMA complex (Lang, et al., 1991; Shibasaki & Halett, 2006; Yazawa et al., 2000), suggesting that this area might be one of the key areas of voluntary action genesis. The onset of this activity, usually called readiness potential, is considered to be the initiation of a cascade of neural activity causing action. As Figure 1 (panel a) illustrates, the voluntary action network is distributed across the brain and includes also areas other than M1 and the preSMA/SMA complex.

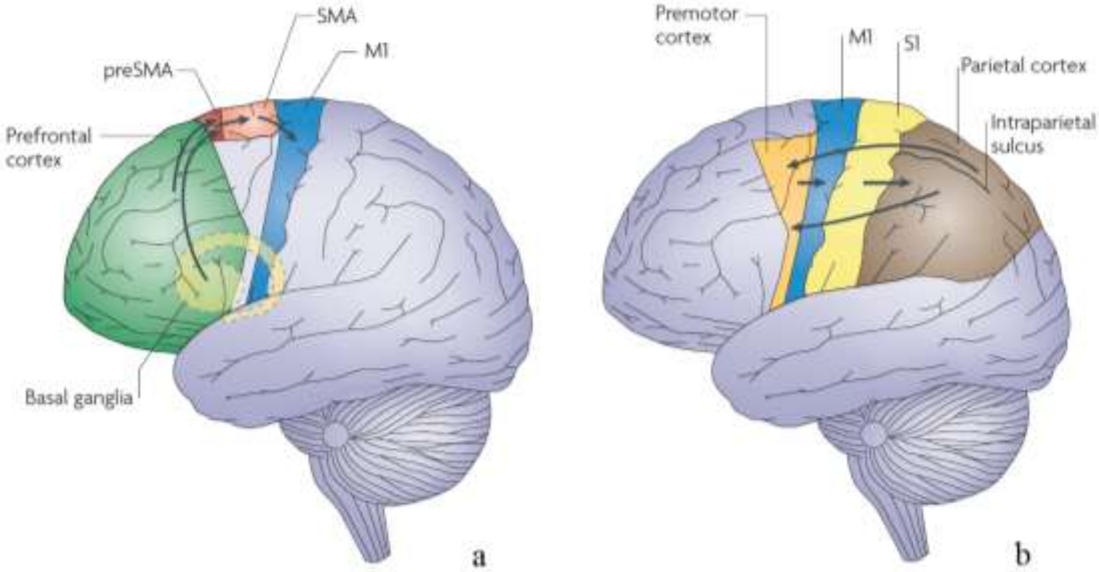

**Figure 1 :** panel a) Voluntary action network; panel b) automatic action network. From Haggard (2008).

As concerns stimulus-driven actions (see Figure 1 panel b), it is thought that processing takes place notably in sensory areas, the parietal lobe and lateral premotor cortex. This latter area then projects to M1 (Rizzolati, Lupino & Matelli, 1998).

Recent studies indicate that this dissociation might not be as strict as it first appeared. Notably, it has been shown that motor preparation during voluntary action can be transferred to stimulus-driven actions (Hughes, Schütz-Bosbach & Waszak, 2011). This might be taken to indicate that, just as behavior can be situated somewhere between fully automatic (such like the myotatic reflex) and fully "voluntary" (such like going to the cinema alone), action-related brain networks can be more or less involved during action preparation depending on the origin (endogenous or exogenous) and the complexity of the action.

However, despite the lack of a strict distinction between voluntary and stimulus driven action, the use of distinct concepts is useful. For example, experimental research demonstrated a number of effects which arise notably in the presence of voluntary action, notably sensory attenuation and intentional binding.

As will be outlined below, sensory attenuation and intentional binding are thought to be related to action effect anticipation. Recent studies suggest the involvement of two motor areas in the anticipation of action effects (SMA and M1). Haggard and Whitford (2004) showed that Transcranial Magnetic Stimulation (TMS) pulses disrupting the SMA abolish sensory attenuation. They suggest that the SMA is in charge of the generation of action-effect anticipation (see also Voss, Ingram, Haggard, & Wolpert, 2006). More recently, Hughes and Waszak (2011), using EEG, explored pre-movement differences between an action-to-effect and an action only condition, that is between conditions in which an action-effect was predicted or not, respectively. Rather than observing differences in the readiness potential (RP; a slow negative deflection in the EEG thought to reflect unspecific motor preparation in SMA observed 1-2s before action execution), the authors observed an increased lateralized readiness potential (LRP observed 700-400ms before action execution) for action-to-effect trials compared to action-only trials. The LRP is thought to reflect lateralized action-specific activation in M1 (e.g., Haggard & Eimer, 1999; Hughes et al., 2011; Hughes & Waszak, 2011; Leuthold & Jentzsch, 2002; Libet, Gleason, Wright, & Pearl, 1983), suggesting that the M1 might be involved in the prediction of action–effect.

# C. Specific effects in voluntary action

Two effects are typically associated with the execution/preparation of voluntary actions, namely sensory attenuation and intentional binding.

Sensory attenuation refers to the observation that action consequences are attenuated both in terms of their phenomenology and their cortical response (Gallistel, 1986; Schafer & Marcus, 1973). Sensory attenuation is commonly illustrated by the fact that it is impossible to tickle oneself (cf. Blakemore, Frith, & Wolpert, 1998; Weiskrantz, Elliott, & Darlington, 1971). Blakemore et al. (1998) investigated sensory attenuation using a robotic arm. Participants controlled this arm while it made tickling movements on themselves. They observed that the activity in the somatosensory cortex was reduced when the participant was in control of the robotic arm compared to the condition in which the robotic arm was controlled by the experimenter.

Sensory attenuation has also been studied using psychophysical paradigms assessing perception. For example, using subjective reports, Blakemore, Frith, and Wolpert (1999) showed that the introduction of a temporal or spatial distortion between the subjects' movement and the robot's reaction caused an increased feeling of ticklishness. The phenomenon of sensory attenuation has also been observed in other than the tactile modality. In the auditory domain, it has been observed that self-generated sounds are perceived as less intense than externally generated sounds (Sato, 2008). In the visual domain, using methods based on signal detection theory, it has been shown that predicted action-effects are less well detected compared to unpredicted action-effects (Cardoso-Leite et al., 2010). This is in agreement with neurophysiological findings showing attenuated cortical responses when stimuli are self-generated (Aliu, Houde, & Nagarajan, 2009; Baess, Jacobsen, & Schröger, 2008; Blakemore et al., 1998; 1999; Martikainen, Kaneko, & Hari, 2005; Gentsch & Schütz-Bosbach, 2011).

Thus, it appears that the neural processing of stimuli generated by a voluntary action is different compared to the processing of the same stimuli but externally generated.

In a similar way studies investigating intentional binding observed that events that are generated by a voluntary action are perceived as occurring earlier in time compared to the same events when externally generated (e.g., Haggard, Clark, & Kalogeras, 2002, for a review see Moore & Obhi, 2012). Haggard et al., (2002) demonstrated this effect by adapting the clock method originally used by Libet et al., (1983) to assess the perceived time of a conscious decision to move. In their study participants executed voluntary actions (key-presses) which triggered the occurrence of an auditory stimulus (a tone). Participants were then asked to indicate the time of occurrence of their action or of the tone by reporting the position of the clock-hand when these events occurred. This condition was compared to a condition in which the tone was preceded by an involuntary muscle twitch induced by Transcranial Magnetic Stimulation (TMS). Haggard and colleagues observed that voluntary actions were perceived as occurring later in time compared to involuntary movements. Similarly, the tone was perceived as occurring earlier in time when generated by a voluntary movement compared to when it was preceded by an involuntary movement. Accordingly action and its outcome were perceived as occurring close together in time. This intentional binding phenomenon has been replicated using other methods such as temporal order judgments (Cravo, Claessens, & Baldo, 2011), interval replication tasks (Humphreys & Buehner, 2009, 2010) and interval estimation (Engbert, Wohlsschäger, Thomas & Haggard, 2007; Engbert & Wohlsschäger, 2007).

## D. Predictive Mechanisms and voluntary action

During the last decade researchers aimed at investigating the neural and functional mechanisms underlying intentional binding and sensory attenuation.

It has been suggested that the difference in the processing of self vs. externally generated events depends on differences in the prediction of self-generated and externally generated stimuli (Blakemore et al., 1999; Wolpert & Ghahramani, 2000). Wolpert, Ghahramani and Jordan (1995) proposed a model of sensorimotor integration that involves inverse and forward models. Inverse models determine the appropriate motor command to achieve a goal given a particular sensory state; while forward models predict the expected sensory consequence of the action given the motor command (see Figure2).

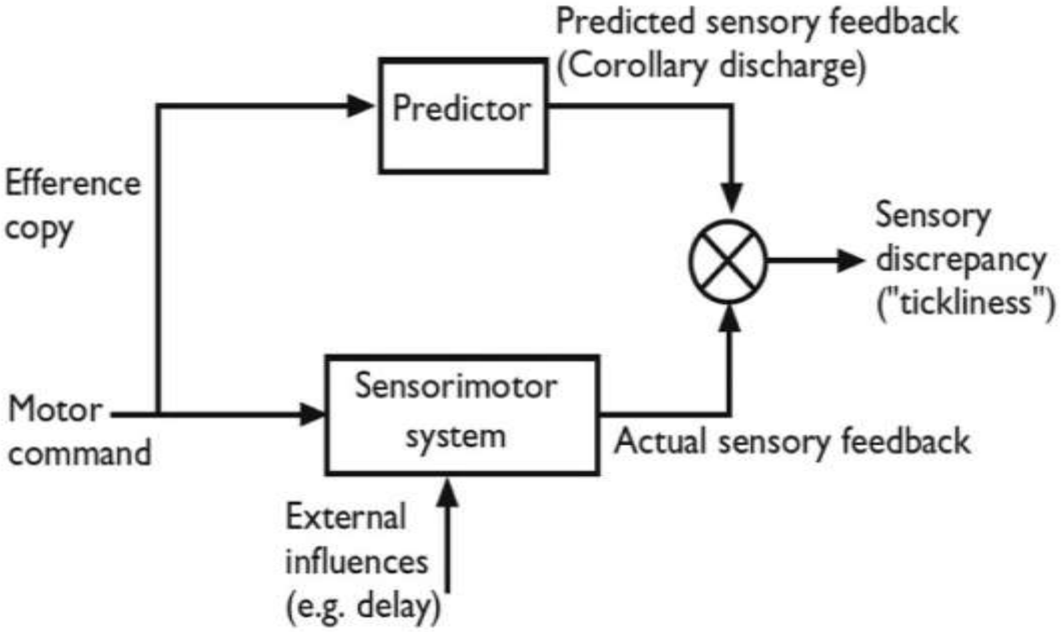

**Figure 2 :** Representation of a Forward model determining the sensory consequence of an action. When an action is performed the motor command generates an efference copy, this allows the systems to predict the sensory consequences of the action (corollary discharge). This prediction will be compared to the actual outcome of the action. If the prediction and the outcome matches, the prediction is cancelled from the experience, resulting in sensory attenuation. From Waszak et al. (2012) adapted from Frith, Blakemore and Wolpet (2000).

These models are similar to the concept of corollary discharge. The corollary discharge is a copy of the motor program (referred to as efference copy; see Figure 2) fed into the sensory network to predict the consequences of the action (for a review see Bridgeman 1995). One of the most popular examples of corollary discharge can be found in the mormyrid electric fish (Bell, 1989; Bell & Grant, 1989). This particular type of fish is able to generate an electrical field used for electrolocation and electrocomunication through an organ named electric organ discharge (EOD). It is also, able to perceive electrical fields. It has been shown that the cortical regions responsible to process the electrical afferences are activated at the time of the EOD motor activity through electric organ corollary discharge (EOCD). More evidence to support the fact that sensory systems respond to cues generated by an animal's own behavior has been provided, for example, by Poulet and Hedwig (2007). In humans, the ability of cognitive systems to predict action effects has been supported by a number of studies (Miall, Weir, Wolpert, 1993; Kawato, 1999; Wolpert et al., 1995, Blakemore et al., 1998, 1999; for a review see Desmurget & Grafton, 2000).

The use of these sensory predictions is threefold. First, it allows the system to prepare the next motor command given the consequences of the first in a sequential action process, such as playing the piano or video gaming (inverse model). Second, it allows the comparison of the prediction to the actual feedback resulting in a prediction error when they do not match (forward model). This error had been thought to be an important component to learning (Wolpert & Kawato, 1998) and the basis of sensory attenuation. For example, Blakemore and colleagues (1998, 1999) hypothesize that a match between the actual and the predicted consequence would result in a 'subtraction' of the prediction from the actual effect. This subtraction would be reflected in the decrease in observed neural activity, and the attenuation of the perceived strength of the stimulus Third the prediction of the sensory consequences will be compared to the goal for fine tuning of the motor command before the actual sensory feedback (feed-forward model; see Desmurget & Grafton, 2000). It has been proposed that Intentional Binding, too, may rely on the same internal model mechanisms (Haggard & Clark, 2003). However the predictive origin of intentional binding was subject to debate and until now remains unclear. According to Moore and Obhi (2012) the fact that intentional binding is specific to voluntary movements suggests the implication of motor based predictive processes. This conclusion is corroborated by a study from Haggard and Clark (2003) who investigated intentional binding in the case of disrupted motor intentions. To do so they applied, in some blocks, a random TMS pulse while participants made voluntary key presses. Doing so, they were able to disrupt the completion of intentional actions by triggering identical involuntary movements. They found that intentional binding was weakened in these trials. They concluded that the intention to produce the tone was necessary to evoke intentional binding. According to this view, intentional binding is a consequence of processes involved in the preparation of intentional action.

However more recently, further experiments investigating the perceived time of an action demonstrated that both predictive and retrospective mechanism are involved in intentional binding (Engbert & Wohlschläger, 2007; Moore & Haggard, 2008; Moore, Wegner, Haggard, 2009; Moore & Haggard., 2010). For example, Moore and Haggard (2008) investigated the contribution of prediction and retrospective inference, respectively, to the action component of intentional binding (i.e., the perceived onset of the action). Participants were instructed to report the time of the onset of their voluntary action on a Libet clock. In one condition, the key press generated a tone on 75% of trials (predictable condition). In another condition the key press produced a tone on 50% of trials (unpredictable condition). By comparing the perceived time of action in the trials without tone of the predictable and the unpredictable condition they were able to assess the contribution of predictability to intentional binding. They found an increased shift in the perceived time of action in the predictable condition, indicating the implication of predictive mechanisms in the effect. By comparing the perceived time of action in the unpredictable blocks between the trials with and without tone they could assess the retrospective contribution of the tone apparition to intentional binding. They found that the presence of the tone shifted the perception of the action's onset toward the tone, suggesting that the presence of the tone retrospectively influenced the perceived action time. The authors concluded that intentional binding emerges from both predictive and retrospective cues (see also Moore et al., 2009).

However, note that the study of Moore and Haggard (2008; see also Moore et al., 2009) concerned only the perceived time of action. It is, thus, unclear whether the same mechanisms are responsible for the shift of the perceived onset of the action's sensory consequence. So far the only conclusion that can be drawn from these experiments is that the perceived time shift of the action is influenced by both predictive and retrospective cues, while the mechanisms responsible for the perceived time shift of the action effect remains unclear (cf., Hughes, Desantis & Waszak, 2013, more detail about this topic will be provided in the general discussion).

Even if the role of motor prediction in intentional binding remains unclear, it appears that the prediction of the sensory consequence of the action plays a central role in voluntary action control. So far the forward/inverse model is the best candidate to account for the cognitive processes underlying voluntary action. Pacherie (2008) recently proposed to use the notion of forward models not only to explain sensory attenuation and intentional binding but also to provide a more general framework for a wide range of phenomena related to the "will".

# E. From the forward model to a more precise conception of voluntary action

Pacherie (2008) proposed a multifactorial model of action control to explain a large number of phenomena such as the sense of agency, free will, awareness of action goals, awareness of movements, sense of activity, sense of control, experience of authorship, experience of intentionality, and experience of mental causation. Most important in the context of the present thesis is how this framework relates the two effects outlined above (sensory attenuation and intentional binding) with a more general conception of how voluntary action is implemented in the mind, using the forward/inverse model conception. The concept of voluntary action implies that the agent performs a motor act caused by internal mental processes or ‘intentions’. Most philosophers working on the concept of intention concluded that it refers to multiple processes in action initiation, guidance and control (Searle, 1983; Bratman, 1987; Mele, 1992; Pacherie, 2007). The same holds for Pacherie (2008) who considers intentions to be hierarchically organized, and each level of intention to have a dynamics with two phases. The upstream dynamics results in the formation of the intention and the downstream dynamics result in its execution.

In Pacherie’s framework (2008) distal (D-)intentions are high level cognitive processes directed toward future events. On this level the upstream dynamics, e.g., the dynamics of decision-making that lead to the formation of an intention, involves the selection of a goal. In its downstream dynamics, D-intentions are concerned with the guidance and the monitoring of the action. They are maintained in prospective memory until they are no more necessary, and they ensure by means of the downstream dynamics the proper control of the action.

Proximal (P-)intentions often ‘inherit’ an higher-order action plan from a D-intention. Its role is to anchor the action plan in the current context, such as temporal anchoring or the decision to start acting. In the upstream dynamics they are formed by the integration of the conceptual information from the D-intention with the perceptual information of the current context and the behavioral possibilities of the agent. This dynamics results in a more concrete representation of the action. In the downstream dynamics, P-intentions ensure the guidance and the control of the action in the context of the action. In this framework both intentional binding and sensory attenuation are caused by the comparison of the predicted outcome and the actual feedback at the P-level.

The third type of intentions are motor (M-)intentions. These intentions refer to the motor representation of the action in terms of a precise neural code. Their upstream dynamics leads to the selection of the appropriate motor program. In their downstream dynamics the precise parameters of the motor commands are set and the movement is adjusted during its execution.

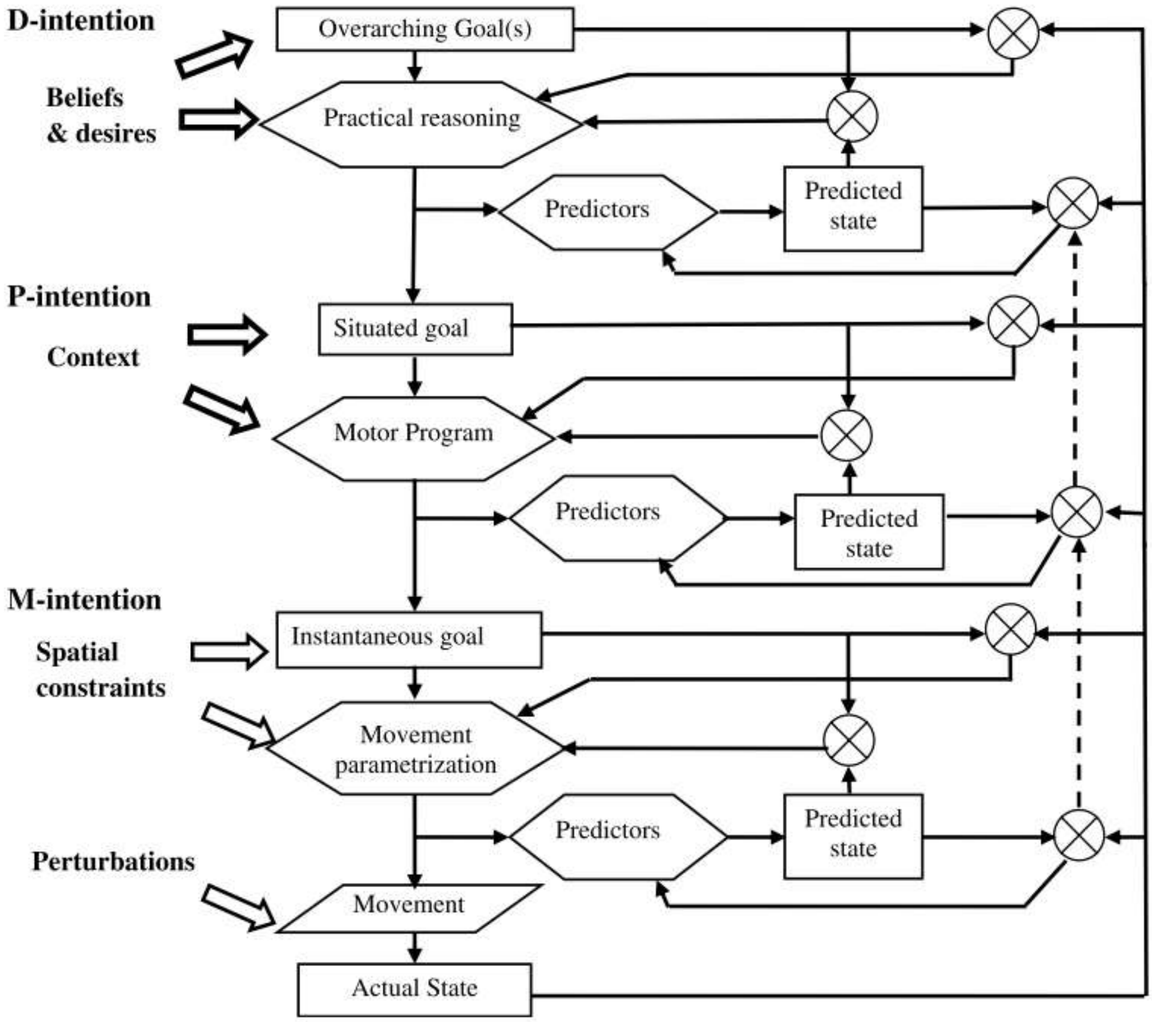

**Figure 3:** Hierarchical model of action specification based on the Inverse/Forward model framework. Each level of intention in the process of action specification functions as a forward/inverse model, in constant interaction with the others levels. From Pacherie, 2008.

To sum up, according to Pacherie's framework (2008), voluntary action may depend on a cascade of three levels of intentions. These three levels are tightly linked and are constituted by forward and inverse models (see Figure 3). That is to say, the precision of the representation of the goal on one hand and the action on the other depends on the intentional level. For example, if I sit in my living room and the sun is not shining yet, then I may form the intention to switch on the light. However at this stage the representation of my goal is abstract (I want more light), and so is my action representation. At the time of my decision to act my representations are already more precise. The kinematics of my movements are set as a function of the context (for example obstacles between the current position of my arm and the switch), and the goal representation includes what will happen in terms of a proprioceptive and somatosensory states. The execution of the movement involves a very precise representation of the muscles to use in order to achieve the goal. Along with this latter representation goes a precise prediction of the proprioceptive consequence of the action and a precise representation of the goal on a sensory level. This idea that the action's goals are at a certain stage represented on a sensory level is in line with the ideomotor principle, which states that actions are coded in terms of their perceptual consequences.

## F. The Ideomotor Theory

First introduced by German philosophy (Herbart, 1816; Lotze, 1852; Harless, 1861), the term ideomotor action has later been picked up by William James (1950). In his terms, it refers to the initiation of "movement upon the mere thought of it" (p. 522). James proposed that a direct relation between a desired "idea" and a movement are necessary to generate a voluntary movement. In other words, ideomotor action necessitates an association between actions and their consequences. Ideomotor actions, thus, are represented in terms of the actions' effect (for reviews see Shin, Proctor & Capaldi, 2010; Stock & Stock, 2004).

Greenwald (1970) was the first to address the notion of ideomotor action experimentally. Based on his experiments, he concluded that sensory consequence of an action is critical for action control. He suggested that three events are necessary in ideomotor action, the

Stimulus (S), the Response (R), and the Effect (E). According to Greenwald repetition of the sequence S-R-E results in an association between the events S and E, such that S will evoke the representation of E in a “conditioned anticipatory image of response feedback”. According to Greenwald the activation of this anticipatory image will, even in the absence of the stimulus, trigger the associated response. When the “anticipatory image” actually triggers the action then the action is not a basic response to a stimulus anymore but rather a goal directed action. This goal directed action relies on the fact that a stimulus or a context induces the activation of the anticipatory image of the consequence and, therefore, activates the appropriate motor code.

The ideomotor theory is also strongly reverberated in the Theory of event coding (TEC, Hommel, Müssler, Aschersleben & Prinz 2001). The TEC framework hypothesizes a common representational system of perception and action codes. This hypothesis is based on several assumptions such as: Perceived and to-be-produced events are represented in a common domain; actions are represented in a similar distributed fashion as perception; event codes are structured within several hierarchical levels (Hommel et al., 2001; for more detail see, Shin et al., 2010). In this framework, as in a more general view of the ideomotor principle, event codes for action and perception refer to the same distal features in the environment, coded in terms of the perceiver-and-environment relation. TEC and ideomotor theory, thus, state that the shared code between action and perception should be a reference to distal events in the environment (Prinz, 1997; Hommel et al, 2001). One might also think of this code as a modality-free representation of objects, comprising information about the object and the possible interactions that the system can have with this object. This concept of distal coding shares characteristics with the concepts of “affordance” (Gibbson, 1977) in the sense that action does not require the transformation of the contents of the perceptual analysis into actionable codes. This concept of affordance has been demonstrated experimentally in a number of studies (Grèzes & Decety, 2002; Grèzes, Tucker, Armony, Ellis, & Passingham, 2003; Tucker & Ellis, 1998).

There are several lines of evidence demonstrating that the cognitive system acquires bidirectional associations between actions and their effects and that these associations are used to initiate voluntary actions by the internal anticipation of the action’s outcome ( Elsner & Hommel, 2001; Greenwald, 1970; Herwig, Prinz & Waszak, 2007; Hoffman, Sebald & Stöcker, 2001; Kunde, 2001; Waszak & Herwig, 2007; Ziessler, 1998; see also the common coding theory, e.g. Hommel et al., 2001; Prinz, 1997). For example, Elsner and Hommel (2001) trained participants to associate left- and right-hand keypresses with the presentation of high and low frequency tones. In the test phase participants had to respond to the tones that had previously been the outcome of their actions. The participants who had to respond to the tone with a congruent mapping (e.g. with the same hand they used to generate the effect) were significantly faster and more accurate than the participant who had to respond with the opposite mapping. This highlights that an action can be influenced by a previously associated effect stimulus. More recent replications and extensions of this effect show that selection and execution or preparation of an action involves the internal anticipation of the specific sensory effect predicted by the action (Kunde, 2003; Pfister, Kiesel, & Melcher, 2010; Waszak & Herwig, 2007). These studies, therefore, support the idea that actions are selected by the internal prediction of the specific action outcome.

Some authors also suggested that an ideomotor mechanism could serve an evaluative function of the match between the action outcome and the intended effect (Band, van Steenbergen, Ridderinkhof, Falkenstein, & Hommel, 2009; Ziessler, Nattkemper & Frensch, 2004). This notion has already been mentioned above in the context of inverse/forward models (Wolpert, et al., 1995). To give one example that is rarely discussed in the literature, eliminating sensory feedback by destroying the sense organ that signals the leg position of insects does not result in the disruption of locomotion. However, a continuous distortion of the sensory feedback does (that is, when the organ gives wrong information about the leg position). This suggests that, once acquired, the anticipatory image of the action consequences is sufficient to induce action. But a systematical distortion of the sensory feedback results in drastic impairments of motor performance (Bässler, 1977, 1983, 1987).

One might go to such lengths as to claim that ‘intention’ is precisely this ability to select an action by activating a desired consequence. This relation between intention and the ability to select an action by activating the desired consequence can be found in Ziessler et al. (2004). In this experiment participant had to produce a visual stimulus with a particular key press (right or left for example) in response to a given target. Such that for example they had to respond to the target “W” with a right key press which will produce a “Y” on the screen. They learned different mapping between targets, responses and outcomes during the acquisition phases. During this acquisition phase the outcome of the response was either a specific letter either a specific digit (50% of the outcome of a right press response to a “W” was a “Y” and 50% was a “2”). During the acquisition phase, half of the participants were instructed to generate letters as outcome of their action, and the other half to generate digits. Such that for one group the outcome “letter” was considered to be an intended outcome while the outcome “digit” was considered to be not intentionally generated and reciprocally for the other group. In the test phase they tested Response-Effect learning by assessing the reaction time to the target flanked with distractors (for example for group 1 subject had to respond to “2W2” or “YWY”). They founded Response-Effect learning (e.g. facilitation in reaction time due to flankers) only for the instructed effect in the acquisition phase and therefore concluded that learning of action effects occurs only for the intended events. In other words action-effect learning appears to be related to intention.

## G. Intention and action-effect prediction

*This chapter is based on: Desantis, A., **Roussel**, C., & Waszak, F. (2011). On the influence of causal beliefs on the feeling of agency. Consciousness and Cognition, 20(4), 1211–1220.*

The study of Desantis, Roussel, & Waszak (2011) also supports the notion of a tight relationship between intention and the prediction of an action’s outcome. They manipulated participants’ prior belief of causality (belief of agency) and assessed whether this manipulation modulated intentional binding. Participants were made to believe that an auditory stimulus was either triggered by themselves or by somebody else (a confederate of the experimenter), although, in reality, the sound was always triggered by the participants. They used the Libet-clock method to measure intentional binding (Haggard et al., 2002; Libet et al., 1983). They predicted that intentional binding would be larger when participants are convinced that their action triggered the tone, compared to when they believe that somebody else’s action triggered the tone. As outlined above intentional binding is thought to rely on the prediction of the sensory consequences of the action (Haggard & Clark, 2003; Pacherie, 2008). An effect of prior beliefs on intentional binding would, thus, imply that intentional binding is a consequence of our prior beliefs about the causality between our action and their outcomes. Similarly to Ziessler et al. (2004, see above), where the intention to produce a stimulus is necessary to induce Response-Effect learning, here the intention to produce the tone (driven by the belief of agency) would be necessary to induce intentional binding.

The study that is outlined in more detail in the appendix (1) provides support for this hypothesis. Desantis et al. (2011) found that the perceived time of a tone was significantly shifted when participants were convinced that their actions triggered the tone, compared to when they were convinced that the confederate's action triggered the tone. Notice that both belief conditions matched in terms of temporal expectancy and action, the tone was always triggered by the participants' actions and the participant's actions were always identical.

As one possible explantation, Desantis et al. (2011) referred to the comparator model of motor control to account for their finding. According to the comparator model of motor control, internal prediction is used to provide internal feedback of the predicted outcome of an action which can be used before sensory feedback is available, thereby shifting earlier in time the perceived occurrence of action effects (Wolpert, 1997; Pacherie, 2008; Haggard & Clark, 2003). This would suggest that the prediction is somehow contingent on the causality belief, since in the experiment the perceived time shift occurs only when the subjects believed to produce the outcome.

According to the ideomotor theory action is driven and selected by its consequence. In such a scenario the same action could be generated by different predicted consequences depending on the subject's intentions. For example the key press may be driven by the tone representation in the subject belief condition since s/he intended to produce the tone, while this same key press might be only driven by proprioceptive and somato-sensory representation in the confederate condition. This claim is in line with the fact that Response-Effect learning is elicited only for instructed outcome (Ziessler et al. (2004) see previous part for more details). This argument is supported by the fact that no significant shift of the perceived stimulus onset-time was found in the confederate condition. As a matter of fact if the perceived time of the stimulus depends on the prediction of that stimulus, then no prediction about the tone was made in the confederate condition. See Appendix 1 for more details about this study.

## H. Outcome prediction: A stimulus specific preactivation?

According to previous findings the prediction of the sensory consequences of actions plays a central role in motor intentions and voluntary action control. However, the question of how the brain implements this prediction remains open.

It has been shown that the preSMA/SMA complex (see Chapter I.B), M1 and the cerebellum are key regions in action driven prediction of sensory action outcomes; however, very little is known about the content of the prediction and the way the brain produces it. As a matter of course, this question is tightly related to the question of how the external world is represented in the brain. It is commonly assumed that the brain possesses a neural code representing external events and objects. Obviously, the ability to predict an event might be the ability to activate the representation of this event before it actually happens. One question in this context is what is the neural code referring to the event.

It has been shown that the brain works in a modular fashion (Fodor, 1983). This modularity is reflected in the fact that higher-order cognitive functions (such as language, perception …) are subdivided into more or less autonomous sub functions. It has also been shown that these sub-functions, even at the single cell level, constitute an integration of the information, one example being the integration of retinal cells into on-off cells. As a matter of fact, if it is undeniable that neurons are tuned to respond to a limited range of stimulus, there is growing evidence that world events are represented by patterns of activation, rather than the mere response of localized cells. To give one example, Del cul et al. (2007) demonstrated that the conscious perception of a visual stimulus relies on activation in a loop of occipital, pre-frontal, and temporal regions.

According to the view that an object representation consists in the stream of activity it evokes from sensory to higher order areas, the prediction of a specific event/object might involve (at least partially) the activation of this stream from high order integrative areas to the sensory cortex. Hence according to the ideomotor theory, preparing or executing an action associated to an outcome should pre-activate the sensory pathway that represents the expected effect (Waszak, Cardoso-Leite & Hughes, 2012, see Figure 4).

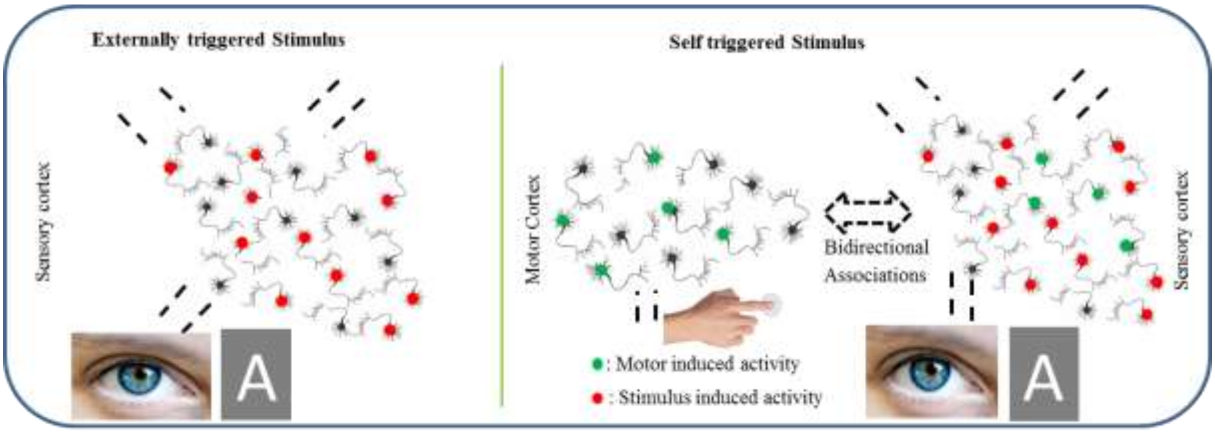

**Figure 4:** Schematic representation of the preactivation hypothesis through bidirectional associations. Left Panel) Activity induced by a visual stimulus. Right panel) The activation of a motor network will induce activity in sensory areas. When the stimulus is processed the activity globally increases in the sensory cortex.

This notion is supported by recent findings using neurophysiological methods. Kühn, Seunrinck, Fias & Waszak (2010), for example, harnessed the modularity of perceptual category representation in the human brain to assess action-induced activity in areas involved in the perception of particular classes of stimuli. They made participants acquire an association between left and right button presses and face and house stimuli, respectively, as action effects. During the test phase, participants continued to make left- and right-hand responses but no action-effects were presented. Nonetheless, they observed that activity in fusiform face area (FFA, which shows selective activation for faces, e.g. Kanwisher, McDermott & Chun, 1997) was increased for actions associated to face stimuli, whereas activity in the parahippocampal place area (PPA, which shows selective activation for places, e.g., Epstein and Kanwisher, 1998) was increased for actions associated to house stimuli. Note that these activations were observed in the absence of any visual stimulation. It is thus the action itself that induces activity in FFA and PPA. The results thus demonstrate that performing an action entails activity in perceptual areas as if the anticipated sensory action effect is actually perceived, corroborating the notion of action-induced preactivation in perceptual networks representing the expected sensory effects. (For a review of functional and neurophysiological mechanisms of effect anticipation see Waszak et al., 2012.)

Preactivation had also been suggested by Cardoso-Leite et al. (2010) to explain sensory attenuation. Along the lines of the ideomotor theory, they suggested that discrimination of the signal activation (of the real stimulus) from the pedestal level (due to this preactivation) is more difficult than discrimination of the signal activation from the baseline level (i.e., in a condition without effect anticipation, in which baseline activity is not raised to pedestal levels). This reasoning relies on the Signal Detection Theory (SDT) and will be explained in detail in the next part.

# II. Preactivation Theory

# A. Study 1: A model of neural Preactivation

*This chapter is based on: **Roussel**, C., Hughes, G., & Waszak, F. (2013). A preactivation account of sensory attenuation. Neuropsychologia, 51(5), 922–929.*

According to signal detection theory (Green & Swets, 1966; McMillan & Creelman, 1991) a sensory process transforms the stimulus energy into an internal response (or representation) and a decision process decides on each trial whether or not to consider the stimulus to be present based on the current value of the internal response and the decision criterion. The difference in the mean and standard deviation of the distributions of internal responses elicited in trials in which the stimulus is physically present ("signal" distribution) vs. in which the stimulus is physically absent ("noise" distribution) characterizes sensitivity (d’). The higher d’ (i.e., the further apart the two distributions are), the better the perceptual system is able to differentiate signal from noise. The decision or response criterion (c) is the value that the internal response has to exceed in order for the participant to report the stimulus to be present. The lower c, the smaller are the internal responses that the perceptual system still accepts as "stimulus present". Cardoso-Leite et al. (2010) assessed participants’ detection performance to stimuli (tilted Gabor patches presented at detection threshold in 50% of the trials) in three different conditions: The stimuli were triggered by an action involving the internal anticipation of a learned visual effect that is either 1) congruent or 2) incongruent to the to-be-detected threshold stimulus, or 3) neutral. Using signal detection methodology (Green & Swets, 1966) Cardoso-Leite et al. were able to test for the influence of the congruency between anticipated and actual action effect on the detection of the latter, separately for sensitivity (d’) and response criterion (c). They found that sensitivity (d’) was reduced in the congruent condition compared to the neutral and incongruent conditions, reflecting sensory attenuation. At the same time the response criterion (c) was identical in the three conditions. This suggests that internal action-effect anticipation truly affects perception. That is, signal and noise distributions overlap more with than without internal effect anticipation.

But how, precisely, does internal effect anticipation result in larger overlap of signal and noise distributions? One possible scenario is that the signal distribution draws closer to the noise distribution because the anticipated sensory effect is inhibited (see Figure 5, panel a). While many studies do not state a precise mechanism to explain sensory attenuation, some seem to reverberate the notion of a predictive inhibition as they state that the predicted effect is "subtracted" from the actual sensory effect or that the predicted effect is "cancelled" (Bays, Flanagan, & Wolpert, 2006; Bays & Wolpert, 2007; Blakemore, et al., 1998). Revealing in this context is also the fact that the effect is usually called sensory suppression.

However, Waszak, Cardodo-Leite and Hughes (2012) put forward a different scenario that can be directly derived from the ideomotor theory of action control (cf., Harless, 1861; James, 1890; Lotze, 1852) and, extending it, the common coding principle (e.g. Prinz, 1997, Hommel et al., 2001). The common coding principle claims that perception and action share a common representational code: Actions are coded in terms of the distal perceptual effects they evoke in the environment. Consequently, perceiving an action effect involves the same representation as performing the associated action and, conversely, performing an action involves the same representation as perceiving the effect it is associated to effects. In other words, performing an action results in the internal pre-activation of the sensory representation of the action’s expected perceptual consequence. According to this scenario, internal action effect anticipation increases the mean level of activity in the network representing the expected effect to some pedestal level (see Figure 5, panel b).

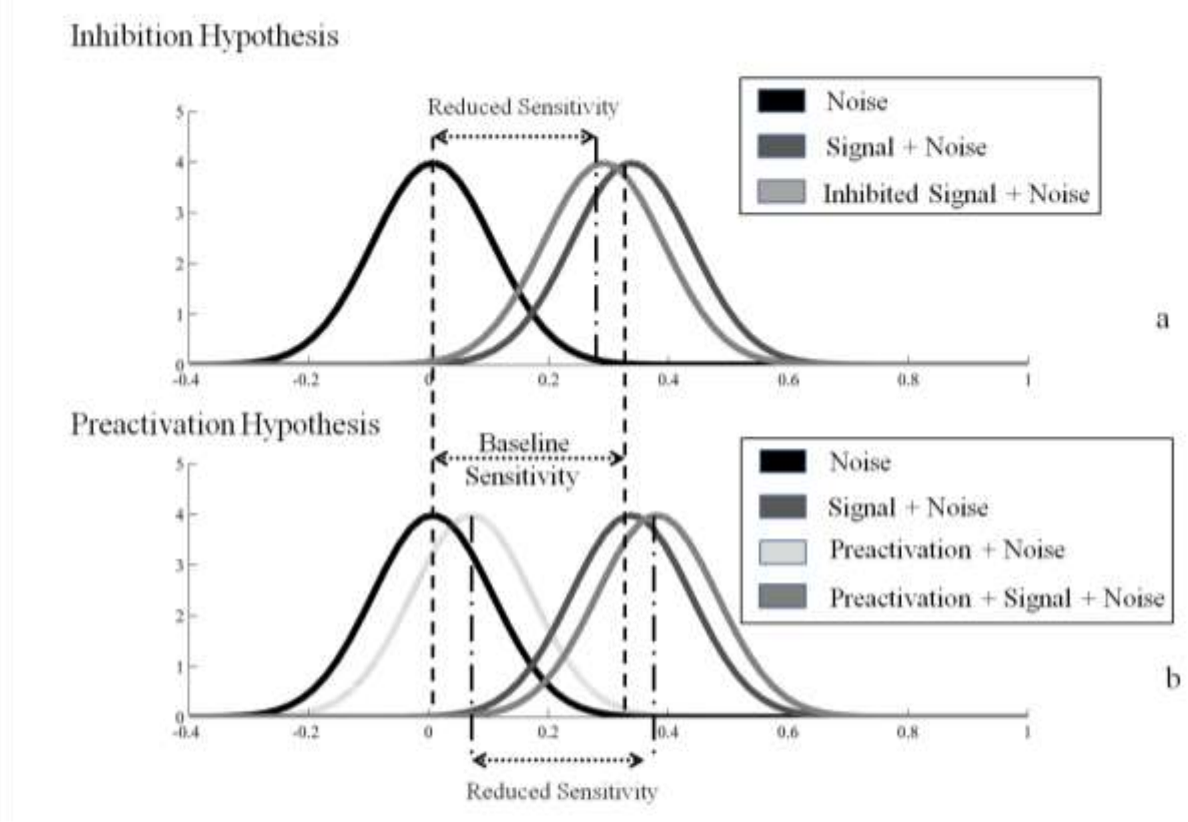

**Figure 5:** Panel a) Illustration of the inhibition hypothesis where a correct prediction decreases the internal response of the signal (dark gray to gray distribution). Panel b) Illustration of the preactivation hypothesis where the internal response increases more for the noise distribution (black to light gray distribution) than for the Stimulus distribution (dark gray to gray distribution).

In terms of the signal detection theory, under this latter account the sensitivity loss for congruent action effects 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 and the neutral conditions of the study of Cardoso-Leite et al. (2010), in which baseline activity in the neurons coding the action effect is not raised to pedestal levels). That is, according to the preactivation account, sensitivity in congruent trials is reduced because the mean of the pre-stimulus activity distribution is increased, not because the mean of the signal distribution is decreased.

To formalize this notion, we considered the shape of the neural response to stimulus intensity (or relative contrast) to be non-linear and saturating (Ohzawa et al 1985, Dean 1983, Albrecht et al 1984, Saul & Cynader, 1989a; see Figure 6). This response is usually fitted by the hyperbolic ratio function (Sclar et al., 1985; Heeger, 1992), $R = R_{max}\frac{c^n}{\sigma^n+c^n} + M$, but in the current example we used a Weibull repartion function,

$F(x;k,\lambda) = 1 - e^{-(\frac{x}{\lambda})^k}$, such as to avoid making any specific hypotheses regarding the maximum firing rate of a neuron or a population. In Figure 6 the parameters λ and k were set to 10 and 1, respectively. In this function we considered $x$ to be the neural entry and F(*x*) the associate neural response. Furthermore this function appears to be a good fit of the visual system's response to stimulus contrast (Burr, Morrone, & Ross, 1994; Reynolds, Pasternak, & Desimone, 2000). Based on the idea that the neural firing rate is a function of the entry stimulation (McCulloch & Pitts, 1943) we used this function as an integrator of the sum of the entry stimulation where $x = \sum_{i=1}^{n} x_i$, with *i* being the activation source. There are three possible activation sources: Noise (n), Stimulus contrast (c) or Preactivation (p)). That is, the network is either activated by noise, internal preactivation and/or stimulus contrast.

As the noise baseline response function we took a spontaneous internal response activity (internal noise) of about 1% such that $\mathrm{x_{noise}}(\mathrm{x_n})\,; 1 - \mathrm{e}^{-(\frac{\mathrm{x_n}}{10})^1}$=0.01. This noise baseline is represented by the x-axis in Figure 6. It reflects the network's mean activity when internal noise is the only source of neural activation. However, when a stimulus is presented, the neurons can be activated due to noise and the stimulus. The corresponding function $F(\sum_{i=1}^{n} x_i\,; 1{,}10) = 1 - e^{-(\frac{x_n + x_c}{10})^1}$, thus, integrates stimulus activity and noise baseline activity (dashed function in Figure 6). It represents the system's response to stimulus contrast without internal preactivation that is when it does not anticipate the stimulus represented by this particular set of neurons. In case of internal preactivation, $x_{preactivation}(x_p)\,; 1 - e^{-(\frac{x_p}{10})^1} = \omega$, the baseline is raised to a pedestal level (dashed horizontal line in Figure 6). When a stimulus is presented in this situation, the neurons can be activated by noise, preactivation and the stimulus contrast. The corresponding neural response function $\mathrm{F}(\sum_{i=1}^{n} \mathrm{x_i}\,; 1{,}10) = 1 - \mathrm{e}^{-(\frac{\mathrm{xc+xn+xp}}{10})^1}$ integrates stimulus contrast activity, noise baseline, and preactivation (solid function in Figure 6). Note that the function is shifted upward resulting in stronger internal responses for identical stimulus contrasts as well as faster saturation of the internal response (i.e., saturation at a lower contrast values). Note also that the increase in internal response from the “without preactivation” (dashed line in Figure 6) to the “with preactivation” function (solid line in Figure 6) decreases with increasing level of activity. That is, the higher the activity level, the smaller the increase in internal response due to preactivation.

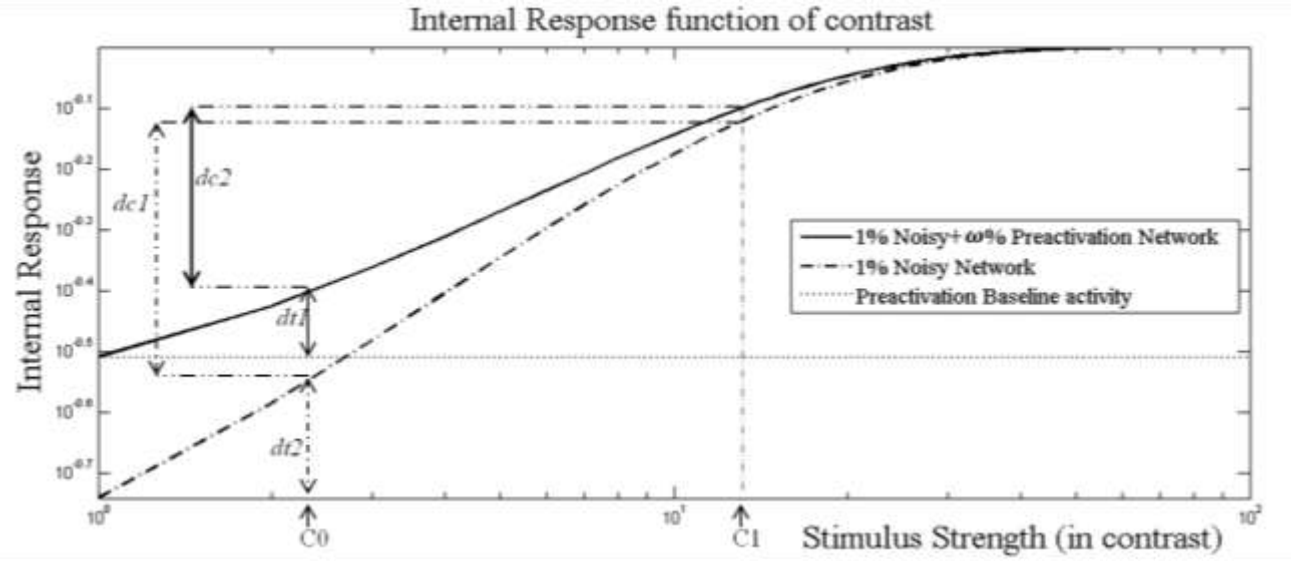

**Figure 6 :** Internal response as a function of contrast. S*ensitivity* is reflected in the distance between internal responses. C0: Weak stimulus contrast; *dt1*: Detection sensitivity when there is no preactivation; *dt2*: Detection sensitivity when there is preactivation for the same stimulus (C0). The difference between $x_c = 0$ (baseline activity) and $x_c = C0$ is smaller with than without preactivation ($dt1 < dt2$) due to a smaller stimulus-driven increase of the internal response when there is preactivation than when there is not. C1: Strong stimulus contrast; *dc1*: Discrimination sensitivity between C0 and C1 when there is no preactivation; *dc2*: Discrimination sensitivity between C0 and C1 when there is preactivation. The model predicts that, for any C0 and C1, discrimination is better without than with preactivation ($dc1 < dc2$).

As noted earlier, according to Signal Detection Theory (SDT) *detection sensitivity*, as studied by Cardoso-Leite et al. (2010), is represented by the difference between the baseline internal responses (in Figure 6 the x-axis line when there is no preactivation and the dotted horizontal line when there is preactivation) and the internal response evoked by the stimulus added to the baseline. As we can see in Figure 5b and 6 our model predicts this difference to be smaller when there is a preactivation in the network than when there is not. Hence our model predicts smaller sensitivity with than without preactivation due to a smaller stimulus-driven increase of the internal response when there is preactivation than when there is not (see detection *sensitivities* (*dt*) *dt1* and *dt2* for a stimulus of the same contrast (C0) in Figure 6).

In addition to this predicted difference in detection sensitivity our model predicts that the *discrimination performance* between stimuli of two different contrasts should also be reduced in the presence of preactivation compared to without preactivation (see discrimination sensitivities (*dc*) *dc1* and *dc2* for two stimuli of different contrasts (C0 and C1) in Figure 6). Furthermore, our model predicts that this reduced contrast discrimination for trials with preactivation should largely be driven by a change in the internal response of the weaker stimulus (stimulus C0 in Figure 6), since at higher contrast values (stimulus C1 in Figure 6), the difference between preactivation and no preactivation should be reduced or even abolished due to saturation. In contrast, cancellation accounts of sensory attenuation have been described such that "a cancellation mechanism that specifically affects self-generated input may nonetheless attenuate all self-generated input equally, irrespective of intensity" (Bays & Wolpert, 2007, page 30 line 28). In such a scenario contrast discrimination should not be reduced, since the perceptual distance between the two stimuli should remain the same if cancellation affects stimuli of all intensity equally. In our model the reduced effect of preactivation on the internal response at high stimulus strength, allows us to make a quite different prediction. As described above, discrimination sensitivity should be reduced since preactivation will influence the low contrast stimuli more than the high contrast stimuli. These basic features of our model are supported by a number of previous findings. Firstly, neural saturation and the non-linearity of the neural response to intensity (Ohzawa et al 1985, Dean 1983, Albrecht et al 1984, Saul & Cynader, 1989a, Nieder & Miller 2003) suggest that a linear increase in intensity translates to a non linear log-like increase in neural response. As such, an incremental increase in stimulus intensity will result in smaller increases in the neural response as intensity increases. Secondly perception of intensities has been seen to follow a Weber-Fechner law (Dehaene 2003 for numerosity, Gorea & Sagi 2001 for contrast) meaning that a given increase in stimulus energy will affect perception more if the stimulus is weak. Taken together these findings provide strong support for the basis of our experimental prediction, namely that a fixed amount of stimulus preactivation will influence the internal response to a greater degree for low contrast than for high contrast stimuli.

We tested these predictions by asking participants to perform a contrast discrimination task for stimuli that were congruent (trials where the preactivation matches the stimulus) or incongruent (trials where the preactivation doesn’t match the stimulus) with previously learnt action effect associations. Participants performed left- or right-hand voluntary actions on each trial that had previously been associated with the letters A and H respectively. In the test phase these stimuli were presented at one of two contrast values (C0 or C1) and participants were required to report the perceived contrast at the end of each trial on a 100 point scale, where ratings of below 50 were classed as C0 and ratings above 50 as C1. According to our model, congruent trials should result in reduced contrast discrimination compared to incongruent trials, with this difference largely determined by changes in the internal response for the weaker stimulus.

## a Materials and Methods

### ***Stimuli***

Experimental stimuli were generated and presented with Matlab 2007b using the psychophysics toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007). The stimuli were two white letters (A and H) displayed on a 15 inches CRT monitor at a 60hz refresh rate and fitted into a virtual square of 2.9 degrees of visual angle. The screen resolution was set at 800*600 pixels. These two stimuli were presented at two different contrast values (C0 and C1; see contrast determination phase) at the center of the screen. In the test phase we used a uniform noise texture to increase perceptual variance. This noise was a 100*100 matrix filled with an equal number of white and black pixels. The matrix was re-sampled (using the randperm function in Matlab) on each screen refresh. The mean luminance of the noise was then equal to the gray background.

### ***Contrast determination phase.***

In order to determine individual contrast values C0 and C1 yielding a discrimination d' of about 2, every participant completed a psychophysical staircase converging on 90% correct responses in a letter identification task (A vs. H). We employed an adaptive staircase to manipulate stimulus transparency using an accelerated stochastic approximation algorithm as described by Kesten (1958). The initial step size was 20 and we stopped the staircase when the step size was 1 (in transparency). The correct response rated was used to ensure that the stimuli were supraliminal and that we could then independently manipulate discrimination. We used this contrast value as the referential contrast C0 in a 2AFC paradigm (with constant stimuli ranging from C0 to C0+12%) in order to calculate, the contrast value of C1 yielding 85% correct responses in a luminance discrimination task (C0 vs. C1). For an ideal observer, this contrast yields a discrimination d' of around 2 (Mcmillan & Creelman, 1991). Pilot experiments showed this procedure to work reliably.

### ***Association phase.***

Participants fixated on a 3.3 degrees visual angle square located at the center of the screen. They were asked to press with their right/left index finger one of two keys (P and A on a standard French (AZERTY) keyboard), each key press triggering presentation of a visual effect (A or H). The key-letter mapping was counterbalanced across participants. The letters appeared 100ms after the key press at full contrast in the square at the center of the screen.

There were two types of association phase. First, in the free association ("FreeAsso") blocks the action sequence (left / right) was freely generate by the participants at a pace of about 1 key press every seconds for 50 seconds. In 5% of the trials the visual effect was a W. In these catch trials, the participant had to press both buttons within 1s of the appearance of the stimulus. Catch trials were meant to ensure that participants paid attention to the effect stimuli. Second, in memory association blocks ("MemoryAsso") random lists of As and Hs were presented to the participants (the average list size was 5). The lists were presented via headphones as spoken letters. After the lists were presented, participants had to reproduce the sequence by pressing the corresponding button sequence.

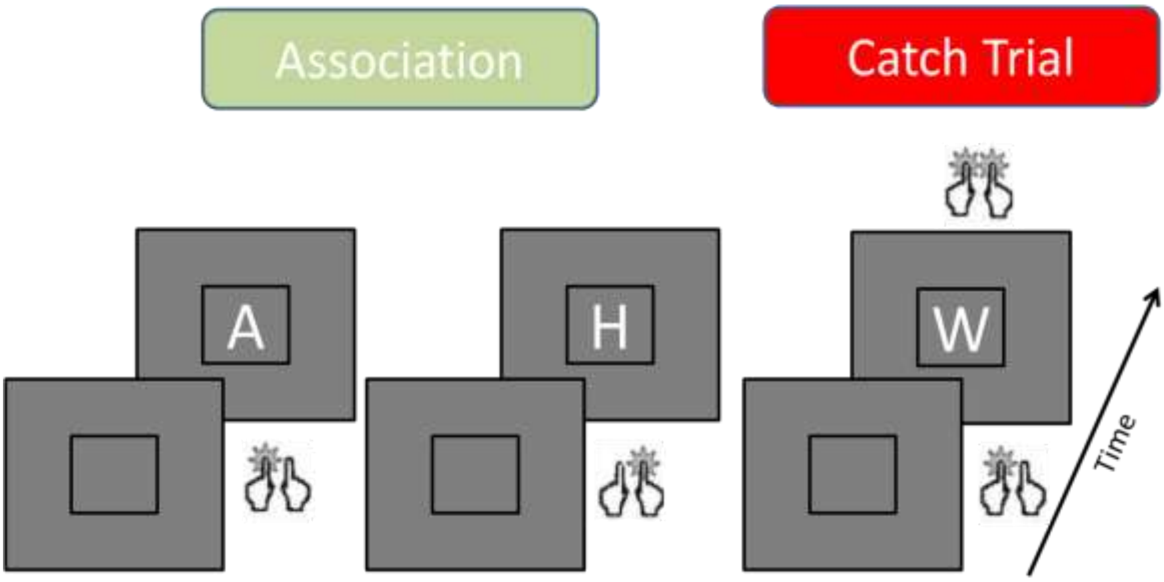

**Figure 7:** Representation of the “FreeAsso” association phase.

The association phase consisted of three FreeAsso blocks and two MemoryAsso blocks. Each FreeAsso block contained 50 trials. Each MemoryAsso block contained 30 sequences of, on average, 5 items. Each Participant ran 3 FreeAsso and 2 MemoryAsso blocks.

### ***Test Phase***

Participants fixated a square at the center of the screen, just as in the association phase. They were asked to produce, at random, right and left key presses. Again, the key presses triggered presentation of letter stimuli 100ms after the key press for a duration of 200ms. In this phase, however, Hs and As were presented randomly after each key press, such that 50% of the generated stimuli were congruent with the previous association (i.e., the letter corresponded to the one associated to that key press in the association phase), and 50% were incongruent.

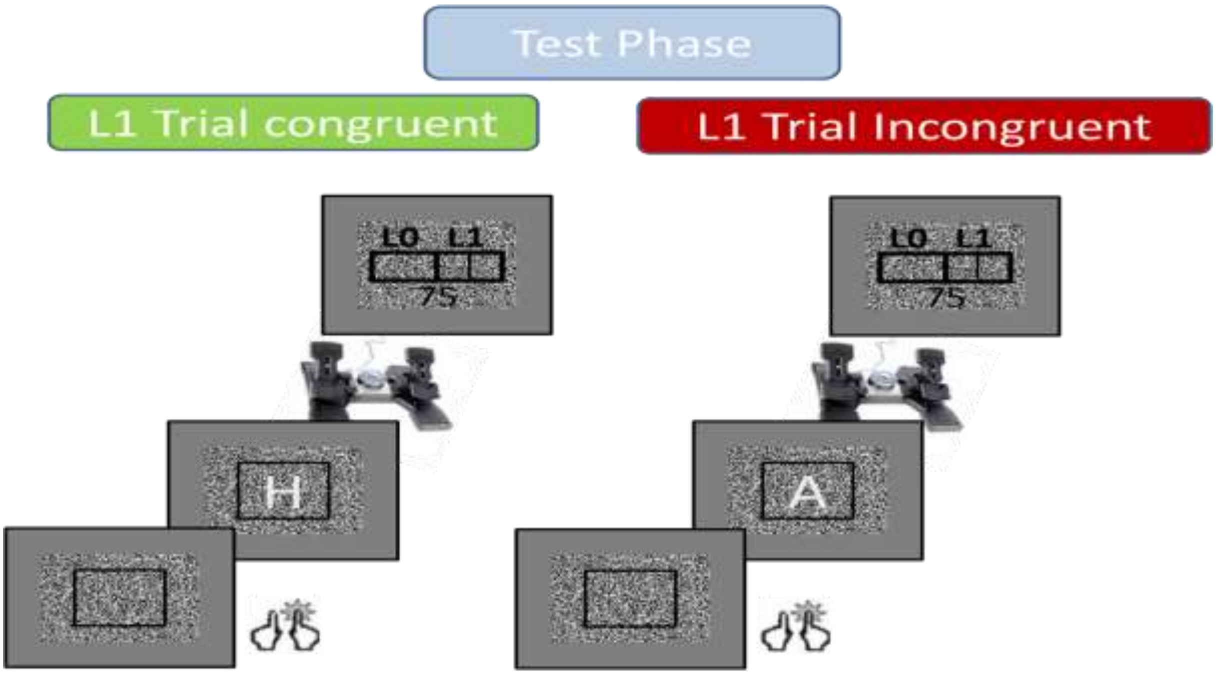

**Figure 8:** Test phase protocol representation

The stimuli appeared randomly (but in equal proportions) with the luminance C0 or the luminance C1. Participants were told that there were two categories of luminance ranging from the value 0 to 50 for the C0 category and from 50 to 100 for the C1 category. In order to maintain this uncertainty about the contrast on 5% of the trial stimuli appeared with a random contrast between C0-15% contrast and C1+15% contrast. After the stimulus had disappeared participants were required to judge the luminance value of the stimulus on a luminance response bar. On this bar participants could place the cursor on the perceived contrast value (from 0 to 100 except 50) with values under 50 corresponding to C0 and over 50 corresponding to C1. Participants completed 3 tests blocks of 44 trials before being in a re-association phase composed by one of each association blocks and ran 3 others tests blocks. In total participants responded to 264 test trials.

### ***Participants.***

Fifteen participants took part in the experiment. They were naive to the purpose of the experiment. Three of these fifteen participants were excluded from the analysis as their luminance discrimination d's were almost 0 (mean d'=0.038 SEM=0.124). Six of the remaining participants (7 women, 5 men; mean age = 24 years, *SEM* = 3.69 years) had action-effect mapping 1 (left→A, right → B), and six had mapping 2.

### ***Analysis of discrimination performance***

The luminance discrimination task was considered to be a yes/no protocol, with C1 being the target. That is, a C1 response to a C1 stimulus is a hit, a C1 response to a C0 stimulus is a false alarm, etc. d' and c are calculated using $d' = z(\text{hit rate}) - z(\text{false alarm rate})$ and $c = -0.5 \times [z(\text{hit rate}) + z(\text{false alarm rate})]$[3]. D' and c were calculated separately for congruent and incongruent trials.

## b Results

We first analyzed our data dependent on participants' contrast discrimination, such that we divided our contrast rating into two classes of contrasts (corresponding to C0 and C1). Discrimination performance (d') was lower in the congruent condition than in the incongruent condition (Congruent: M=1.82, SEM=0.54 Incongruent: M=1.98, SEM=0.59). A repeated measure one factor ANOVA with the factor of congruency showed this effect of congruency on d' to be significant (F(1,11)=5.59, p=0.037). At the same time, the criterion was not different in the two congruency conditions (Congruent: M=0.32, SEM=0.23 Incongruent: M=0.34, SEM=0.33; F(1,11)=0.15, p=0.69).

We performed separate repeated measure ANOVAs for Side*Congruency (interaction: F(1,11)=0.99, p=0.33) and for Mapping group*Congruency (interaction: F(1,5)=0.064, p=0.81) what revealed the congruency effect to be identical for both response sides and for both mapping groups. Taken together, these finding supports the hypothesis that preactivation of predicted action-effects reduces discrimination sensitivity.

Since participants provided their judgments of contrast using a continuous scale from 1 to 100 this allowed us to analyze not only their overall contrast judgment (C0 or C1) but also their rating of the perceived intensity (contrast) of the stimulus. We used these ratings to compute ROC (receiver operating characteristic) curves for each participant. We calculated the Area under the curve ($A' = 1/2 * \Sigma(F_{i+1} - F_i)(H_{i+1} + H_i)$) separately for every participant and condition. A one tailed T-test revealed a near significant difference such that incongruent trials were perceived as being of higher contrast that congruent trials (Congruent: M=0.871, SEM=0.05 Incongruent: M=0.88, SEM=0.06; t(11)=0.8866, p=0.0515). As we can see on Figure 9, the distribution of A' tends to be in majority above the equality line (Cong = Incong). This fits the distribution pattern of d' (see figure 9).

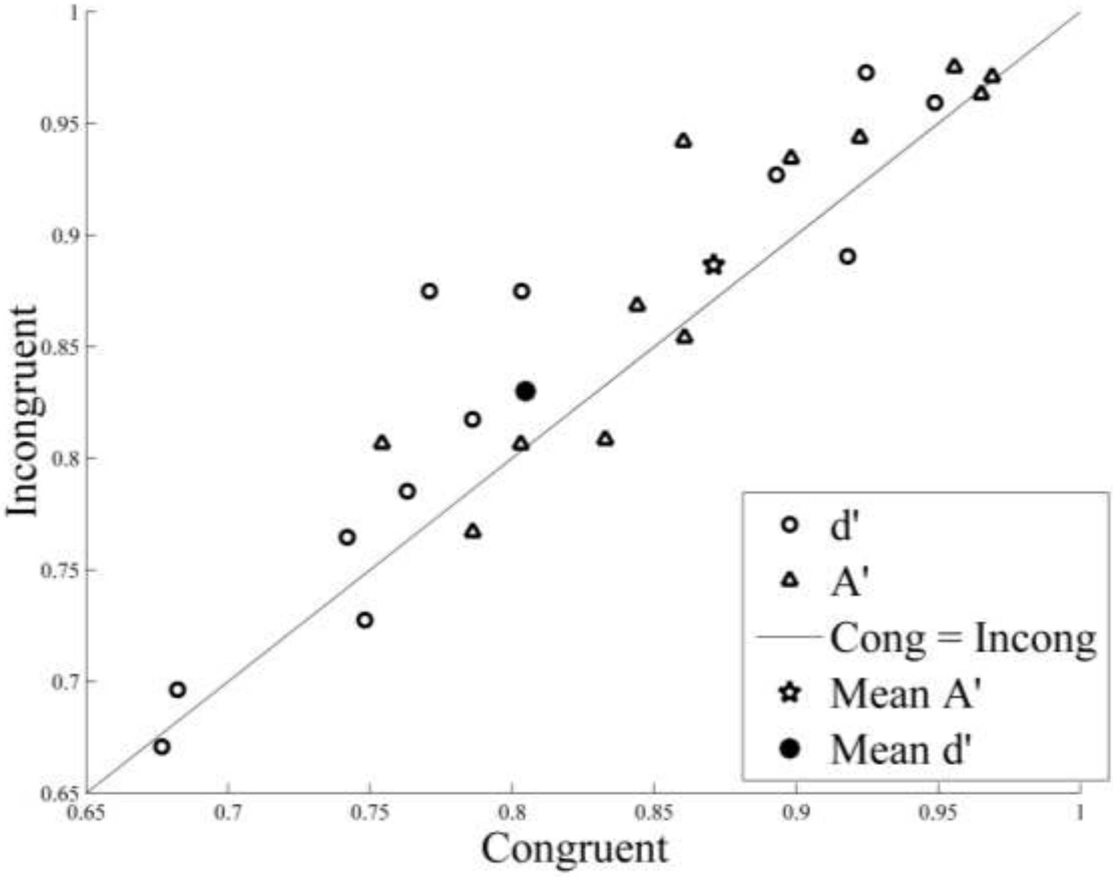

**Figure 9 :** Distribution of individual results for d' and A', above the diagonal Incongruent>Congruent.

### c Discussion

We trained participants to acquire associations between left- and right-hand key presses and the visual presentation of the letters A and H, respectively. In a subsequent test phase, participants performed right- and left-hand key presses that triggered the presentation of either and H or an A. The letters appeared in one of two contrasts (C0/C1). The stimulus could either be congruent or incongruent with respect to the learnt action-effect contingency. We showed that luminance discrimination between the two contrasts yielded a smaller d' for congruent action-letter combinations (i.e., when the letter corresponded to the one associated to that key press in the association phase) than for incongruent combinations. Subsequent analysis of participants' luminance ratings showed that this reduction in d' was due to internal responses of C0 being increased for congruent compared to incongruent trials, while internal responses of C1 not being different for congruent and incongruent trials. This brings about that the internal responses of C0 and C1 are drawn together for congruent compared to incongruent trials, making the luminance discrimination more difficult in the former than in the latter type of trial. Note that this pattern of results corresponds to what the preactivation account outlined above predicts. The account predicts a reduced contrast discrimination performance for congruent compared to incongruent stimuli, because the stimulus-driven internal response gain is lower with than without preactivation. The stronger the stimulus the smaller the increase of the internal response due to the preactivation (i.e., with preactivation (congruent) compared to without preactivation (incongruent)).

The presence of a shift of the internal response for C0 but not for C1 in our data confirms the model's prediction.

So, detection/discrimination is determined by the difference between Signal and Noise with and without preactivation ((Signal without – noise without) > (Signal with – noise with) in the case of detection, and by the difference between C1 and C0 with and without preactivation in case of discrimination (C1 without – C0 without) > (C1 with – C0 with). In a nutshell, detection/discrimination is worse because preactivation shifts the perceived contrast/rating more to the right for noise than signal (detection), and for C0 than C1 (discrimination).

In order to estimate the level of preactivation that would result in a difference in discrimination d' between a situation without and a situation with preactivation as observed in our experiment, we attempted to use our model to recreate our observed results. We first fitted the exact discrimination d' we observed in the incongruent condition with a Weibull function, using the contrast values from our experiment C0 (36%) and C1 (51%). To be able to calculate d' (Distance of the distributions/Variance of the distributions) we needed to estimate the variance of the internal response of a given contrast. Since the variance of the participants' perceived contrast ratings did not differ between our conditions any variance satisfying the constraint of returning the exact experimental d' could be chosen. Moreover, we constrained the function to minimize the differences in internal response to C1 across all possible preactivation levels. We found that, with a Weibull repartition function having the parameters λ=14.8 and k=1 we were able to fit a 3% variance of the internal response to yield exactly to the observed d' for incongruent trials (1.98). The internal responses for C0 and C1 were 90.7% and 96.6%, respectively. This fits well the fact that our stimuli were highly supraliminal (see methods and materials).

Using the function for the incongruent condition as a starting point, we estimated the level of preactivation necessary to yield a d' reduction as observed in the experiment. We found that a preactivation activity of 8% was sufficient to reduce d' from 1.98 in the incongruent condition to 1.82 in the congruent condition. The internal responses to C0 and C1 in the latter condition were 91.4% and 96.9%. Hence, the internal response for C0 increased much more (0.7%) than the internal response of C1 (0.3%), replicating what we observed in our experimental data.

To further validate our model, we assessed whether the same amount of preactivation would explain the decrease in detection d' observed by Cardoso-Leite et al. (2010). In a detection task sensitivity is represented by the distance between the mean baseline internal activity ('noise') and the mean internal activity driven by the stimulus ('signal' + 'noise'; see Figure 1 and Figure 2a). Cardoso et al. observed a reduction in detection d' from the incongruent condition (2.55) to the congruent condition (2.37). We modeled detection d' as the difference in internal response between the baseline activity corresponding to 1% of internal activity due to noise and the internal response to a stimulus at threshold (9% of internal activity yielding a detection d' of 2.55). We found that adding 8% of preactivation activity (as estimated above) made detection d' drop to 2.34. Hence, the same preactivation level of 8% could also explain the results from Cardoso-Leite et al. (2010).

In this study we introduce a new preactivation-based model to account for the role of motor prediction in sensory attenuation. This model allowed us to derive novel hypotheses regarding participants contrast discrimination performance. We should note such preactivation induced sensory attention might not be limited to action prediction, but may also result from other non-motor prediction mechanisms (see Waszak et al., 2012 for a discussion).

# B. Study 2: Investigate the link between Neurophysiology and psychophysics

*This chapter is based on:* ***Roussel****, C., Hughes, G., & Waszak, F. Action prediction modulates both neurophysiological and psychophysical indices of sensory attenuation. [Under Review]*

It has been shown that self-generated stimuli are perceived as less intense than externally generated stimuli, a phenomenon known as sensory attenuation as outlined above. Sensory attenuation has been demonstrated in the somatosensory (Blakemore et al., 1998), the auditory (Sato, 2008) and the visual domain (Cardoso-Leite et al., 2010; Roussel et al., 2013). Studies investigating sensory attenuation as a perceptual phenomenon have been complemented by studies investigating neurophysiological correlates of anticipated action effects (e.g., Aliu et al., 2009; Baess et al., 2008; Blakemore et al., 1998; Gentsch and Schütz-Bosbach, 2011; Hughes and Waszak, 2011; Hughes et al., 2013b; Schäfer and Marcus, 1973). To give an example, Baess et al. (2008) found a reduced auditory N1 component for action-triggered tones compared to externally triggered tones, suggesting that cortical activity was attenuated for the former.

Both attenuated phenomenological and neurophysiological responses are usually interpreted along the same lines in terms of forward models introduced above (e.g., Wolpert and Miall, 1996), as if they reflect the same mechanism. At the same time, this mechanism has usually been considered to be relatively low-level (‘sensory’). However, a systematic investigation of the relationship between sensory attenuation as a perceptual phenomenon, on the one side, and as a neurophysiological phenomenon, on the other side, is missing. As a consequence, it is impossible to tell how the attenuation of perceptual awareness is related to the attenuation of cortical responses that have been observed a number of times in separate experiments. Moreover, concerning the locus of the effect, differences in cortical responses between conditions with and without effect anticipation cannot always be unequivocally attributed to sensory processing. Often they may also be caused by other differences in attentional and cognitive processing (cf., Waszak et al., 2012; Hughes et al., 2013a). As a consequence, different studies do not always converge to the same conclusions. For example, Baess et al. (2008) observed attenuated fronto-central negativity when comparing action-triggered vs. externally triggered auditory stimuli. They concluded that early sensory processing in the auditory cortex is reduced. By contrast, Hughes and Waszak (2011) compared ERPs to action-triggered vs. externally triggered visual stimuli. They observed an increased, not a decreased, visual P1 component. In this study, attenuated cortical responses were observed in a frontoparietal network, starting 150 ms after stimulus. This result would be in line with the findings of Del Cul et al. (2007) showing that subjective thresholds of visual stimuli is reflected in later processing in a fronto-parietal network, rather than in early visual areas.

The aim of the current experiment was to shed new light on two interrelated questions. First, we investigated whether neurophysiological indices of sensory attenuation reflect early, low-level or later, higher-level mechanisms. Second, we explored how neurophysiological and perceptual indices of sensory attenuation relate. To do so, we adapted, using EEG, a luminance discrimination protocol that has been used before successfully to assess perceptual sensitivity and response bias of anticipated and unanticipated visual action effects (Roussel et al., 2013; see chapter II.A). Roussel et al. made participants learn an association between left and right key presses and the presentation of the letters A and H, respectively. They then made participants perform left and right key presses that randomly triggered presentation of either an H or and A at one of two possible contrasts. Participants were required to make discrimination judgments between the two contrasts. They showed contrast discrimination to be worse when the prediction (H or A, as learned during the association phase of the experiment) matches the true stimulus. Importantly, this paradigm does not only manipulate whether an action effect is predicted or not, but also the action effects' physical energy (contrast, as we used visual stimuli). It, thus, allows us to test whether or not prediction influences the same early components in the EEG as physical stimulus energy. If this is the case, then the effect of prediction (sensory attenuation) is likely an early, low-level phenomenon. Moreover, assessing both psychophysical and neurophysiological measures of sensory attenuation enables us to tell how neurophysiological components and reduced awareness of the action effects interrelate.

# Experimental part: Action prediction modulates both neurophysiological and psychophysical indices of sensory attenuation

---

## a Materials and Methods

### ***Stimuli.***

The stimuli were two white letters (A and H) presented within a virtual square of 3.3 degrees of visual angle and displayed on a 24 inch LED monitor at a 60hz refresh. These two stimuli were presented at two different contrast values (C0 and C1, determined for each subject; see Contrast determination phase) at the center of the screen. In the test phase we used a uniform noise texture to increase perceptual variance. This noise texture was re-sampled on each screen refresh with always the same number of white and black pixels. The mean luminance of the noise was then equal to the gray background.

### ***Contrast determination phase.***

In order to determine individual contrast values C0 and C1 yielding a discrimination d' of about 1.5, every participant completed a psychophysical staircase converging on 90% correct responses in a letter identification task (A vs. H). This correct response rate was used to ensure that the stimuli were supraliminal and that we could then independently manipulate discrimination. We used the resulting contrast value as the referential contrast C0 in a 2AFC paradigm (with constant stimuli ranging from C0 to C0+12%) in order to calculate the contrast value of C1 yielding 80% correct responses in a luminance discrimination task (C0 vs. C1). For an ideal observer, this contrast yields a discrimination d' of around 1.5 (Mcmillan & Creelman, 1991).

### ***Association phase.***

Participants fixated on a 3.3 degrees visual angle square located at the center of the screen. They were asked to press with their right/left index finger one of two keys (right and left on a response pad), each key press triggering presentation of a visual effect (A or H). The key-letter mapping was counterbalanced across participants. The letters appeared 200ms after the key press at full contrast in the square at the center of the screen.

There were two types of association phase. First, in the free association ("FreeAsso") blocks the action sequence (left / right) was freely generate by the participants at a pace of about 1 key press every second for 50 seconds. In 5% of the trials the visual effect was a W. In these catch trials, the participant had to press both buttons within 1s of the appearance of the stimulus. Catch trials were meant to ensure that participants paid attention to the effect stimuli. Second, in memory association blocks ("MemoryAsso"), random lists of As and Hs were presented to the participants (the average list size was 5). The lists were presented via headphones as spoken letters. After the lists were presented, participants had to reproduce the sequence by pressing the corresponding button sequence.

The association phase consisted of three FreeAsso blocks and two MemoryAsso blocks. Each FreeAsso block contained 50 trials. Each MemoryAsso block contained 30 sequences of, on average, 5 items. Each Participant ran 3 FreeAsso and 2 MemoryAsso blocks.

### ***Test Phase.***

Participants fixated a square at the center of the screen, just as in the association phase. They were asked to produce, at random, right and left key presses. Again, the key presses triggered presentation of letter stimuli 200ms after the key press. In this phase, however, Hs and As were presented randomly after each key press, such that 47.5% of the generated stimuli were congruent with the previous association (i.e., the letter corresponded to the one associated to that key press in the association phase), and 47.5% were incongruent. On the remaining 5% of trials, no stimulus was presented. The stimuli appeared randomly (but in equal proportions) with the luminance C0 or the luminance C1. Participants were told that there were two categories of luminance ranging from the value 0 to 49 for the C0 category and from 51 to 100 for the C1 category. In order to maintain this uncertainty about the contrast on 5% of trials stimuli appeared with a random contrast between C0-15% contrast and C1+15% contrast. After the stimulus had disappeared participants were required to judge the luminance value of the stimulus on a luminance response bar. On this bar participants could place the cursor on the perceived contrast value with values of 49 and under corresponding to C0 and 51 and over corresponding to C1. Participants completed 3 tests blocks of 44 trials before being in a re-association phase composed by one of each association blocks and ran 3 others tests blocks. In total participants responded to 264 test trials.

### ***Analysis of discrimination performance***

The luminance discrimination task was considered to be a yes/no protocol, with C1 being the target. That is, a C1 response to a C1 stimulus is a hit, a C1 response to a C0 stimulus is a false alarm, etc. d’ and c are calculated using d′ = z(hit rate) − z(false alarm rate) and c = −0.5 × [z(hit rate)+ z(false alarm rate)])[1]. D’ and c were calculated separately for congruent and incongruent trials.

### ***EEG recording and data preprocessing.***

EEG was recorded with 64 electrodes (actiCAP, Brain ProductsGmbH, Germany). The EEG was digitized at 500 Hz. EEG analysis was conducted using EEGLAB (Delorme & Makeig, 2004) and custom-built Matlab scripts. The data were resampled offline to a 250 Hz sample rate, with a notch filter from 45 to 55 Hz to remove line noise. Epochs were generated from -500 to 980 ms relative to stimulus onset, with a 200 ms prestimulus baseline correction. Initial artifact rejection was conducted in a semiautomatic manner (in EEGLAB) by rejecting epochs with activity above 100 μV or below -100 μV, as well as rejecting trials where activity at any time point for any electrode was more than 5 standard deviations from the mean activity for that epoch. Any channels that contributed to the rejection of many epochs were considered for removal and later interpolation. Frontal channels that showed large amplitude blink activity were also excluded from this first pass of semiautomatic artifact rejection. Ocular artifact correction was conducted in EEGLAB in Matlab using independent component analysis (Delorme & Makeig, 2004). Following removal of eye blinks and eye movements, noisy channels were replaced by an interpolated weighted average from surrounding electrodes. Data were then re-referenced to the common average. A final round of semiautomatic artifact rejection with a threshold of +/- 80 μV was used to remove any remaining artifacts. All ERPs are presented with a low-pass filter of 20 Hz for visual presentation purposes.

EEG analysis was on averaged ERPs for each participant using ANOVA with the factors congruency (congruent, incongruent) and stimulus contrast (C0, C1). Since our task involved visual stimuli, we focused our analysis of a region of interest on the occipital electrodes (O1 Oz O2). Since our stimuli were degraded and presented in a continuous stream of background visual noise, we postulated that this might influence the latency of the visual response. Therefore we inspected the ERPs over our region of interest to determine the time window corresponding to an apparent peak for the visual stimulus that would also be modulated by the contrast of the stimulus (C0 vs C1).

### ***Participants.***

Nineteen participants took part in the experiment. They were naive to the purpose of the experiment. Four of these nineteen participants were excluded from the analysis as their luminance discrimination d’s were below 0.5 (for 2 of them) or because the ratio of right left key presses during the test phases exceeded a 75% 35% ratio (for one of them). One was rejected because of the poor quality of the EEG recordings. Seven of the remaining participants had action-effect mapping 1 (left→A, right → H), and height had mapping 2 (8 women, 7 men; mean age = 24 years, SEM = 3.69 years).

## b Results

In order to ensure that the data were equivalent between the EEG and the behavioral analysis only trials free from EEG artifacts and trials that were not classified as outliers in the behavioral data were analyzed. The amount of rejected data was less than 10% of the total number of trials.

### *Psychophysical Results*

We analyzed our data dependent on participants' contrast discrimination. Discrimination performance (d') was lower in the congruent condition (d' congruent: 1.22) than in the incongruent (d' incongruent: 1.42) condition (see Table 1). A repeated measure analysis of variance including the factor of congruency showed this effect of congruency on d' to be significant (F(1,14)=5.36, p=0.03). At the same time, the criterion was not different in the two congruency conditions (see Table 1; F(1,14) = 0.165, p = 0.69). An unbiased measure of A' also confirms our finding. Sensitivity appears to be better for incongruent trials (A': 0.81) than for congruent trials (A': 0.79) (F(1,14) = 4.670, p =0.04) (see figure 11).

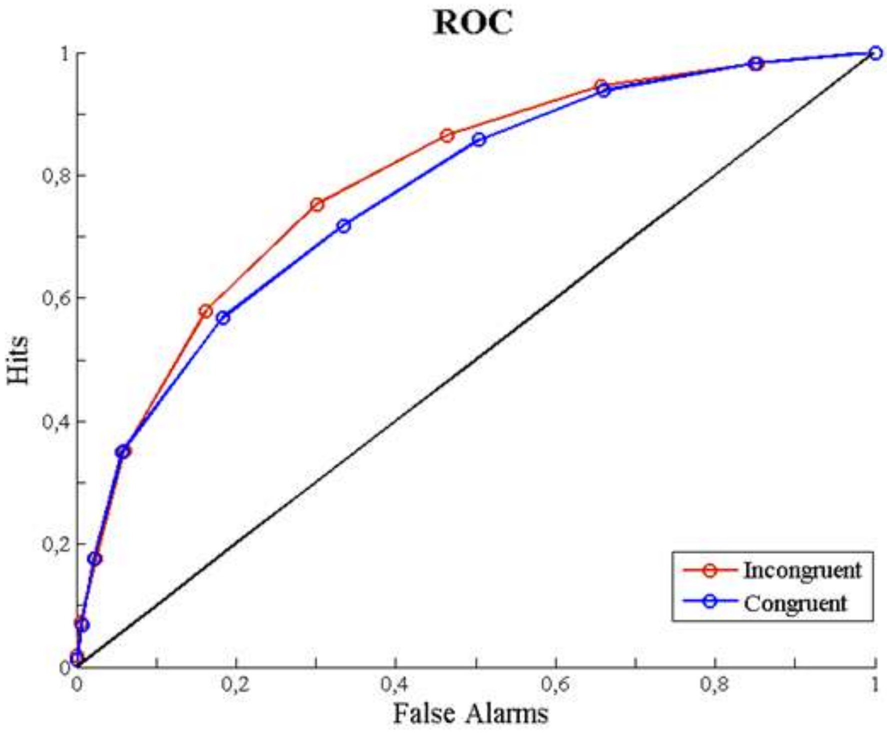

**Figure 10**: ROC curves for congruent and incongruent trials.

### *EEG Results*

In this section we focus on the effect of motor prediction on the neurophysiological indices of visual processing to determine the degree to which behavioral and neural sensory attenuation are related. The ERPs and the topographies for the different conditions are presented in Figures 2 and 3. A large negative deflection is apparent in all the waveforms, peaking at around 250 ms after stimulus onset. This peak appears to be greater for C1 than C0, such that it likely reflects processing of the visual stimulus (a delayed visual N1 component). To quantify these effects we took the average amplitude of each condition in a 140ms time window centered on this peak (180 ms to 320 ms). This analysis revealed a significant main effect of Contrast (F(1,14) = 6.56 p = 0.011), confirming significantly greater amplitude for C1 (mean = -1.67; std = 1.55) compared to C0 (mean = -0.81; std = 1.8). The topography of this difference is consistent with modulation of an occipital ERP component.

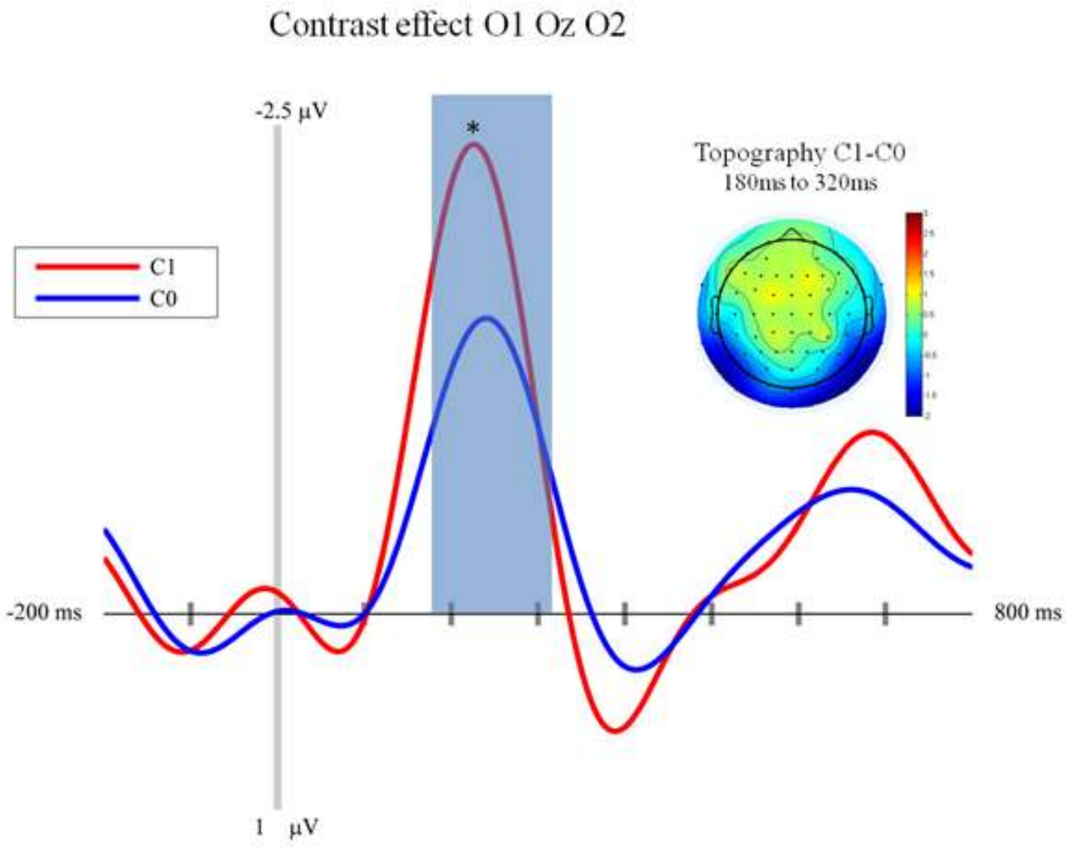

**Figure 11:** Contrast effect: ERP & Topography, This figure presents the mean ERPs on O1 Oz O2 for C0 (in blue) and C1 (in red) from -500 to 980ms relative to stimulus apparition. The blued surface represents the analysis time window (from 180 to 320ms, centered on the pic around 250ms). In the top left corner the topography of the difference (C1-C0) is presented for the analysis time window.

In line with the aim of the current experiment, it was important to determine whether we observed significant attenuation of this visual ERP peak as a function of congruent motor prediction. We observed a significant main effect of Congruency (F(1,14)=6.93 p= 0.009), such that our visual component was of significantly smaller amplitude in the congruent condition (mean = -0.99; std = 1.57), compared to the incongruent condition (mean = -1.50; std = 1.64). The topography of this difference is also consistent with a modulation of visual processing as a function of action prediction. Since we also observed attenuated sensitivity for congruent trials in the behavioral analysis presented above, this provides evidence that neural and behavioral measures of sensory attenuation are likely related.

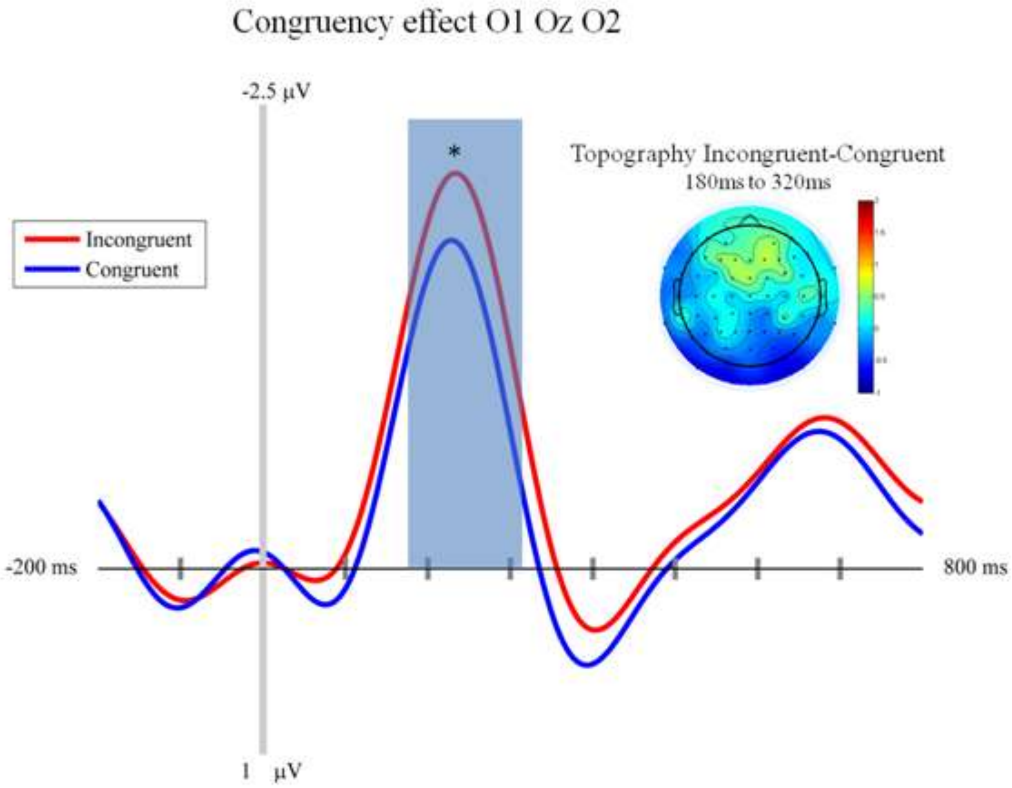

**Figure 12:** Congruency effect: ERP & Topography This figure presents the mean ERPs on O1 Oz O2 for congruent (in blue) and incongruent (in red) from -500 to 980ms relative to stimulus apparition. The blued surface represents the analysis time window (from 180 to 320ms, centered on the pic around 250ms). In the top left corner the topography of the difference (Incongruent – Congruent) is presented for the analysis time window.

### c Discussion

First of all, our experiment shows that neurophysiological sensory attenuation is a phenomenon that is not restricted to the auditory and somatosensory modality, but that it can also be observed in the visual domain. More importantly, one of the aims of the experiment presented above was to investigate the locus of sensory attenuation. This was done by way of comparing the influence of prediction and stimulus contrast, respectively, on the ERPs triggered by visual action effects. We observed that an N1 component was clearly modulated by stimulus contrast, with larger contrasts resulting in a larger deflection. Importantly, congruency affected the very same component, with congruent trials resulting in smaller amplitudes than incongruent trials.

Of course, our data cannot show which processing stage precisely the two factors manipulated in the current experiment influence. The current paradigm differs in important aspects from other studies investigating visual evoked potentials. Notably, we presented stimuli in visual noise resampled at each screen refresh. The actions' effects were, therefore, not presented with a sharp onset. Moreover, in our experiment, stimuli were triggered by an action. It is therefore difficult to compare our results to ERPs found in previous studies. However, previous research seems to suggest that contrast-dependent processes take place rather early in the visual processing stream (e.g., Schadow et al., 2007), while later components are rather modulated by motion and form perception (Bach & Ullrich, 1997; Göpfert et al., 1998). Importantly, our experiment allowed us to directly compare the effect of contrast and prediction. As it demonstrates that motor prediction influences the same processing stage as visual contrast, we assume that motor prediction as manipulated in our experiment influences an early processing stage that is otherwise still modulated by basic stimulus-features. This interpretation is corroborated by the fact that the N1 component in question has an occipital topography. That this component has a relatively late latency is probably due to the fact that the stimuli used in the current experiment were not presented with a sharp onset, but embedded in dynamic pixel noise, such that the detection of a pattern is more time-consuming.

The second aim of the present study was to explore how neurophysiological and perceptual indices of sensory attenuation relate. We used a luminance discrimination protocol to assess perceptual sensitivity and response bias of anticipated and unanticipated visual action effects, assessing EEG activity at the same time. The psychophysical results show that discrimination performance (d', A') was better in the incongruent condition than in the congruent condition. At the same time, the criterion was not different in the two congruency conditions. The results, thus, are in line with the findings of Roussel et al. (2013) and Cardoso-Leite et al. (2010). They show that contrast sensitivity is reduced when a motor act provides an accurate prediction of the ensuing visual stimulus.

As concerns the effect of congruency on ERPs, we observed that the contrast-sensitive visual component was significantly smaller in the congruent condition compared to the incongruent condition. We, thus, observed, to our knowledge for the first time, sensory attenuation in psychophysical and neurophysiological indices at the same time, suggesting that the two measures of sensory attenuation are likely related. However, note that the psychophysical effect corresponds to an interaction between contrast and congruency: The discrimination between the two contrast levels is more difficult in congruent than in incongruent trials. If sensory attenuation assessed with psychophysical methods were a direct reflection of the ERPs assessed at the same time, we would have expected to see an interaction between these two factors in our ERP data as well. However, this was not the case. Of course, it might be that the ERP data simply lack sufficient statistical power. However, it is also possible that psychophysical and ERP indices of sensory attenuation (at least those assessed in the current experiment) are not in a simple one-to-one relationship (See Appendix 2 for an ERP figure of the 2x2 design). Perception might be dependent not only on early cortical responses, but also on later processing and/or recurrent processing, tweaking the relationship between perceptual measures and observable neurophysiological measures.

In conclusion, in our experiment ERP effects of visual sensory attenuation were found to correspond to contrast-dependent processing stages. We conclude that motor prediction, thus, influences quite early processes in the sensory treatment. Moreover, we demonstrated that both psychophysical and ERP indices of sensory attenuation can be observed in the visual modality. However, the exact relationship between the two types of measure needs to be further clarified, as there are not only commonalities, but also differences.

## C. Study 3: Two dimensional extension of the preactivation principle

It has been shown in the present work (Roussel et al. 2013; see chapter II.A) that contrast discrimination is altered when a stimulus was predicted by a keypress.  The model outlined above suggests that we should call this effect sensitivity attenuation, rather than sensory attenuation, since this term better expresses the idea of a decrease in the ability to discriminate intensities. The model is based on the assumption of a non linear increase of the internal response to contrast due to preactivation. Thus, according to the preactivation theory the prediction of the action’s consequence increases the internal response distribution as a function of stimulus intensity, the more intense the stimulus, the less the increase of activity due to pre-activation.

However, note that sensitivity attenuation is due to the relative change of the signal and the noise distribution (or of two different contrast distributions); the absolute activity level of the signal(s) actually increases. Importantly, if the absolute internal response to a stimulus increases with preactivation, then preactivation should result in perceptual facilitation (not attenuation) when the signal has to be identified (instead of detected in noise/ discriminated from another contrast stimulus). This perceptual facilitation would be based on the fact that a preactivated neural population is more easily differentiated from others neural populations. To capture this notion, we extended the preactivation model to comprise a two dimensional decision space. This extension is outlined in the following.

According to signal detection theory, the distribution of the internal response represents the activity of a given neural population along one stimulus dimension. That is, this distribution represents the activity of one neural population coding for a particular stimulus (see Figure 14). To represent two different neural populations coding for two different stimuli, the formal representation needs to be extended to two dimensions (Macmillan & Creelman, 1991; see Figure 14). In this framework, an identification task is based on the comparison of the activity in two different neural populations (instead of one neural population as in the case of detection/ contrast discrimination). As a consequence, in the case of identification the preactivation model predicts a rather different result compared to detection/discrimination tasks (see Figure 14).

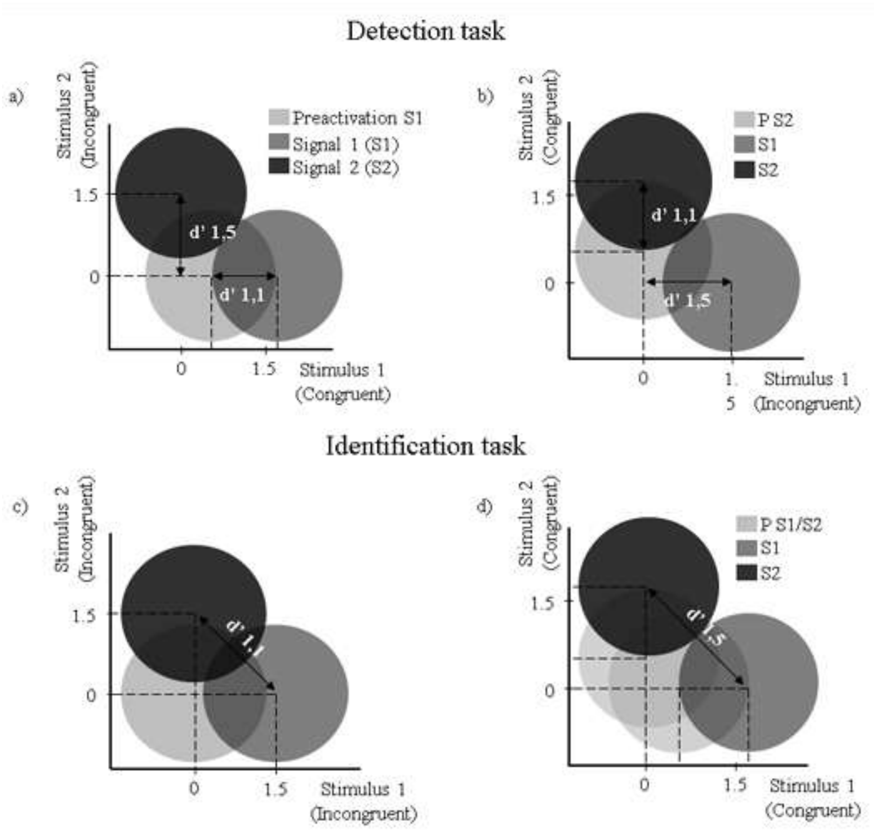

**Figure 13:** Schematic representation of the pre-activation model of action-effect anticipation. Panel (a) and (b) represent detection d' for stimuli that are predicted (congruent stimuli) compared to stimuli that are not predicted (incongruent stimuli). a-b) Lower d' for congruent vs. incongruent stimuli would result from the fact that discrimination of the signal activation from the pedestal level is more difficult than discrimination of the signal activation from the baseline level (i.e., in the incongruent condition without effect anticipation, in which baseline activity is not raised to pedestal levels). Importantly, note that according to Roussel et al., (2013) action-effect pre-activation would not only increase the baseline activity but also, although to a lower degree, the Signal distribution (i.e. Signals 1 in Figure 13a and Signal 2 in Figure 13b). c-d) Consequently, due to this increase in Signal 1 and Signal 2 distribution when these signals are preactivated (Figure 13d), the perceptual distance between S1 and S2 would increase, thus enhancing identification d' for congruent vs. incongruent stimuli (Figure 13d vs. Figure 13c).

Note that in a two dimensional space, too, the sensitivity is still given by the distance between the internal response distributions. However, detection d' and identification d' rely on different comparisons (see Figure 13). Since preactivation is specific to the predicted stimulus, it affects only one dimension. Hence, the increases in the internal response for the predicted stimulus results in the increase of the distance between the internal response distributions of the two stimuli that have to be identified (see Figure 13). Hence while the preactivation hypothesis predicts sensitivity loss in detection and contrast discrimination tasks (with congruent trials being worse than incongruent trials), it predicts an improvement of identification performance (with congruent trials being better than incongruent trials).

By assessing the identification performance after acquisition of action effect-relationships, we aimed to test this prediction of the preactivation hypothesis.

## Experimental part: Sensitivity facilitation in an identification task

### a Method

#### ***Stimuli.***

Experimental stimuli were generated and presented with Matlab 2007b using the psychophysics toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007). The stimuli were two tilted Gabor gratings (45° left, 45° right) displayed on a 15 inches CRT monitor at a 80hz refresh rate and fitted into a virtual square of 2.9 degrees of visual angle. The screen resolution was set at 800*600 pixels. These two stimuli were presented at a single contrast value (C0; see contrast determination phase) at the center of the screen consequently to a key press. We used in all our phases a uniform noise texture to increase perceptual variance. This noise was a 100*100 matrix filled with an equal number of white and black pixels. The matrix was re-sampled (using the randperm function in Matlab) on each screen refresh. The mean luminance of the noise was then equal to the gray background.

#### ***Contrast determination phase.***

In order to determine an individual contrast values C0 yielding an identification d’ of about 2 (between 2 orientations), every participant completed a psychophysical staircase converging on 85% correct responses in an orientation identification phase (rightward versus leftward tilt). We employed an adaptive staircase to manipulate stimulus transparency using an accelerated stochastic approximation algorithm as described by Kesten (1958). The initial step size was 20 and we stopped the staircase when the step size was 1 (in transparency). This correct response rate was used to ensure that the stimuli were supraliminal and that we could then independently manipulate identification with associated action effects. We used the resulting contrast value as the individual contrast value C0.

### ***Association phase.***

Participants fixated on a 3.3 degrees visual angle square located at the center of the screen. They were asked to press with their right/left index finger one of two keys (right and left on a response pad), each key press triggering presentation of a visual effect (leftward or rightward tilted Gabor grating). The key Gabor grating mapping was counterbalanced across participants. The gratings appeared 200ms after the key press at full contrast in the square at the center of the screen.

In the association phase the action sequence (left / right) was freely generated by the participants at a pace of about 1 key press every second for 100 trials. In 5% of the trials the visual effect was a yellow grating. In these catch trials, the participant had to press both buttons within 1s of the appearance of the stimulus. Catch trials were meant to ensure that participants paid attention to the effect stimuli.

The association phase consisted of ten association blocks.

### ***Test Phase.***

Participants fixated a square at the center of the screen, just as in the association phase. They were asked to produce, at random, right and left key presses. Again, the key presses triggered presentation of Gabor grating 200ms after the key press. In this phase, however, rightward and leftward tilted gratings were presented randomly after each key press, such that 50% of the generated stimuli were congruent with the previous association (i.e., the letter corresponded to the one associated to that key press in the association phase), and 50% were incongruent. After the stimulus had disappeared participants were required to judge the orientation of the stimulus. Participants completed 10 tests blocks of 20 trials interleaved with the 10 association blocks.

### ***Analysis of identification performance***

The Identification task was considered to be a yes/no protocol, with right being the target. That is, a right response to a right stimulus is a hit, a right response to a left stimulus is a false alarm, etc. d’ and c are calculated using $d' = z(\text{hit rate}) - z(\text{false alarm rate})$ and $c = -0.5 \times [z(\text{hit rate}) + z(\text{false alarm rate})]$)[1]. D’ and c were calculated separately for congruent and incongruent trials.

## b Results

The results were analyzed for the 10 participants who passed the experiment.

We first analyzed our data dependent on participants’ orientation identification. Identification performance (d’) was lower in the incongruent condition than in the congruent condition (Incongruent: M= 1.5, SEM=0.75; Congruent: M=2.12, SEM=0.6). A repeated measure one factor ANOVA with the factor of congruency showed this effect of congruency on d’ to be significant (F(1,9)= 11.8, p=0.008 ). At the same time, the criterion was not different in the two congruency conditions (Congruent: M=-0.13, SEM=0.28; Incongruent: M=-0.025, SEM=0.18 ; F(1,9)=1.8 , p=0.2 ).

However in this protocol it was possible that participants, in case of uncertainty, have indicated that the orientation of the grating that was congruent to the action they executed.

To test for this bias we computed *c* for trials in which participants executed the action that in the previous association phase triggered the rightward tilted grating (defined to be the “signal”), i.e., right key-press for half of the subjects and left key-press for the rest of the subjects. We did the same for the trials in which participants executed the action that was associated with the leftward tilted grating (defined to be “noise”). We reasoned that if participants were biased in responding that the dots moved in the direction that was congruent to the action they executed, they should have been *liberal* (negative *c*) when they executed the action associated with the rightward tilted grating and *conservative* (positive *c*) when they executed the action associated with the leftward tilted grating. Our analysis showed that for the left tilted grating the criterion did not different from 0 ( mean : -0.01; F(1,9)=0.036 ,p=0.85 ; no bias), but the right tilted grating criterion appeared to be significantly different from 0 (mean=-0.12; F(1,9)=5.401 ,p=0.0452).

### c Discussion

In the current experiment we aimed to investigate identification performances in case of motor induced outcome prediction. Participants first associated right and left key presses with rightward and leftward tilted Gabor gratings. In the test phase participants had to freely produce a keypress followed randomly by one of the two gratings presented at identification threshold. They then had to indicate the stimulus orientation. The stimulus could either be congruent or incongruent with the learnt action-effect association.. We observed that identification d’ was higher for congruent trials than for incongruent trials. However, the method we used was not bias free, as we varied the stimulus dimension along the response dimension. Indeed, further analysis assessing the response bias showed that for one stimulus orientation the criterion was different from 0.

This effect on the criterion thus raises the question as to the origin of our effect, as it renders the interpretation of the d’ effect difficult. However we would like to argue that if the results were mostly driven by a response strategy then both *criteria* should be different from 0. Hence, the asymmetry in our results suggests that our effect was at least partly driven by sensory processes and not uniquely by response bias.

## D. Study 4: A closer look on the dynamics of preactivation

*This chapter is based on: Desantis, A., **Roussel**, C. & Waszak, F. The temporal dynamics of the perceptual consequences of action-effect prediction. [Under Review]*

To recapitulate: The ideomotor theory claims that performing an action results in a bidirectional association between the action’s motor code and the sensory effects the action produces. Once acquired, these associations can be used to select an action by anticipating or internally activating their perceptual consequences (e.g., Elsner & Hommel, 2001; Herwig, Prinz, & Waszak, 2007; Prinz, 1997). This notion has been recently formulated in terms of a neural pre-activation, according to which action preparation/execution results in the activation of the sensory network that represent the sensory action effect (see Kuhn et al., 2010; Waszak, Cardoso-Leite, & Hughes, 2012; Roussel, Hughes & Waszak, 2013; SanMiguel et al., 2013).

The ideomotor principle has been corroborated by a number of studies (see Hommel, Müsseler, Aschersleben, & Prinz, 2001; Schütz-Bosbach & Prinz, 2007; Shin et al., 2010; Waszak et al., 2012; see Chapter I.F). However, it remains unclear when action-effect prediction is generated by the brain. The question of whether it is related to preparatory stages of motor processing or rather to the execution of the action is essential, as it differentiates between a strong and a weak version of the ideomotor theory (cf. Ziessler & Nattkemper, 2011). A strong version assumes that effect anticipation is an integral part of action selection (e.g., James, 1950; Prinz, 1997). Effect anticipation should therefore necessarily take place at early stages of motor preparation. However, if effect anticipation is rather used for quality control and error handling, it could occur at later stages of motor preparation or after action execution.

The present study investigates the temporal dynamics of action-effect anticipation (see also Bays, Wolpert, & Flanagan, 2005; Ziessler & Nattkemper, 2011) by assessing the perception of predicted and unpredicted action effects. Participants completed an acquisition phase during which specific actions (left and right key-presses) were associated with specific visual effects (dots moving upward or downward). In the test phase they completed a 2 AFC identification task in which they were required to indicate whether the dots moved upward or downward. To isolate any effects of action-effect prediction, participants were presented with congruent and incongruent dot motion in which the association they learned in the previous acquisition phase was respected or violated, respectively. Crucially, to assess the temporal dynamics of action prediction, congruent and incongruent stimuli were presented at different time points not only after but also before action execution.

We observed higher sensitivity (d') to motion discrimination in congruent vs. incongruent trials only when stimuli were presented from about 220ms before the action to 280ms after the action. The temporal dynamics of our effect suggests that the perceptual modulation of action-effect prediction occurs during motor preparation. As we will discuss below, our results corroborate the predictions of the pre-activation model concerning the influence of action-effect anticipation on identification d'.

## a Materials and Methods

### ***Materials.***

Stimulus presentation and data acquisition were conducted using the psychophysics Toolbox (Brainard, 1997; Pelli, 1997) for Matlab 7.5.0 running on a PC computer connected to a 19-in. 85 Hz CRT monitor. Auditory stimuli were presented via a pair of headphones.

### ***Participants.***

Sixteen volunteers (average age = 26.34 years, *SD* = 5.42 years) participated in the experiment for an allowance of € 10/h. All had normal or corrected-to-normal vision and hearing and were naïve as to the hypothesis under investigation. They all gave written informed consent.

Participants completed 40 acquisition and 40 test phases presented in an ABAB order.

### ***Acquisition phases.***

The aim of the acquisition phases was to build action-effect associations. Participants were presented with a Random Dot Kinematogram (RDK) in which 100 dots (dots size: 0.107deg) were displayed within a circular aperture of 7deg of diameter. They executed left or right key-presses in a random order and about equally often. Feedback of the proportion of right and left key-presses was provided every 20 trials. Participants' actions generated a coherent dot motion: 90% of the dots moved either upward or downward during 100ms. For half of the participants the left key-press triggered an upward dot motion and the right key-press generated a downward motion. For the other half of the participants the reversed mapping was used. In 10% of the trials, after participants' actions, the dots moved obliquely either downward or upward. In these trials participants were required to immediately press the space bar. Each acquisition phase consisted of 60 trials except for the first acquisition phase (100 trials), for a total of 2440 trials.

### ***Test phases.***

As for the acquisition phases, in the test phases participants were presented with a RDK. They executed random left and right key-presses and about equally often. Feedback of the proportion of right and left key-presses was provided every 5 trials. Contrary to the acquisition phase, they were asked to execute left/right actions at about one second after a go signal (a 700Hz sinusoidal tone, 100ms of duration, presented at 70dB). Only the trials in which participants performed an action within 900ms and 1300ms after the go signal were considered as “correct” trials and analyzed. The rest of the trials were replaced with new trials until a minimum of 40 trials per condition was reached. Coherent dot motion (upward/downward motion) was delivered at different Stimulus Onset Asynchrony (SOA), i.e., 400, 600, 800, 1000 or 1200ms after the go signal. Consequently, the dots moved coherently before or after action execution. The coherent dot motion lasted 100ms. Thereafter, the dots moved randomly. In addition, the percentage of dots moving coherently was individually determined for each SOA and each participant in a preliminary experiment to produce a discrimination threshold of 75% of correct responses.

Dot motion direction was either congruent or incongruent with respect to the action-effect association participants learnt in the previous acquisition phase. On congruent trials, participants’ key-press was preceded/followed by the dot motion direction that was associated with that same key-press in the acquisition phase. On incongruent trials, their action was preceded/followed by the dot motion direction that was associated with the other hand in the acquisition phase. At the end of each trial participants indicated whether the dots were moving upward or downward (Figure 14).

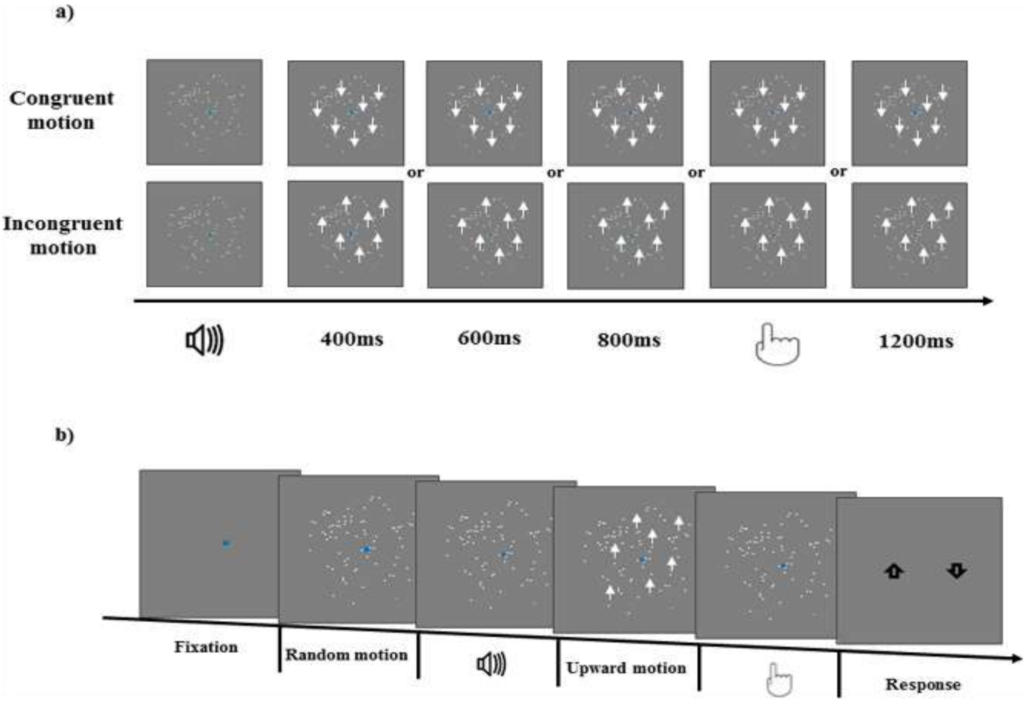

Figure 14 a) Illustration of the factorial design. Participants' executed a left/right key-press at about 1 second after a go signal (pure tone). They were presented with congruent or incongruent trials in which the action-effect association participants learnt in the acquisition phase was respected or violated, respectively. Congruent/ incongruent motion was presented at different interval before or after the action. b) Illustration of a trial from the test phase. At the end of each trial, participants were required to indicate whether the dots moved upward or downward. The percentage of dots moving upward/downward was individually determined for each SOA in a preliminary experiment to produce a discrimination threshold of 75% of correct responses.

Each test phase consisted of 10 trials for a total of 400 trials (40 x 5 SOAs x 2 Congruency conditions).

### b Results

To assess the temporal dynamics of action-effect anticipation, we firstly divided the stimulus-before-action trials into four time intervals: *bin1*, from -880ms to -661ms; *bin2*, from -660ms to -441ms; *bin3*, form -440ms to -221ms; *bin4*, from -220ms to -1ms. The stimulus-after-action trials formed one bin: *bin5*, from 0ms to 280ms. Then, we computed participants' sensitivity to motion discrimination (upward vs. downward) for the 2 Congruency conditions (congruent and incongruent) and the 5 Intervals.

We conducted a repeated measure of variance (ANOVA) on *d'* values with Congruency (congruent and incongruent) and Intervals (bins 1, 2, 3, 4 and 5) as factors. The analysis showed no main effect of Congruency: *F*(1, 16) = .2459, p = .6266. Similarly, no main effect of Intervals was observed: *F*(4, 64) = .8469, p = .5007. However, the interaction was significant: *F*(4, 64) = 2.9282, p = .0274. Further two-tailed paired *t*-tests showed higher *d'* in congruent vs. incongruent trials for the *bin4* (-220ms : -1ms) and the stimulus-after-action trials (0ms : 280ms), *t*(15) = 2.8093, p = .0132 and *t*(15) = 3.0098, p = .0087, respectively. We did not find any differences between congruent and incongruent trials for the other three before-action intervals: *t*(15) = -0.8361, p = .4161 (*bin1*, -880ms : -661ms), *t*(15) = -1.1432, p = .2708 (*bin2*, -660ms : -441ms), *t*(15) = -0.6252, p = .5411 (*bin3*, -440ms : -221ms;). In summary, participants showed better discrimination in the congruent vs. the incongruent trials when stimuli where displayed from 220ms before the action to 280ms after the action (for an illustration of the temporal dynamics of action-effect prediction see Figure 2).

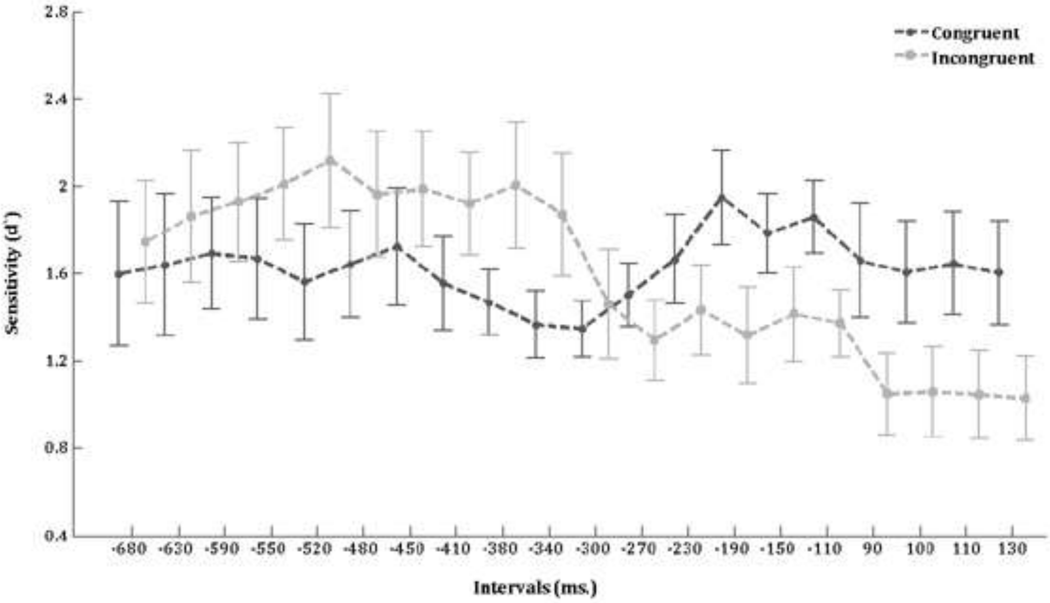

Figure 15: The ordinate axis shows participants' running mean sensitivity (d') to motion discrimination for congruent and incongruent trials. The abscissa shows the mean Interval at which the dots moved upward/downward with respect to participants' action execution. The running means were computed as follows: for each participant, we started with about 40 trials (20% of the total number of trials per condition) having the longest negative SOA (the leftmost interval in the figure). We then discarded the 5% of trials having the longest SOAs within this bin of 40 trials and added the 5% of trials having the next shorter SOA; and so on until the bin included the 40 trials longest positive SOAs (the rightmost interval in the figure). The graphic shows that around 230ms participants' sensitivity for congruent stimuli increase compared to sensitivity for incongruent stimuli. Bars represent standard error.

To check whether the increase in d′ was truly a sensitivity change we computed participants' response criterion c = −0.5 × [z(hit rate) + z(false alarm rate)]). In particular, in order to rule out the possibility that participants, in case of uncertainty, might have indicated that the dots moved in the way that was congruent to the action they executed we run the following analysis. We computed *c* for the trials in which participants executed the action that in the previous association phase triggered the upward motion (defined to be the "signal"), i.e., right key-press for half of the subjects and left key-press for the rest of the subjects. We did the same for the trials in which participants executed the action that was associated with downward motion (defined to be "noise"). We reasoned that if participants were biased in responding that the dots moved in the direction that was congruent to the action they executed, they should have been *liberal* (negative *c*) when they executed the action associated with the upward motion (action-signal trials) and *conservative* (positive *c*)when they executed the action associated with the downward motion (action-noise trials). Our analysis showed that *c* did not different from 0 (no bias) for the 5 intervals for both the action-signal and the action-noise trials, confirming that the increase in *d'* for congruent trials was truly due to a sensitivity change.

## c Discussion

The current experiment sought to elucidate the dynamics of action effect anticipation by tracing the time course of its perceptual consequences. Participants learned that specific actions were associated with specific visual effects (upward and downward dots motion direction). They then performed a 2 AFC identification task in which they were required to indicate whether the dots moved upward or downward. Participants were presented with congruent and incongruent dot motion in which the association they learned in the previous acquisition phase was respected or violated, respectively. Crucially, congruent and incongruent stimuli were presented at different time points before or after the execution of one of the actions. Our results showed higher sensitivity (d') to motion discrimination (i.e., upward vs. downward) for congruent than incongruent stimuli when these stimuli where presented later than about 220ms before action onset. This result demonstrates that the anticipation of the action's effect takes place during action preparation.

The current study is important in the context of the comparison of voluntary and stimulus-driven action control. Ziessler and Nattkemper (2011) tested for effect anticipation at different stages of the preparation/execution of stimulus-triggered actions (contrary to the present study, in which actions were voluntary). Their participants learned that the responses to certain target letters triggered contingently another letter on the screen. Thereafter, the target letters were presented along with the effects of the correct response, effects of other responses, or neutral letters. Ziessler and Nattkemper manipulated the stimulus onset asynchrony between target stimuli and the flanker. Contrary to the results reported above, they found that the flankers influenced response times only at 0 or positive SOAs (flanker presented after the target). They conclude that effect-related information from the flanker stimuli is involved only in late phases of response preparation/execution to enable the evaluation of the action. The fact that we observed evidence of effect anticipation up to 220ms before movement onset might suggest that stimulus-driven and voluntary actions differ in the role effect anticipation plays in action preparation and execution, with effect anticipation being involved earlier in the course of motor preparation for voluntary than for stimulus-driven action control. This notion would fit to the suggestion of Herwig, Prinz and Waszak (2007) that stimulus-driven and voluntary action control draw to a different degree on anticipatory mechanisms.

Note that the effect we assessed in the present experiment should be considered to be a lower bound of the start of effect anticipation. This is because we assess effect anticipation when it already affects perception. Our results, thus, likely reflect a processing stage that is already rather advanced, in that it involves a perceptual representation of the effect. Prior to this stage, the anticipated effect might be represented in a format that is not so closely related to perception and, therefore, did not influence the d'-s in our experiment.

We need to point out that our results are in line with the prediction made by the preactivation hypothesis (see previous chapter). In the present experiment participants were required to identify the stimuli. Our result is predicted by the pre-activation model of action-effect prediction. This model suggests that voluntary action results in the pre-activation of the sensory network that represent the action's expected perceptual consequence, thus increasing the mean level of activity in this network to some pedestal level (Cardoso-Leite et al., 2010; Roussel, Hughes, & Waszak, 2013; Waszak et al., 2012). Although if congruent stimuli are perceived less easily when they have to be detected in noise, they are identified more easily when they have to be discriminated from another stimulus (cf. Study 6; Roussel, Allenmark, & Waszak, in preparation). Indeed, note that identification differs from detection, in that the differentiation between 2 stimuli is not based on the activity in one selective neural population. Instead, in identification tasks the perceived difference between stimuli relies on the competition between different neural populations (see previous chapter).

In sum, the current experiment sought to elucidate the dynamics of action-effect anticipation by tracing the time course of its perceptual consequences. The temporal dynamics of our effect suggest that the perceptual modulation of action-effect prediction occurs at later stages of motor preparation (220ms the execution of an action - note that pre-movement activity later than 150ms before movement is most likely attributable to motor execution processes, see Gratton, Coles, Sirevaag, Eriksen, & Donchin, 1988), that is when specific motor command is going to be delivered to muscles. Thus if specific preactivation of the coding sensory network there is it seems to be driven by the latter stage of motor preparation. This appears to be in line with Pacherie's model of voluntary action (2008) were the representation of the effect outcome gets more and more specified in terms of sensory consequence along the specification of the motor command.

## E. Study 5: Preactivation due to sensory input

*This chapter is based on: **Roussel**, C., Hsu, Y-F., Waszak, F., Repetition priming results in sensitivity attenuation. [In Prep]*

In the previous chapters we outlined the preactivation model and showed that the model can explain certain perceptual consequences of action effect anticipation, such as sensory attenuation. The preactivation hypothesis is based on the notion that a sensory pathway can be activated before the actual presence of the stimulus. This idea, thus, implies that the preactivation itself can be considered as a sensory input (Waszak et al, 2012). As a matter of course, action-based preactivation does not need to be the only source of neural preactivation. For instance, it has been shown that mental imagery of a stimulus interferes with the true perception of the same stimulus, just as action-based prediction does (Craver-Lamley, 1987; Perky, 1910; see Waszak et al., 2012, for a review; a more detailed discussion of that topic is provided in the general discussion).

Another parallel can be drawn between action-based preactivation and repetition-based practivation of sensory pathways. Notably, it has been shown that when a stimulus is repeated the associated neural activity is decreased ("repetition suppression", see Grill-Spector, Henson & Martin, 2006, for a review). It has been suggested that this phenomenon could be related to the neural correlates of sensory attenuation (Waszak et al, 2012). In the present study we discuss in detail this notion and we assessed whether stimulus repetition leads to the same perceptual effect as the prediction of an actions outcome.

Repetition priming refers to the change in the ability to perform a task on a stimulus as a consequence of a former encounter with that very same item (for a review see Schacter & Slotnick, 2004). Repetition priming has been used extensively in behavioural experiments to investigate implicit memory. It is usually assessed using cognitive tasks in which participants have to identify repeated and non-repeated stimuli, most of the time words or objects, or to make some sort of decision based on one or more features of the item. The main finding of priming studies is that task performance is improved when the target stimulus is repeated compared to when it is not repeated. Since priming can be dissociated from explicit recall and recognition of the items, this has been taken as evidence for the existence of implicit memory (e.g., Hamann & Squire, 1997).

Repetition priming has been shown to take place on different levels, from perceptual-conceptual (Sayres & Grill-Spector, 2006; Friese, Supp, Hipp, Engel & Gruber, 2012) to motor levels (Dobbins, Schnyer, Verfaellie & Schacter, 2004; Hsu & Waszak, 2012; Moutsopoulou & Waszak, 2012). On all these levels, repetition results in faster and more accurate performance. As concerns the underlying brain mechanisms, neuroimaging studies have demonstrated repetition priming to be based on decreased brain activity following item repetition (for reviews see Grill-Spector et al., 2006; Henson & Rugg, 2003). This decrease in activity across stimulus repetition has been referred to as repetition suppression. It has been demonstrated with single-cell recordings in monkeys (Miller & Desimone, 1994), as well as in humans using fMRI (Grill-Spector & Malach, 2001) and EEG/MEG (Henson, Rylands, Ross, Vuilleumeir & Rugg, 2004). Several models have been proposed in the literature to account for neural repetition suppression (see Grill-Spector et al., 2006). The sharpening model assumes that repetition suppression is due to a repeated stimulus being represented more sparsely, because the firing rate of neurons responding to irrelevant features decreases across repetitions (Desimone, 1996; Wiggs & Martin, 1998; Kok, Jehee, & de Lange, 2012). According to the facilitation model, repetition suppression is due to stimuli being processed more quickly when presented more than once (i.e., neurons firing for a shorter period of time) (Henson & Rugg, 2003; James & Gauthier, 2006). The fatigue model, finally, suggests that repetition suppression results from neurons being less responsive when a stimulus is repeated due to firing rate adaptation and synaptic depression (Grill-Spector & Malach, 2001; Kaliukhovich & Vogels, 2011). Grill-Spector et al. (2006) speculate that the different models explain repetition suppression at different time scales, with the sharpening model and the facilitation model accounting for long-term repetition suppression across intervening trials and the fatigue model accounting for short-lived repetition suppression that operates within a few hundred milliseconds after immediate repetitions of a stimulus.

Recently, repetition suppression has been discussed as a possible mechanism for a phenomenon called sensory attenuation (cf., Waszak et al., 2012). Sensory attenuation refers to the finding that self-generated stimuli are perceived as less intense than externally generated stimuli (e.g., Blakemore et al., 1998; Sato, 2008; Cardoso-Leite et al., 2010; Roussel et al., 2013). Several studies also investigated this effect with neurophysiological correlates of anticipated action effects. It has been shown that when the stimulus is anticipated the neural response is reduced compared to externally triggered or not anticipated stimuli (e.g., Schäfer & Marcus, 1973; Blakemore et al., 1998; Baess et al., 2008; Aliu et al., 2009; Gentsch & Schütz-Bosbach, 2011; Hughes & Waszak, 2011; Hughes et al., 2013b; Roussel et al., under review). Waszak et al. (2012) suggest that sensory attenuation of self-produced stimuli (i.e., anticipated action effects) is based on similar if not identical neural mechanisms as repetition suppression. As outline above, they suggest that sensory attenuation may be due to a change in the baseline-to-signal ratio due to a preactivation of the sensory network (Cardoso-Leite et al., 2011; Waszak et al., 2012; Roussel et al., 2012). Under this account action effect anticipation would result in the specific preactivation of the coding network. Sensory attenuation would reflect the decrease in differential activity between an increased baseline due to preactivation and the signal activity.

Importantly, if action effect anticipation, as suggested by the preactivation account, results in activity similar to the activity triggered by the true perception of the anticipated stimulus, then the activity triggered by a correctly anticipated action effect is actually the repetition of the anticipation-based activity. As a consequence, a correctly anticipated action effect should be subject to the same neural dynamics as a stimulus that is truly repeated, at least to a certain degree. Sensory attenuation could, thus, be due to neural sharpening, facilitation, and/or fatigue. For example, fatigue of a network of neurons coding for a stimulus is basically equivalent to an increase of the baseline activity as described above. Under a fatigue hypothesis, a neuron that has been preactivated as a consequence of internal effect anticipation would not be available to code for the same stimulus when the true action effect appears.

Note that the effects of repetition priming and action effect anticipation on performance seem to be conflictive. Repetition priming has been demonstrated to enhance performance (Morton, 1969; Tenpenny, 1995), whereas action effect anticipation results in sensory attenuation, i.e., a decrease in perceptual performance (Cardoso-Leite et al., 2010; Roussel et al., 2013). However, studies on sensory attenuation are usually psychophysical studies in which participants are required to detect a stimulus in noise or to indicate the stimulus’ intensity. Repetition priming experiments, on the other side, never used a psychophysical detection protocol. Instead, participants are usually required to respond the identity of the stimulus. Identification differs from detection, in that the differentiation between two stimuli is not based on the activity in one selective neural population. Instead, in identification tasks the perceived difference between stimuli relies on the competition between different neural populations. As has been outlined above, a preactivation account predicts that although anticipated stimuli are perceived less easily when they have to be detected in noise, they are identified more easily when they have to be discriminated from another stimulus (cf., Roussel, Allenmark, & Waszak, in prep. see Study 6). Improved identification of correctly anticipated action effects has indeed been demonstrated recently (Desantis, Roussel, & Waszak, under review; study 4).

If mechanisms of action effect anticipation and repetition priming are similar, then we should observe the same perceptual effects in case of a repeated stimulus as in case of a stimulus that is anticipated by an action. That is, a repeated stimulus should be perceptually attenuated. The present study tests this notion. We adapted a luminance discrimination protocol that has been used before successfully to assess perceptual sensitivity and response bias of anticipated and unanticipated visual action effects (Roussel et al., 2013, Roussel et al., [under review]). Roussel et al. made participants learn an association between left and right key presses and the presentation of the letters A and H, respectively. They then made participants perform left and right key presses that randomly triggered presentation of either an H or and A at one of two possible contrasts. Participants were required to make discrimination judgment between the two contrasts. They showed contrast discrimination to be worse when the prediction (H or A, as learned during the association phase of the experiment) matches the true stimulus, demonstrating sensory attenuation. The present study uses the same protocol. However, instead of testing contrast discrimination of stimuli that have been correctly anticipated by an action or not, we assessed contrast discrimination when a (Gabor grating) stimulus is congruently repeated (same orientation) versus when it is incongruently repeated (different orientation). We predicted that contrast discrimination performance will be worse for congruent than for incongruent stimuli.

## Experimental part: Repetition Priming results in sensitivity attenuation

### a Material & Method

#### ***Stimuli***

Experimental stimuli were generated and presented with Matlab 2007b using the psychophysics toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007). The stimuli were two tilted Gabor gratings (45° left, 45° right) displayed on a 15 inches CRT monitor at a 80Hz refresh rate and fitted into a virtual square of 2.9 degrees of visual angle. The screen resolution was set at 800*600 pixels. These two stimuli were presented at two different contrast values (C0 and C1; see contrast determination phase) at the center of the screen. The primes were identical to the target stimuli except for the luminance which was set at 10% more than C1 (to be sure that the prime exceeded the detection threshold). We used in all our phases a uniform noise texture to increase perceptual variance. This noise was a 100*100 matrix filled with an equal number of white and black pixels. The matrix was re-sampled (using the randperm function in Matlab) on each screen refresh. The mean luminance of the noise was then equal to the gray background.

### ***Contrast determination phase***

In order to determine individual contrast values C0 and C1 yielding a discrimination d' of about 2, every participant completed a psychophysical staircase converging on 85% correct responses in an orientation identification phase (Right versus Left tilt). We employed an adaptive staircase to manipulate stimulus transparency using an accelerated stochastic approximation algorithm as described by Kesten (1958). The initial step size was 20 and we stopped the staircase when the step size was 1 (in transparency). The correct response rated was used to ensure that the stimuli were supraliminal and that we could then independently manipulate discrimination. We used this contrast value as the referential contrast C0 in a 2AFC paradigm (with constant stimuli ranging from C0 to C0+12%) in order to calculate, the contrast value of C1 yielding 85% correct responses in a contrast discrimination task (C0 vs. C1). For an ideal observer, this contrast yields a discrimination d' of around 2 (Mcmillan & Creelman, 1991). Previous experiments showed this procedure to work reliably.

For each subject the prime contrast was set at C1+10% transparency to ensure perfect visibility of it.

### ***Test Phase.***

Participants fixated a square at the center of the screen, just as in the contrast determination phase. The offset of the prime was followed in a fix interval of 100ms by the target onset. Both prime and target had a 200 ms duration. The right or left orientation were presented randomly, such that 50% of the stimuli were congruent with the prime (i.e., the Gabor grating were identical), and 50% were incongruent (i.e., the

Gabor grating had the opposite orientation). The target stimuli appeared randomly (but in equal proportions) with the luminance C0 or the luminance C1. Participants were told that there were two luminance values and examples of these two were shown before each test block. After the target stimulus had disappeared participants were required to judge the luminance value of the stimulus in a forced choice. Participants completed 10 tests blocks of 40 trials.

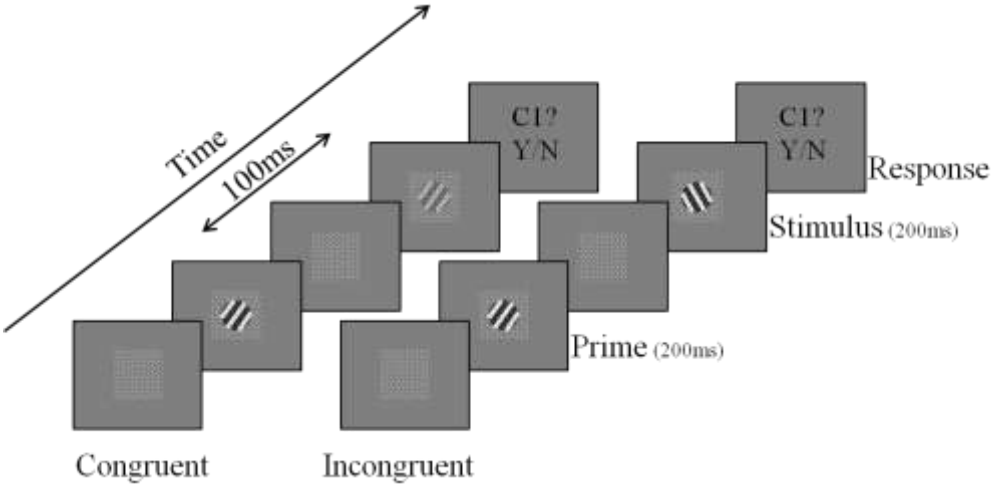

**Figure 16:** Representation of the experimental design. Congruency was manipulated by the match between the prime and the target. The target could appear at 2 different contrast value, C0 or C1. The Participant had to indicate the contrast at which the target appeared.

### ***Participants***

Fifteen participants took part in the experiment. They were naive to the purpose of the experiment. Two of these fifteen participants were excluded from the analysis as their mean luminance discrimination d’ was below 1.

### ***Analysis of discrimination performance***

The luminance discrimination task was considered to be a yes/no protocol, with C1 being the target. That is, a C1 response to a C1 stimulus is a hit, a C1 response to a C0 stimulus is a false alarm, etc. d’ and c are calculated using $d' = z(\text{hit rate}) - z(\text{false alarm rate})$ and $c = -0.5 \times [z(\text{hit rate}) + z(\text{false alarm rate})])^3$. D’ and c were calculated separately for congruent and incongruent trials.

## b Results

We analyzed our data dependent on participants' contrast discrimination. Discrimination performance (d') was lower in the congruent condition (d' congruent: 1.79) than in the incongruent (d' incongruent: 2.04) condition. A repeated measure analysis of variance including the factor of congruency showed this effect of congruency on d' to be significant (F(1,12)=7.033, p=0.02). At the same time, the criterion was not different in the two congruency conditions (F(1,12) = 0.324, p = 0.58) (congruent: 0.40 , incongruent : 0.44).

## c Discussion

In this study we assessed the hypothesis that stimulus repetition, which is found to induce repetition suppression at the neuronal level, is associated with decreased performance in contrast discrimination. We showed that when the prime and the target were identical, participants were less able to discriminate stimulus intensity. The finding that a repeated stimulus was perceptually attenuated resembles the common finding of sensory attenuation research, where self-generated stimuli are perceived as less intense and more difficult to be detected. The decreased performance is likely due to the change in the baseline-to-signal ratio, where there is a decrease in differential activity between an increased baseline and the signal activity.

This result supports the idea that repetition suppression may be the underlying mechanism for sensory attenuation (Caroso-Leite et al., 2011; Waszak et al., 2012; Roussel et al., 2013). The preactivation account in action effect anticipation research predicts that anticipated stimuli are perceived less easily when they have to be detected in noise because there is a decrease in differential activity between an increased baseline and the signal activity.

A possible scenario is that anticipation preactivates the corresponding neuronal representation, increasing the neuronal activity at the baseline. Given that neurophysiological data are baseline-corrected, trial-by-trial, more neuronal activity was removed from the activation triggered by anticipated stimuli. This manifests as a reduction in neuronal responses triggered by anticipated stimuli.

Our interpretation is also compatible with the predictive coding model (Friston, 2005, 2009; Egner, Monti, & Summerfield, 2010). The predictive coding model postulates that perception entails two distinct neurocomputation components, the top-down propagation of predictions and the bottom-up propagation of prediction errors. The brain is believed to work as a predictive machine constantly attempting to match sensory inputs with predictions (Clark, 2013). The preactivation account mainly describes the activity in the prediction units, where there is increased activation before stimulus onset. Such preactivation of the corresponding neuronal representation would lead to a reduction in neuronal responses triggered by anticipated stimuli as well as improved stimulus identification. Using multivariate pattern analysis techniques on fMRI data, Kok and colleagues (Kok, Jehee, & de Lange, 2012) also found that perceptual expectation reduces the neuronal response amplitude but improves the stimulus representation in the primary visual cortex.

Obviously, the predictive coding model shares a number of features with the preactivation account outlined above. Therefore, in the following chapter, we aimed to unify the two models.

# F. Study 6: A Predictive coding account of preactivation

*This Chapter is based on: **Roussel**, C., Allenmark, F. & Waszak, F., Toward a better understanding of action effect prediction: A predictive coding model of sensory preactivation. [In Prep]*

Philosophers have for a long time been referring to the concept of *Umwelt*, which is the fact that our knowledge of the world relies on the physical properties of our bodies. The notion that perception depends on the perceptual organs that allow us to perceive the external world raised another question. How is it possible to recognize a given object although the sensory input we receive depends on the context (for example light, orientation …)? Helmholtz (1860), for example, distinguished “perception” and “sensation” in order to explain why from which ever any angle I observe my desk, it is only the sensation that changes (e.g., the visual information) whereas the perception remains the same (e.g., the knowledge of what it is). Hence perception appears to be the process by which we recognize the causes of our sensory inputs. The existence of such inferential processes in the perceptual system has been hypothesized and supported in visual and auditory domains (see Kersten, Mamassian & Yuille, 2004; Pressnitzer, Suied, & Shamma, 2011; for a review in the visual and auditory domain, respectively). These inferential processes have been studied under the Bayesian framework. In computational modeling the Bayesian framework has been implemented into empirical Bayes (a statistical implementation of the Bayesian probability theory). Empirical Bayes claims that prior expectations can be abstracted from the sensory data, using a hierarchical model of how those data were caused (Friston, 2002, 2003). Predictive coding theory implements empirical Bayes methods using a modern statistical approach, in which prior expectations are used to compute a prediction error, which, in turn, is used to adjust the state of the system until this error is minimized. This idea is also related to the forward/inverse model outlined above. In the comparator model (Wolpert et al., 1995; Wolpert & Ghahramani, 2000) the system compares input and expectation to infer whether action consequences are as expected. But unlike the forward/inverse model, the predictive coding theory, implement priors, prediction error and inference of the cause of the activity at every level of the perceptual process (for more details, see Friston, 2005).

Predictive coding is based on a hierarchical organization of the brain; each level tries to explain the activity at this level and the level below through the use of priors. These priors are implemented by two components. First, the synaptic organization between two types of units, Error units (Eneurons) and Representational units (Rneurons). This synaptic organization is shaped through learning and “represents” the statistical regularities of the systems environment. Second the activity coming from the backward connections, which represent the systems current expectation of the incoming signal. The activity in the Representational units will be compared by the Error units to the incoming signal from the level below. The activity in the Representational units is then adjusted to minimize the prediction error (see Figure 17).

The preactivation model postulates that future sensory input is predicted by the preactivation of the neural network coding for the actual stimulus. In the preactivation model of sensory attenuation the sensory pathway is preactivated through motor preparatory processes (see Chapter I.E and Study 1 & 2). Thus, the architecture of the predictive coding model can give a more detailed neural basis for this notion.

As a consequence, we crafted a visual predictive coding system, which can be preactivated by the motor system (or otherwise) and tested whether the model can explain some of data provided above.

## Modeling work:

### a Model

We based our model on Friston's (2005) predictive coding scheme. This model is a simplified version of the predictive coding model, without learning phase (that is, the synaptic weights of the connections between the errors neurons and the representational neurons are pre-set). It is constituted of one layer (e.g., level of interaction between R and E units; see Figure 17). In this layer two Rneurons were modeled and each was optimally tuned to respond to an oriented Gabor grating (45° and – 45° from the vertical orientation respectively). The receptive field of those neurons covered a square of 2.9° visual angle, such that they could process the exact same stimulus as presented in Experiments 3 and 5.

We modeled 1681 error neurons corresponding to the number of pixel contain in our stimulus. Each of these Eneurons compares the Rneurons activity to its corresponding pixel. The stimulus was a Gabor grating fitted into a square of 2.9° visual angle oriented in one of the two possible orientations. For modeling purposes we represented the stimulus in a 1x1681 vector containing values from -1 to 1, -1 coding for fully black and 1 fully white (noted I). In our model this stimulus was considered to be at the maximal strength. That is to say, the contrast of this stimulus would saturate the R neurons' activity.

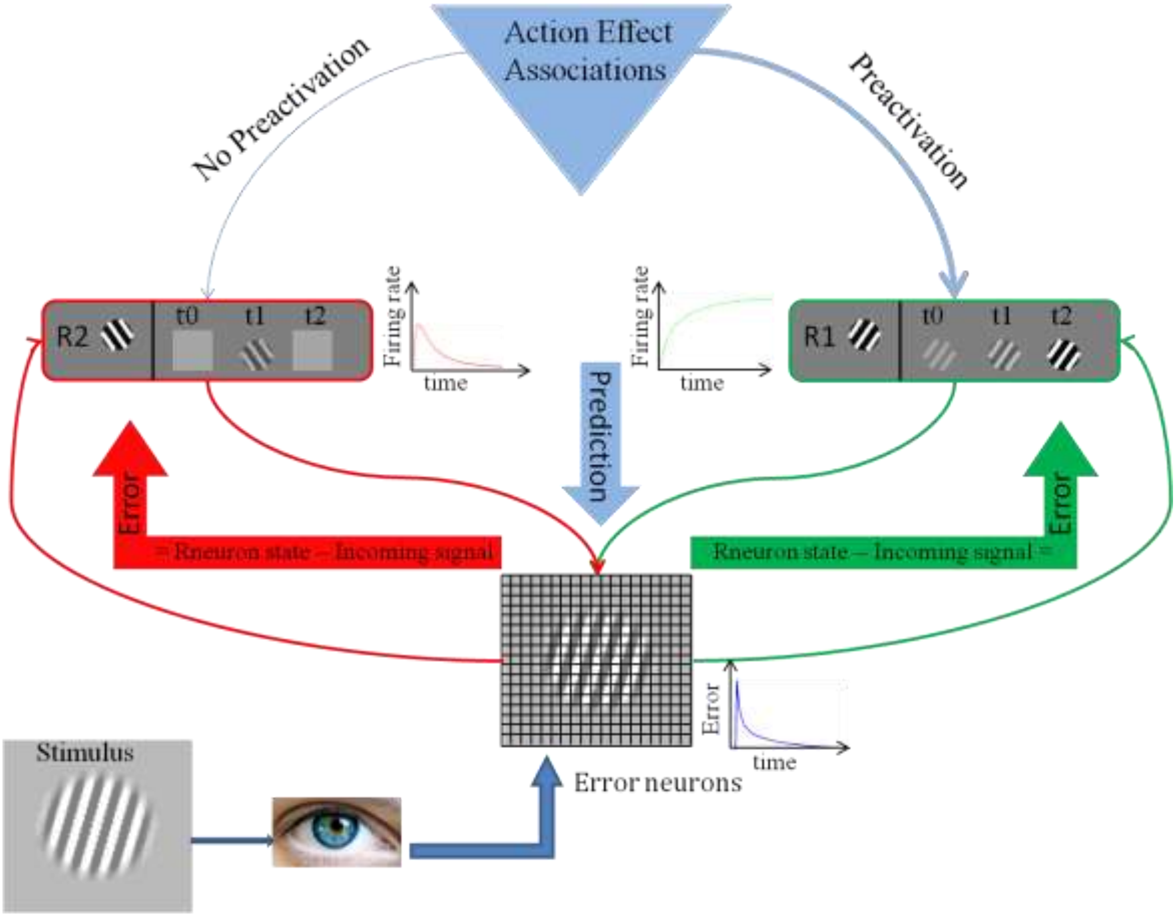

**Figure 17:** Representation of our predictive coding based model of sensory preactivation. When a stimulus is presented the Errors neurons computes the difference between the incoming signal and the corresponding Representational neurons activity. This error will by way of forward connections be transmitted to the Representational neurons which will adjust their activity to diminish the prediction error by an optimization process tuned to infer the cause of the stimulation. Activity in R1 predicts that the source of the stimulation is a rightward tilted grating, while activity in R2 predicts the presence of a leftward tilted stimulus. In the case of a preactivation induced prediction of a rightward tilted stimulus, R1 is already activated before the stimulus onset (t0) while R2 is not. At the time of stimulus onset (t1) the error is maximal and different hypothesis about the source of this prediction error are "tested" (both R1 and R2 fires). At the end of the optimization process (t2) the causes of the activity (a rightward tilted stimulus) had been inferred, thus the prediction error is null and R1 is the only representational neuron to fire.

To do so we modeled the Eneurons and Rneurons with a hyperbolic tangent function such that the neural activity was nonlinear and constraint between -1 and 1 (noted *th*(x)). In order to vary the stimulus contrast we set a parameter varying from 0 to 1 to multiply the Image vector(I).

According to the Predictive Coding framework any stimulus presentation will generate error given the synaptic arrangement of the system. Then an optimization process will by gradient descent modify the activity of the R-neurons in order to minimize the total error.

The error (E) equals the difference between the product of the synaptic weight matrix (w) with the previous state of the Rneuron and the current input. Such that for each step k of the process $E_k = th(s * w * th(R_{k-1} * s) - th(s * I))$.

The optimization by gradient descent giving the subsequent activity in the Rneurons was a minimization of the error as a function of the Rneuron activity. The gradient descent is performed in the algorithm by taking a step $\gamma$ in the direction of the gradient at each step of the optimization process. In our algorithm the gradient was implemented such that for each step k of the process

$$R_k = R_{k-1} - (diag(1 - th(w * R_{k-1} * s)) * w * diag(1 - th(R_{k-1} * s) * E_k * \gamma$$

The algorithm was programmed to stop the optimization process when the error changes by less than 0.00001 between steps. Finally to study the Preactivation Hypothesis we could preactivate the Rneurons. Since the neuronal activity was constrained between -1 and 1 the preactivation value could be between these values.

## b Results

By looking at the model results in the total absence of noise we can observe that the Rneuron response to a stimulus converges to an activity equal to the presented intensity (Figure 18).

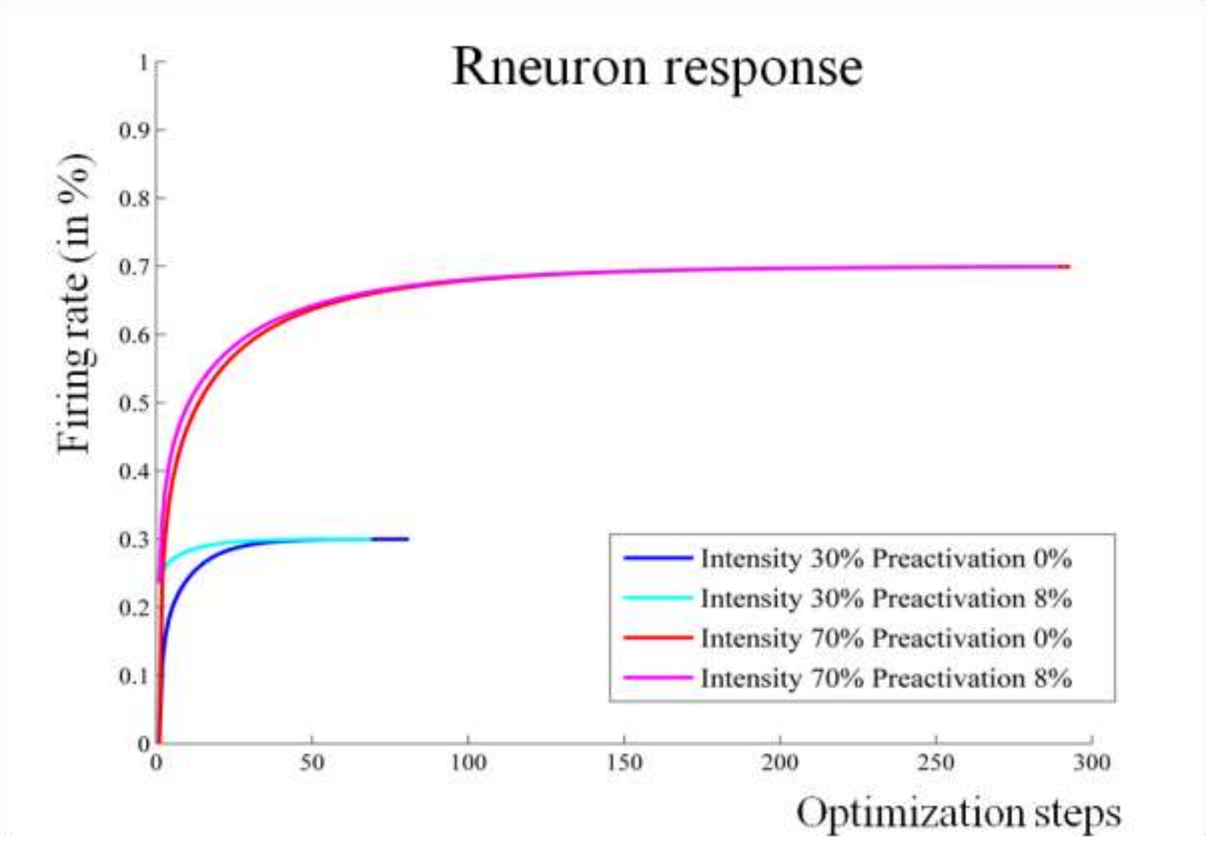

**Figure 18:** This figure shows the activity contained in R1 when a rightward tilted stimulus is presented to the model at different intensities, when there is or not preactivation in R1.

That is, the model could represent perfectly the stimulus at any given intensity. Despite the apparent similarity with our previous model (Roussel et al., 2013; Study 1), we would like to point out that there are substantial differences between this and the former model. The two main differences are the following. First the neural activity is changing over time for a given stimulus intensity while the former model considered just a global activity function of the stimulus contrast. Second in this model the activity with or without preactivation finally converges to the same point, just as in the predictive coding model, indicating that the cause of the activity had been properly inferred. However, the integration over time of the neural activity is higher when there is preactivation than when there is not. Note, however, that just as in the preactivation hypothesis this increase due to preactivation is a nonlinear function of contrast. That is to say, the gain in activity due to preactivation decreases the more the contrast increases. To further test the preactivation hypothesis with the current model, we tested whether the model could fit the results we obtained in Experiments 3 and 5. To do so we choose to base the model response on the 50$^{th}$ step in the optimization process (more detail about this will be provided in the discussion).

### ***Sensitivity Attenuation / Repetition suppression***

To model Sensitivity Attenuation we reproduced the protocol introduced by Roussel et al. (2013). These authors made participants learn an association between left and right key presses and the presentation of two different stimuli respectively. They then made participants perform left and right key presses that randomly triggered a stimulus presented at one of two possible contrasts. Participants were required to make a discrimination judgment between the two contrasts. This paradigm was equivalent to the paradigm used in Experiment 5, with the difference that the congruency was not driven by the action effect contingencies but by the fast repetition of two identical stimuli. For the sake of coherence with the modeling in Experiment 1 the preactivation was set to a level of 8%. Moreover, C0 and C1 values were set to the mean C0 and C1 values used in Experiment 5 (in terms of the maximal contrast value (just as in our first model)). C0 and C1 were, thus, set to 27.13% and 32.17% of the maximal contrast respectively. In order to model the discrimination response we set a threshold at the mean of C0 and C1 (29.65). That is, once the activity of the R neuron reached the threshold the model would concludes that the contrast is C1, C0 otherwise. To keep the model test as similar as possible to the experiment we added noise to the stimulus. The signal to noise ratio was estimated in order to fit the experimental d'. We found that with 13% of noise the d' with preactivation was 2.07 while it was 2.28 without preactivation. This discrepancy with the actual experimental d' in study 5 (d': 1.79 and 2.04 with and without preactivation respectively) can be explained by the fact that we didn't yet find the best model parameters to fit the data. But it is to be noticed that 8% of preactivation results in a d' reduction of about 10% just as in our experiments, in the first model and in Cardoso-Leite et al. (2010).

Regarding to the Error Response (figure 19) until the error converges to the minimal point preactivation induces globally less prediction error.

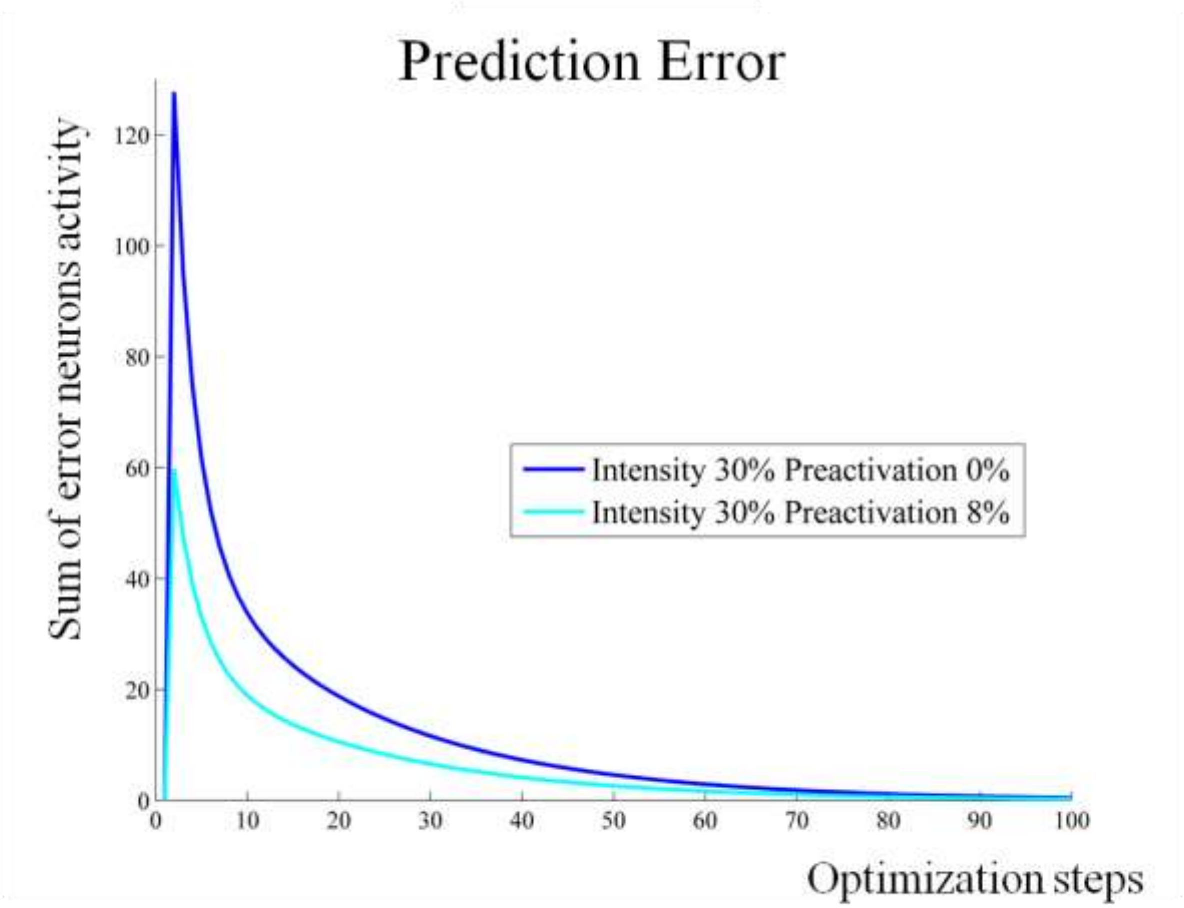

**Figure 19:** Sum of the absolute value of the error neurons' response. It appears that when there is preactivation the error induced by the presentation of the stimulus is less important until the prediction error had been minimized.

### ***Increased Identification***

It has been shown by previous experiments that action-effect associations have a facilitatory effect on identification performance (Study 3 & 4). The preactivation hypothesis postulates that, while preactivation leads to sensitivity attenuation along one dimension, it leads to increased perceptual distance between two different dimensions (e.g., it will reduce the competition between the possible alternatives; see Figure 13 in part II.C). To test whether the predictive coding model behaves in line with this hypothesis we looked at the response given by our two representational neurons for a rightward tilted stimulus. The response pattern of both R1 and R2 without noise added to the stimulus is presented in Figure 20.

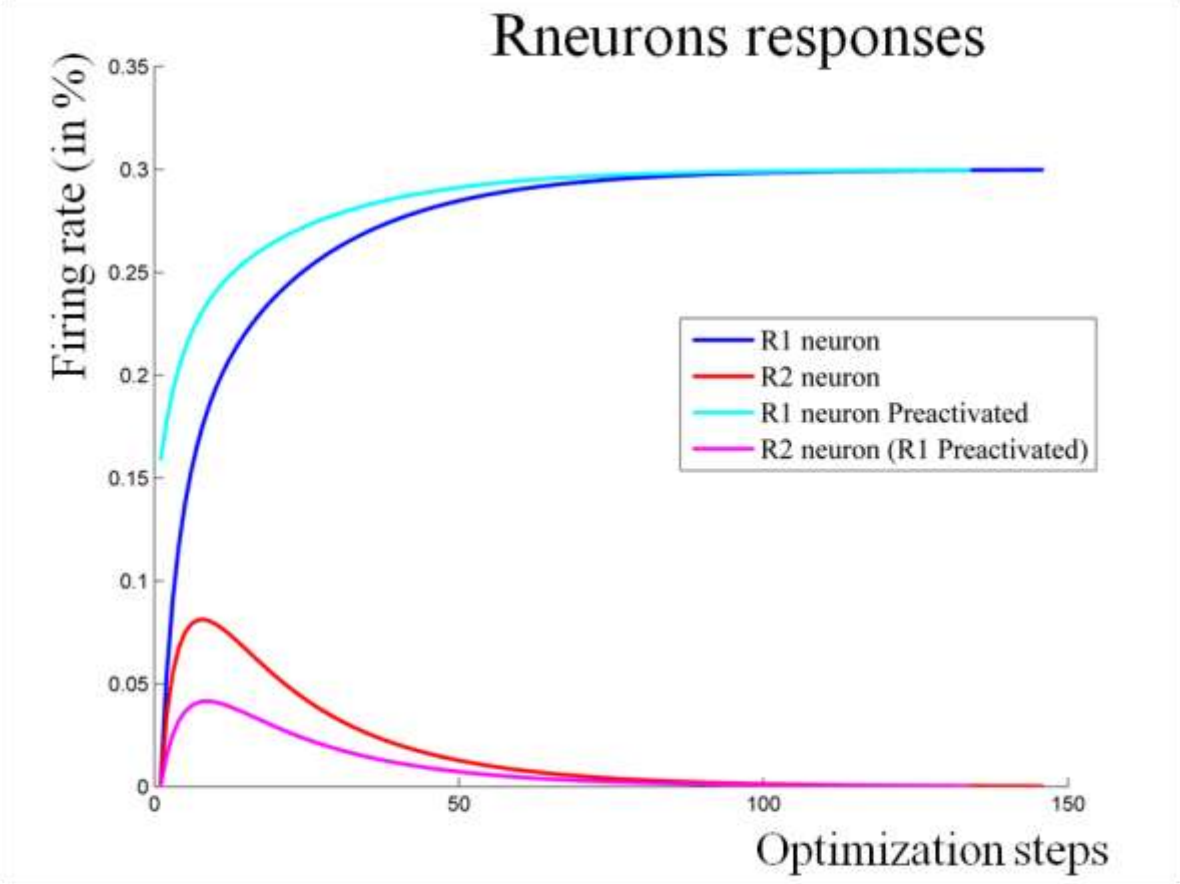

**Figure 19:** Activity in the Rneurons R1 and R2 for a rightward tilted stimulus, with and without preactivation.

The response of R1 is similar to the response observed in Figure18. Concerning the response of R2 to a rightward tilted stimulus it is to be noted that there is activity in R2. This suggests that the model first also “tests” the hypothesis that the source of the activity is a leftward tilted stimulus. But since this activity in R2 does not minimize the error it decreases until the error is minimized by the activity in R1.

Moreover preactivation (of R1) increases the activity in R1 (just as in Figure 18) and reduces the activity in R2. This is due to the fact that the response in R1 and R2 are linked through the error neurons. A preactivation in R1 reduces from the first step on the prediction error, making the hypothesis of a leftward tilted stimulus “less likely”.

In order to fit the experimental data provided in Experiment 3 we used the mean contrast value of Experiment 3 (in percentage of the maximal contrast (22%)). To obtain a close fit of the experimental d’ we had to decrease the signal to noise ratio. To do so we increased to 40% the noise intensity. This decrease in the signal to noise ratio was necessary because our simplification of the predictive coding model made the Rneurons very resistant to noise. The model responded following a simple competition rule. The “perceived” orientation was given by the Rneuron discharging the most (i.e., if the activity in R1 was higher than in R2 the model responded that the stimulus was rightward tilted and vice versa). Doing so, the model resulted in d’-s of about 2 and 1.8 with and without preactivation, respectively (we discuss this fit in detail in the discussion).

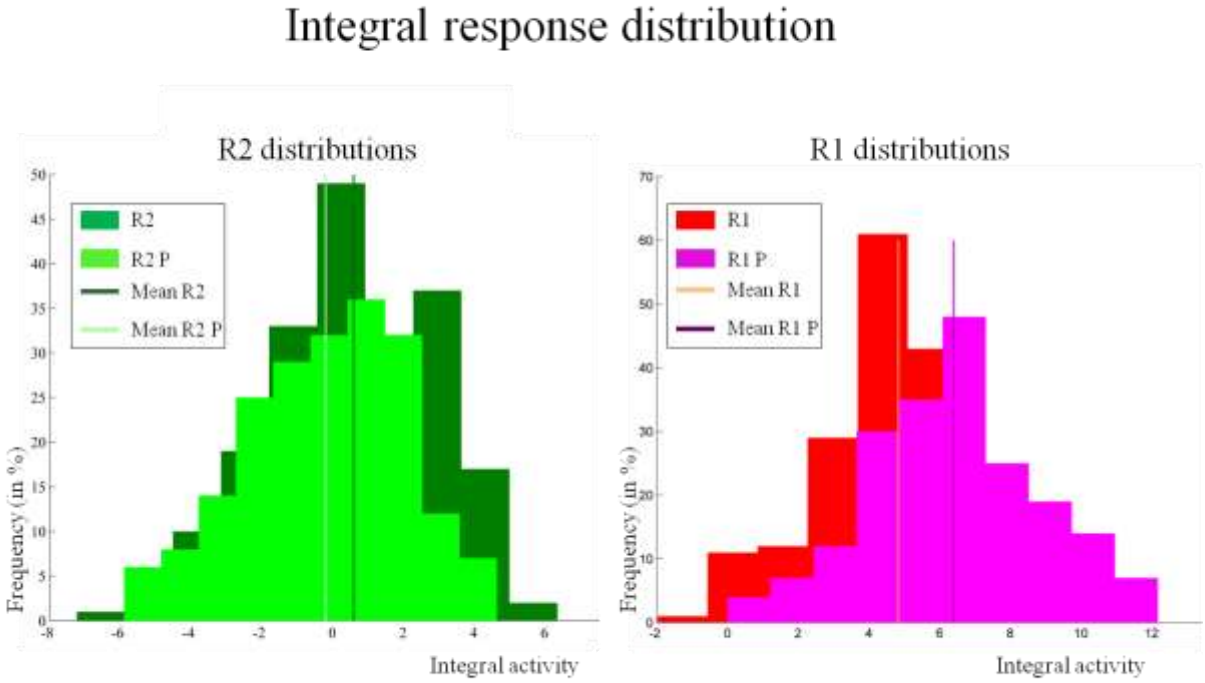

**Figure 21:** Left panel) Integral distribution of R2 responses to a rightward tilted stimulus, when R1 is preactivated (light green) or not (dark green). Right panel) Integral distribution of R1 responses to a rightward tilted stimulus when R1 is preactivated (magenta) or not (red).

Regarding the integral response distributions, the pattern is similar to the one presented in Figure 20. The distribution of R1 responses (Figure 21 left panel) to a rightward tilted stimulus shifts toward more activity when R1 is preactivated while the distribution of R2 responses (Figure 21 right panel) to a rightward tilted stimulus shifts toward less activity when R1 is preactivated.

## c Discussion

In this study we crafted a predictive coding model detecting Gabor gratings of two different orientations. Based on this architecture we tested the preactivation hypothesis of action effect prediction. We postulated that preactivation results in pre-stimulus activation of the corresponding representational units. We have shown that our model is able to explain perceptual consequences of effect anticipation such as sensitivity attenuation and the improvement of identification performance.

However, we would like to point out that the fit of our experimental data is not optimal yet, as we did not yet find the best model parameters. This is mostly due to the fact that the simulation takes a lot of time to estimate the parameters. Note that the fits of Experiment 3 (Identification paradigm) are more deviant from the actual data than the fits of Experiment 5. This is due to the fact that we modeled a congruent vs. neutral condition since we assessed the contrast between with preactivation vs. without preactivation. That is, we showed that the identification d' in congruent trials is better than in neutral trials, whereas the d' in incongruent trials is worse than in neutral trials (see Appendix 3).

It has been suggested that EEG signals are primarily based on activity of the error neurons (Friston, 2005; Friston & Kiebel, 2009). This suggestion is based on the assumption that the Error units consist in the superficial pyramidal cells of the brain while the Representational neurons are thought to be located in deeper layers (Mumford, 1992; Friston, 2005; Friston & Kieble, 2009). The notion that neurophysiological data, like early sensory ERPs, reflect the error signal provides a nice account of the compatibility between the preactivation hypothesis and the neural correlates of sensory attenuation and repetition priming (for short terms repetition). On the one hand, the motor based prediction of an action effect (as well as the repetition of a stimulus) results in a decrease of the observed ERP components. This has been taken to indicate that the predicted stimulus induces less activity than the unpredicted stimulus. On the other hand, the preactivation hypothesis claims that action effect prediction should results in a general increase of activity in the network representing the effect. However, as we demonstrated above, the prediction error elicited by a stimulus decreases in the presence of preactivation. It is only the activity in the Rneurons that increases.

Our decision to extract the model data at the 50th step as the model response on a given trial needs to be discussed. Since the model converged on the same activity for a given contrast value, irrespective of whether there is preactivation or not, we could not simply use the end point of the process. Apart of using a particular step of the algorithm, we saw two other possibilities. First, we could integrate the Rneuron activity until the end of the process. Second, we could set a threshold the model response has to reach.

However, the integration of activity until the end of the optimization does not lead to comparable results between conditions, since the number of steps the model needs to converge differs depending on the stimulus intensity and the preactivation activity.

The absolute threshold method is not a suitable option either. As a matter of fact, since without the addition of noise the convergence activity is identical irrespective of the preactivation, a threshold in noisy conditions would merely reflect the effect of the noise.

In order to capture the effect of preactivation in a comparable way between the conditions, the intensity values and the different fits, the best option consisted in finding a processing step were we could compare every combination of these parameters. We opted for step 50.

Admittedly, it is difficult to relate the optimization steps of our model to the temporal dynamics of the neural process. The brain does not necessarily minimize prediction error by gradient descent. However, we would like to point out that the model dynamics very well represents well known RT phenomena. It has been shown, for example, that the response time to a stimulus decreases with increasing stimulus intensity (Luce, 1959). Our model makes very similar predictions. The stronger the stimulus is, the fewer optimization steps are necessary to get to the same threshold activity (see Figure 18; the increase of R1 activity is sharper for intensity 0.7 than 0.3). It has also been shown that the perceived intensity of a stimulus varies as a function of stimulus duration until it reaches a saturation point (Stevens & Hall, 1966). This is in line with the activity in the Rneurons of our model. Moreover, there is ample evidence for the existence of speed accuracy trade off (for review see Wickelgren, 1977): accuracy is an inverse function of the speed of the perceptual decision. In our model the step at which the perceptual decision is taken fits exactly this notion: the sooner the optimization process takes a decision the less probable it is that the inferred cause of the activity is accurate. Thus, we consider the optimization process to represent the temporal dynamic of the perceptual process rather well.

Note that by assessing the model data before the end of the optimization we postulate that the perceptual difference between predicted and unpredicted stimuli emerge during the stimulus processing rather than at the end of the process.

The decision to use exactly the 50th steps is based on more practical reasons. For instance, the number of steps needed to get to the convergence activity differed trial by trial as a function of the noise, the stimulus intensity and the preactivation. Hence, in order to have the same number of observations in the different noise values (Sensitivity attenuation fit and Identification fit), the different conditions (with or without preactivation), and the different contrast values (C0, C1 and the contrast used in the identification fit), we had to estimate the latest step reached by any of the studied combination.

To conclude, we provided with this model a general support for the preactivation hypothesis. The model shows that a single neural mechanism can explain sensitivity attenuation and facilitation of identification performance (in the perceptual domain), as well as sensory and repetition attenuation (in the neurophysiological domain).

# III. General Discussion

The ideomotor principle claims that an action and its outcome are linked by bidirectional associations and that actions are controlled by the internal activation of the desired/anticipated effect (Lotze, 1852; Greenwald, 1970; Hommel et al., 2001). Based on this notion, Waszak et al. (2012) introduced the preactivation account supposed to explain several perceptual phenomena observed in the domain of action control, for example, sensory attenuation. In this present thesis this preactivation hypothesis was assessed in detail.

In study 1, we introduced a model of the preactivation hypothesis derived a new hypothesis about sensitivity attenuation (i.e., we predicted a decrease in contrast discrimination performance (without prediction vs. with prediction)), and tested this hypothesis experimentally.

In study 2, we addressed whether neurophysiological indices of sensory attenuation reflected early, low-level or later, higher-level mechanisms. Moreover, we explored how neurophysiological and perceptual indices of sensory attenuation relate.

In study 3 we extended the preactivation model to include identification decisions.

In study 4 we assessed the temporal dynamics of the prediction of action effect and we tested the extension of the preactivation hypothesis provided in study 3.

In study 5 we addressed the possible relationships between repetition priming and the preactivation hypothesis.

In study 6, finally, we introduced a new preactivation model based on a predictive coding architecture. We also showed that this model explains the effects demonstrated in the previous studies.

In this discussion we discuss three main points. First, we will redefine ‘preactivation’ in the light of the findings outlined above. Second, we discuss our results and the preactivation hypothesis within a more general view of the literature. Third, we address the question of how the preactivation hypothesis might help us to better understand voluntary action control.

## What is preactivation?

The preactivation hypothesis refers to the idea that the prediction of an action effect relies on the preactivation of the specific sensory network which would be involved in the processing of the (predicted) stimulus. In this thesis we demonstrated that such a mechanism indeed might explain a range of perceptual phenomena that have been observed in the domain of action control.

Preactivation is thought to be specific to the predicted stimulus. All studies presented above confirmed this claim, by assessing the effect of preactivation systematically with stimuli that matched or did not match a previously learnt action-effect association. Also the term preactivation refers to the activation of the specific sensory pathway *before* the presentation of the stimulus. In Experiment 4 we demonstrated that effect anticipation can influence perception from about 220ms before the action's onset. This is in line with the idea that preactivation is related to the preparatory stages of the action and also, of course, with the ideomotor theory of action control claiming that an action is selected by the internal anticipation of its consequence (Greenwald, 1970; Hommel et al., 2001; Prinz, 1997). In study 1, 2, 5 and 6, we demonstrated that the preactivation model predicts a sensitivity loss in contrast discrimination tasks (congruent trials worse than incongruent trials) and we provided support in favor of this hypothesis. In study 3, 4 and 6, we demonstrated that the preactivation of the sensory pathway should result in the facilitation of identification performance and provided support for this hypothesis. Hence, it appears that the preactivation model could explain how the prediction of action effects is implemented in the brain.

## How is the preactivation hypothesis related to the existing literature?

It has been highlighted in the literature (e.g., Blakemore et al., 1998, 2000) that the prediction of action effects results in a diminution of the perceived intensity of the effect. Under the preactivation hypothesis, this diminution in the perceived intensity is accompanied by a decrease in strength discrimination performance (thus making more difficult the perception of variations stimulus in strength). According to Blakemore et al., (1998, 2000) tickling oneself is impossible, because the sensation of the consequences of our own actions are attenuated. According to the preactivation tickling oneself is impossible because we are less able to feel the variation of strength during a self generated tactile stimulation.

However, while the link between the claim that the perceptual strength of a self-generated stimulus is reduced and the observation that the neural response to a self-generated stimulus is attenuated is obvious (Blakemore et al., 1998, 2000), the link between the preactivation hypothesis and the observation of a reduced neural activity is less clear.

In this thesis we proposed several hypotheses of how the preactivation model and the neural correlate of sensory attenuation could relate.

The first possibility follows the fatigue model of repetition suppression (see Grill-Spector et al., 2006 and Waszak et al., 2012). This model postulates that a neural network is less responsive to the same stimulation presented in rapid succession due to firing rate adaptation and synaptic depression. In this scenario neural fatigue due to preactivation would lead to an attenuated response of the network in response to the true effect. Note, however, that this model can hardly explain the facilitation in identification performance for predicted action effects. As a matter of course the decrease in the activity of the relevant neural population increases the probability of higher responses from irrelevant neural populations. This would result in a decrease in identification performances.

The second scenario introduced above addresses this problem in that it postulates that preactivation of neuronal representations increases the baseline activity. Given that neurophysiological data are baseline-corrected, more neuronal activity is removed from the activation triggered by the stimulus if the stimulus is anticipated than if it is not anticipated. This would manifest as a reduction in neuronal responses triggered by anticipated stimuli. However, using our protocol we could not test this hypothesis (in study 2), because both actions were associated with an effect. We were thus not able to assess baseline activity.  Future research should address this notion directly.

The third scenario sketched above relies on a predictive coding architecture. It has been claimed that the stimulus-driven EEG signal typically observed over perceptual areas is actually based on activity of error neurons (superficial pyramidal cells) whose signal is easily picked up by EEG (Clark, 2013; Mumford, 1992; Friston, 2005; Friston & Kiebel, 2009). In such a scenario (see study 6), it is perfectly possible that the activity in the representational unit influenced by the preactivation increases while the neurophysiological signal measured with EEG decreases.

For the time being, it is impossible to decide which of these (mutually not exclusive) scenarios is true. However, it is evident that the notion of preactivation as introduced above does not contradict the neurophysiological data reported in study 2 and in the literature.

Waszak et al. (2012) suggested that the preactivation hypothesis could also explain the phenomenon of intentional binding. As we mentioned already (Chapter I.C and I.D), intentional binding refers to both a shift of the perceived action onset time toward the stimulus and to a shift of the perceived onset time of the effect toward the action. Waszak et al. (2012) suggested that the preactivation of the action effect would lead to a decrease in the perceptual latency of the predicted stimulus, since processing has a head start when there is preactivation. This could account for the shift in the perceived time of the action's consequence in intentional binding. The model in study 6 is in line with this notion, since it shows that fewer steps are needed to reach the threshold when an action effect had been predicted compared to when it has not.

However, whether or not intentional binding is caused by motor predictive mechanisms has been critically discussed by Hughes, Desantis, and Waszak (2012). In this review the authors address the question of whether sensory attenuation and intentional binding are based on the same mechanisms. They differentiated four types of mechanism: motor identity prediction, identity prediction, temporal control, and temporal expectation. They concluded that it is likely that sensory attenuation is driven by motor predictive processing (motor identity prediction). The experiments reported above corroborate this conclusion. However, they concluded that the case is much less clear concerning intentional binding. This conclusion is based, amongst others, on a study in which intentional binding was assessed while manipulating motor predictive processes and keeping the other factors constant (Desantis, Hughes, & Waszak, 2012). This study failed to show intentional binding in the critical contrast. Hughes et al. (2013) concluded that intentional binding might be driven by temporal control mechanisms rather than identity prediction. However in Desantis et al., 2011 (see Chapter I.G and Appendix 1) it has been shown that temporal control is not sufficient to induce binding. In this study the authors manipulated the belief of agency (i.e., the participants believed in some trials that they produced the tone and in some trials they believed that the same tone was produced by somebody else). In reality the participant always produced the tone. The results showed a shift of the perceived time of the action effect only when the subject believed to be the author of the tone. Thus, even if they were in control of the time of occurrence of the action consequence no intentional binding was observed. The underpinnings of intentional binding thus need to be further clarified. For example, assessing the perceptual latency of the effect outcome using psychophysical techniques (such as temporal order judgement) would be more appropriate to test whether there is an actual decrease of perceptual latency in case of action effect prediction.

As it has been suggested in the literature and in the present work, the prediction of action-effects might result in a preactivation of the perceptual representation of the anticipated stimulus (Waszak et al, 2012). In this thesis we attempted to demonstrate that the preactivation of the stimulus representation indeed results in observable effects, but that the source of the preactivation has not necessarily to be motor based, but could also rely on the repetition of a stimulus (Study 5). In Experiment 5 we tried to relate the preactivation hypothesis to an effect known as repetition priming. We showed that the effect of repetition was comparable to the one observed for motor based effect prediction. Our next aim is to show that in the exact same repetition condition we will find a facilitatory effect on identification performance.

Note that repetition priming is not the only phenomenon that relates to the concept of preactivation. Mental imagery (i.e., the mental representation of an object or stimulus) can be also considered as an internal activation of the object representation (see Kaski, 2002 for a review on this topic). Some theories claim that the imagery involves at least partly the same processes as perception (Kossly et al., 2001; Pearson et al., 2008). Interestingly, it has been shown that imagining an object can impair the detection of that object (Perky, 1910; Craver-Lamley and Reeves, 1987; Craver-Lamley, Reeves & Asterberry, 1997). For a more detailed discussion of the relationships between mental imagery and the preactivation hypothesis see Waszak et al., (2012).

**How does the preactivation hypothesis help us to understand voluntary action?**

Humans' behavioral flexibility is largely based on the fact that we are able to produce desired effects in our environment. This ability necessitates that people are capable of predicting the sensory effect of their actions (cf., Hommel et al., 2001; James, 1890; Lotze, 1852; Prinz, 1997; Shin, Proctor, & Capaldi, 2010, Wolpert et al., 1995, Blakemore et al., 2000). It has recently been suggested that motor prediction is based on a template of the brain response to the predicted stimulus (Kuhn et al., 2010; Waszak et al., 2012; SanMiguel et al., 2013). However, although research has gathered a lot of evidence for the existence of motor anticipatory mechanisms (for a review see Waszak et al., 2012), only little is known about how the content of these prediction and the way the brain produces it.

In the present thesis we proposed a theory on the motor based predictive mechanism that is based on the preactivation of the sensory network representing the anticipated effect. The essence of ideomotor control is that action effect anticipation is an integral part of voluntary action selection. From this point of view, the anticipated effect must necessarily be represented in the brain before the action is executed. Otherwise, goal-directed behavior would be impossible. Along the line of the intentional cascade model of Pacherie we argue that the preactivation of the sensory consequences of the action should be situated between the end of the Proximal and the Motor level (i.e, at the end of the preparatory stage of action and during the execution phase). Just as the ideomotor theory we also postulate that the efference copy is not just a one-way road from motor processing to the prediction of the sensory outcome of the action. The predicted consequence of an action can also specify the necessary motor code and/or correct it (see inverse and feed-forward model, Chapter I.D; Wolpert et al, 1995; see Desmurget & Gafton, 2000 for review). This would mean that the preactivation is also involved in the representation of the goal, at least for some aspect of it (e.g. the representation of the situated and instantaneous goal in Pacherie's model (2008); see Chapter I.E Figure 3).

To conclude in this thesis we started with the hypothesis that action-effect prediction should rely on the preactivation of the specific sensory pathway representing the anticipated effect. We conceived a model that allowed us to derive a certain number of novel predictions about the perceptual processing of action consequences. We tested these predictions and provided evidence in favor of the preactivation hypothesis. We also provided evidence to show that the preactivation of the sensory pathway that code for a stimulus can represent a more general mechanism of the way that the brain works by assessing the commonalities between motor and repetition induced preactivation. Finally we assessed whether the preactivation hypothesis was compatible with recent account of the brain functioning.

# References

Albrecht, D. G., Farrar, S. B., & Hamilton, D. B. (1984). Spatial contrast adaptation characteristics of neurones recorded in the cat's visual cortex. Journal of Physiology, 347, 713-739.

Aliu, S. O., Houde, J. F., & Nagarajan, S. S. (2009). Motor-induced suppression of the auditory cortex. *Journal of cognitive neuroscience*, *21*(4), 791–802. doi:10.1162/jocn.2009.21055

Bach, M., & Ullrich, D. (1997). Contrast dependency of motion-onset and pattern-reversal VEPs: interaction of stimulus type, recording site and response component. *Vision Res, 37*(13), 1845-1849.

Band, G. P. H., van Steenbergen, H., Ridderinkhof, K. R., Falkenstein, M., & Hommel, B. (2009). Action-effect negativity: Irrelevant action effects are monitored like relevant feedback. *Biological Psychology*, *82*(3), 211–218.

Bäß, P., Jacobsen, T., & Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones : Evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology*, *70*(2), 137–143. doi:10.1016/j.ijpsycho.2008.06.005

Bässler, U. (1977). Sensory control of leg movement in stick insect Carausius morosus. *Biological Cybernetics*, 25, 61-72.

Bässler, U. (1983). Influence of femoral chordotonal organ afferences on ecdysis and on the development of motor programs in locust larvae. *Physiological Entomology*, 8, 353-357.

Bässler, U. (1987). Timing and shaping influences on the motor output for walking in stick insects. *Biological Cybernetics*, 55, 397-401.

Bays, P. M., Flanagan, J. R., & Wolpert, D. M. (2006). Attenuation of self-generated tactile sensations is predictive, not postdictive. (J. Lackner, Ed.)*PLoS biology*, *4*(2), e28. doi:10.1371/journal.pbio.0040028

Bays, P. M., & Wolpert, D.M. (2007). Predictive attenuation in the perception of touch.

In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition, attention and performance* (pp. 339–358). Oxford, England: Oxford University Press.

Bell, C. (1989). Sensory coding and corollary discharge effects in mormyrid electric fish. *J. Exp. Biol.*, *146*(1), 229–253.

Bell, C., & Grant, K. (1989). Corollary discharge inhibition and preservation of temporal information in a sensory nucleus of mormyrid electric fish. *J. Neurosci.*, *9*(3), 1029–1044.

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature neuroscience*, *1*(7), 635–40. doi:10.1038/2870

Blakemore, S.-J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-Temporal Prediction Modulates the Perception of Self-Produced Stimuli. *Journal of Cognitive Neuroscience*, *11*(5), 551–559. doi:10.1162/089892999563607

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433‑436. doi:10.1163/156856897X00357

Bratman, M. (1987). *Intention, Plans, and Practical Reason.* (Vol. 1999, p. 288). Cambridge, Mass.: Harvard University Press.

Bridgeman, B. (1995). A review of the role of efference copy in sensory and oculomotor control systems. *Annals of Biomedical Engineering*, *23*(4), 409–422. doi:10.1007/BF02584441

Buehner, M. J., & Humphreys, G. R. (2009). Causal binding of actions to their effects. *Psychological science*, 20(10), 1221–8. doi:10.1111/j.1467-9280.2009.02435.x

Burr, D. C., Morrone, M. C., & Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature, 371*(6497), 511-513.

Cardoso-Leite, P., Mamassian, P., Schütz-Bosbach, S., & Waszak, F. (2010). A new look at sensory attenuation: action-effect anticipation affects sensitivity, not response bias. *Psychological science : a journal of the American Psychological Society / APS*, *21*(12), 1740–5. doi:10.1177/0956797610389187

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and brain sciences*, *36*(3), 181–204. doi:10.1017/S0140525X12000477

Craver-Lemley, C., Reeves, A., Arterberry, M.E., 1997. Effects of imagery on vernier acuity under conditions of induced depth. Journal of Experimental Psychology: Human Perception and Performance 23, 3–13.

Craver-Lemley, C., Reeves, A., 1987. Visual imagery selectively reduces vernier acuity. *Perception* 16, 599-614.

Cravo, A. M., Claessens, P. M. E., & Baldo, M. V. C. (2011). The relation between action, predictability and temporal contiguity in temporal binding. *Acta Psychologica*, *136*(1), 157–166.

Dean, A. F. (1983). Adaptation-induced alteration of the relation between response amplitude and contrast in cat striate cortical neurones. *Vision Research*, *23*(3), 249-256.

Dehaene, S. (2003). The neural Basis of the Weber-Fechner law : A logarythmic mental number line. TRENDS *in Cog. Sci.*, 7(4), 145-147.

Deiber, M.-P., Honda, M., Ibanez, V., Sadato, N., & Hallett, M. (1999). Mesial Motor Areas in Self-Initiated Versus Externally Triggered Movements Examined With fMRI: Effect of Movement Type and Rate. *J Neurophysiol*, *81*(6), 3065–3077.

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS biology*, *5*(10), e260. doi:10.1371/journal.pbio.0050260

Desantis, A., Roussel, C., & Waszak, F. (2011). On the influence of causal beliefs on the feeling of agency. *Consciousness and Cognition*, *20*(4), 1211–1220. doi:http://dx.doi.org/10.1016/j.concog.2011.02.012

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences*, *93*(24), 13494–13499. doi:10.1073/pnas.93.24.13494

Desmurget, M., & Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in cognitive sciences*, *4*(11), 423–431.

Dobbins, I. G., Schnyer, D. M., Verfaellie, M., & Schacter, D. L. (2004). Cortical activity reductions during repetition priming can result from rapid response learning. *Nature*, *428*(6980), 316–9. doi:10.1038/nature02400

Domjan, M. (2003). *The Principles of Learning and Behavior, fifth edition, Belmont, CA: Thomson/Wadsworth*

Dum, R. P., & Strick, P. L. (2002). Motor areas in the frontal lobe of the primate. *Physiology & Behavior, 77*(4), 677–682.

Egner, T., Monti, J. M., Summerfield, C. (2010) Expectation and surprise determine neural population responses in the ventral visual stream, J. of Neurosci., 30(49), 16601-16608.

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance,* Vol 27(1), 229-240. doi: 10.1037/0096-1523.27.1.229

Engbert, K., & Wohlschläger, A. (2007). Intentions and expectations in temporal binding. *Consciousness and cognition*, *16*(2), 255–64. doi:10.1016/j.concog.2006.09.010

Engbert, K., Wohlschläger, A., Thomas, R., & Haggard, P. (2007). Agency, subjective time, and other minds. *Journal of experimental psychology. Human perception and performance*, *33*(6), 1261–8. doi:10.1037/0096-1523.33.6.1261

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598–601. doi:10.1038/33402

Fodor, J. A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology,* Cambridge, MA: MIT Press (p. 145).

Friese, U., Supp, G. G., Hipp, J. F., Engel, A. K., & Gruber, T. (2012). Oscillatory MEG gamma band activity dissociates perceptual and conceptual aspects of visual object processing: A combined repetition/conceptual priming study. *NeuroImage*, *59*(1), 861–871.

Friston, K. (2002). Functional integration and inference in the brain. *Progress in Neurobiology*, *68*(2), 113–143.

Friston, K. (2003). Learning and inference in the brain. *Neural Networks*, *16*(9), 1325–1352.

Friston, K. (2005). A theory of cortical responses. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *360*(1456), 815–36. doi:10.1098/rstb.2005.1622

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *364*(1521), 1211–21. doi:10.1098/rstb.2008.0300

Frith, C. D. , Blackemore, S. J., Wolpert, D. (2000) Explaining the symptoms of schyzophrenia: abnormalities in the awarness of action, Brain Res. Rev., 31(2-3), 357-363

Gallistel, C. R. (1986). The role of the dopaminergic projections in MFB self-stimulation. *Behavioural Brain Research*, *22*(2), 97–105.

Gentsch, A., & Schütz-Bosbach, S. (2011). I did it: unconscious expectation of sensory consequences modulates the experience of self-agency and its functional signature. *Journal of cognitive neuroscience*, *23*(12), 3817–28. doi:10.1162/jocn_a_00012

Gibson, J.J. (1977). *The Theory of Affordances* (pp. 67-82). In R. Shaw & J. Bransford (Eds.). Perceiving, Acting, and Knowing: Toward an Ecological Psychology. Hillsdale, NJ: Lawrence Erlbaum.

Gopfert, E., Muller, R., Breuer, D., & Greenlee, M. W. (1998). Similarities and dissimilarities between pattern VEPs and motion VEPs. *Doc Ophthalmol*, 97(1), 67-79.

Gorea, A., Sagi, D.(2001). Disentangling signal from noise in visual contrast discrimination. *Nat. Neuro. Sci.*, 4, 1146-1150.

Gratton, G., Coles, M. G. H., Sirevaag, E. J., Eriksen, C. W., Donchin, E., (1998) Pre- and poststimulus activation of response channels: A psychophysiological analysis. *Journal of Experimental Psychology: Human Perception and Performance*, Vol 14(3), 331-344. doi: 10.1037/0096-1523.14.3.331

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York,: Wiley.

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: with special reference to the ideo-motor mechanism. *Psychol Rev*, *77*(2), 73-99.

Grèzes, J., & Decety, J. (2002). Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia*, *40*(2), 212–222.

Grèzes, J., Tucker, M., Armony, J., Ellis, R., & Passingham, R. E. (2003). Objects automatically potentiate action: an fMRI study of implicit processing. *European Journal of Neuroscience*, *17*(12), 2735–2740. doi:10.1046/j.1460-9568.2003.02695.x

Grill-spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain : neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 17–19. doi:10.1016/j.tics.2005.11.006

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, *107*(1), 293–321.

Haggard, P., & Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research*, *126*(1), 128–133. doi:10.1007/s002210050722

Haggard, Patrick. (2008). Human volition: towards a neuroscience of will. *Nature reviews. Neuroscience*, *9*(12), 934–46. doi:10.1038/nrn2497

Haggard, Patrick, & Clark, S. (2003). Intentional action: Conscious experience and neural prediction. *Consciousness and Cognition*, *12*(4), 695–707.

Haggard, Patrick, Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature neuroscience*, *5*(4), 382–5. doi:10.1038/nn827

Haggard, Patrick, & Whitford, B. (2004). Supplementary motor area provides an efferent signal for sensory suppression. *Cognitive Brain Research*, *19*(1), 52–58.

Hamann, S. B., & Squire, L. R. (1997). Intact perceptual memory in the absence of conscious memory. Behavioral Neuroscience, Vol 111(4), 850-854. doi: 10.1037/0735-7044.111.4.850

Harless, E. (1861). Der Apparat des Willens. *Zeitschrift fuer Philosophie und philosophische Kritik*, (38), 50‑73.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Vision Neuroscience*, *9*(2), 181-197.

Helmholtz, H. (1860) *Handbuch der physiologischen optik* (ed. J. P. C. Southall), vol. 3. New York: Dover (English trans.)

Henson, R. ., Rylands, A., Ross, E., Vuilleumeir, P., & Rugg, M. (2004). The effect of repetition lag on electrophysiological and haemodynamic correlates of visual object priming. *NeuroImage*, *21*(4), 1674–1689.

Henson, R. N. ., & Rugg, M. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia*, *41*(3), 263–270.

Herbart, J. F. (1816). Lehrbuch zur psychologie (textbook for psychology). Königsberg, Germany: unzer

Herwig, A., Prinz, W., & Waszak, F. (2007). Two modes of sensorimotor integration in intention-based and stimulus-based actions. *Quarterly journal of experimental psychology (2006)*, *60*(11), 1540–54. doi:10.1080/17470210601119134

Hoffmann, J., Sebald, A., & Stöcker, C. (2001). Irrelevant response effects improve serial learning in serial reaction time tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol 27(2), Mar 2001, 470-482. doi: 10.1037/0278-7393.27.2.470

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): a framework for perception and action planning. *The Behavioral and brain sciences*, *24*(5), 849–78; discussion 878–937.

Hsu, Y.-F., & Waszak, F. (2012). Stimulus-classification traces are dominant in response learning. *International Journal of Psychophysiology*, *86*(3), 262–268.

Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: the role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological bulletin*, *139*(1), 133–51. doi:10.1037/a0028566

Hughes, G., Schütz-Bosbach, S., & Waszak, F. (2011). One action system or two? Evidence for common central preparatory mechanisms in voluntary and stimulus-driven actions. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *31*(46), 16692–9. doi:10.1523/JNEUROSCI.2256-11.2011

Hughes, G., & Waszak, F. (2011). ERP correlates of action effect prediction and visual sensory attenuation in voluntary action. *NeuroImage*, *56*(3), 1632–40. doi:10.1016/j.neuroimage.2011.02.057

Humphreys, G. R., & Buehner, M. J. (2010). Temporal binding of action and effect in interval reproduction. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, *203*(2), 465–70. doi:10.1007/s00221-010-2199-1

James, K. H., & Gauthier, I. (2006). Letter processing automatically recruits a sensory–motor brain network. *Neuropsychologia*, *44*(14), 2937–2949.

James, W. (1950). *The principles of psychology* : Vol. 2., New York, NY: Dover (original work published in 1890)

Jenkins, I. H. (2000). Self-initiated versus externally triggered movements: II. The effect of movement predictability on regional cerebral blood flow. *Brain*, *123*(6), 1216–1228. doi:10.1093/brain/123.6.1216

Kaliukhovich, D. A., & Vogels, R. (2011). Stimulus repetition probability does not affect repetition suppression in macaque inferior temporal cortex. *Cerebral cortex (New York, N.Y. : 1991)*, *21*(7), 1547–58. doi:10.1093/cercor/bhq207

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *J. Neurosci.*, *17*(11), 4302–4311.

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, *9*(6), 718–727.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual review of psychology*, *55*, 271–304. doi:10.1146/annurev.psych.55.090902.142005

Kesten, H. (1958). Accelerated stochastic approximation. *Annals of Mathematical Statistics*, 29(1), 41-59.

Kleiner M., Brainard D.,& Pelli D., 2007, What's new in Psychtoolbox-3?, Perception 36 ECVP Abstract Supplement.

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron*, 75(2), 265-70. doi: 10.1016/j.neuron.2012.04.034.

Kühn, S., Seurinck, R., Fias, W., & Waszak, F. (2010). The Internal Anticipation of Sensory Action Effects: When Action Induces FFA and PPA Activity. *Frontiers in human neuroscience*, *4*(June), 54. doi:10.3389/fnhum.2010.00054

Kunde, W. (2001). Response-effect compatibility in manual choice reaction tasks. *Journal of Experimental Psychology: Human Perception and Performance*, Vol 27(2), 387-394. doi: 10.1037/0096-1523.27.2.387

Kunde, W. (2003). Temporal response-effect compatibility. *Psychological research*, *67*(3), 153–9. doi:10.1007/s00426-002-0114-5

Lang, W., Cheyne, D., Kristeva, R., Beisteiner, R., Lindinger, G., & Deecke, L. (1991). Three-dimensional localization of SMA activity preceding voluntary movement. *Experimental Brain Research*, *87*(3). doi:10.1007/BF00227095

Leuthold, H., & Jentzsch, I. (2002). Distinguishing neural sources of movement preparation and execution. *Biological Psychology*, *60*(2), 173–198.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). *Brain*, *106*(3), 623–642. doi:10.1093/brain/106.3.623

Lotze, R. H. (1852). *Medicinische Psychologie oder die Physiologie der Seele* (Weidmann'sche Buchhandlung.). Leipzig.

Luce, R., D. (1959) On the possible psychophysical laws, *psychol. rev.*,66,81-95.

McCulloch, W. P., W. (1943). A Logical Calculus of Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics, 5*, 115-133.

Macmillan, N. A., Creelman, C. D. . (1991). *Detection theory: A user's guide.*: Lawrence Erlbaum Associates, Inc.

Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral cortex (New York, N.Y. : 1991)*, *15*(3), 299–302. doi:10.1093/cercor/bhh131

Mele, A. R. (1992) *Springs of action: Understanding intentional behavior.* New York, NY, US: Oxford University Press,(p.272)

Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein, J. F. (1993). Is the cerebellum a smith predictor? *Journal of motor behavior*, *25*(3), 203–16. doi:10.1080/00222895.1993.9942050

Miller, E. K., and R. Desimone. (1994) Parallel neuronal mechanisms for short-term memory. *Science* 263: 520–522

Moore, J., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and cognition*, *17*(1), 136–44. doi:10.1016/j.concog.2006.12.004

Moore, J. W., & Haggard, P. (2010). Intentional binding and higher order agency experience. *Consciousness and cognition*, *19*(1), 490–1. doi:10.1016/j.concog.2009.11.007

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: a review. *Consciousness and cognition*, *21*(1), 546–61. doi:10.1016/j.concog.2011.12.002

Moore, J. W., Wegner, D. M., & Haggard, P. (2009). Modulating the sense of agency with external cues. *Consciousness and cognition*, *18*(4), 1056–64. doi:10.1016/j.concog.2009.05.004

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, Vol 76(2), Mar 1969, 165-178. doi: 10.1037/h0027366

Moutsopoulou, K., & Waszak, F. (2012). Across-task priming revisited: Response and task conflicts disentangled using ex-Gaussian distribution analysis. *Journal of Experimental Psychology: Human Perception and Performance*, Vol 38(2), 367-374. doi: 10.1037/a0025858

Mumford, D. (1992) On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol. Cybern.* 66, 241-251.

Nieder A., Miller E.K. (2003)Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex.*Neuron* 37: 149-157.

Ohzawa, I., Sclar, G., & Freeman, R. D. (1985). Contrast gain control in the cat's visual system. *Journal of Neurophysiology, 54*(3), 651-667.

Pacherie, E. (2007). The Sense of Control and the Sense of Agency. *Psyche*, *13*(1), 1–30.

Pacherie, E. (2008). The phenomenology of action: a conceptual framework. *Cognition*, *107*(1), 179–217. doi:10.1016/j.cognition.2007.09.003

Pavlov, I. P. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex* (Vol. 1, p. 430). London: Oxford University Press.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437‑442. doi:10.1163/156856897X00366

Pfister, R., Kiesel, A., & Melcher, T. (2010). Adaptive control of ideomotor effect anticipations. *Acta psychologica*, *135*(3), 316–22. doi:10.1016/j.actpsy.2010.08.006

Picard, N., & Strick, P. L. (1996). Motor Areas of the Medial Wall: A Review of Their Location and Functional Activation. *Cerebral Cortex*, *6*(3), 342–353. doi:10.1093/cercor/6.3.342

Poulet, J. F. A., & Hedwig, B. (2007). New insights into corollary discharges mediated by identified neural pathways. *Trends in Neurosciences*, *30*(1), 14–21.

Pressnitzer, D., Suied, C., & Shamma, S. a. (2011). Auditory scene analysis: the sweet music of ambiguity. *Frontiers in human neuroscience*, *5*(December), 158. doi:10.3389/fnhum.2011.00158

Prinz, W. (1997). Perception and Action Planning. *European Journal of Cognitive Psychology*, *9*(2), 129–154. doi:10.1080/713752551

Rizzolati, G., Luppino, G., Matelli, M. (1998). The organisation of the cortical motor system : New concepts, Electro.Cli.Neuro., 106(4), 283-296

Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron, 26*(3), 703-714.

Roussel, C., Hughes, G., & Waszak, F. (2013). A preactivation account of sensory attenuation. Neuropsychologia, 51(5), 922-929. doi:10.1016/j.neuropsychologia.2013.02.005

Sato, A. (2008). Action observation modulates auditory perception of the consequence of others ' actions. *Consciousness and Cognition*, *17*(4), 1219–1227. doi:10.1016/j.concog.2008.01.003

SanMiguel, I., Widmann, A., Bendixen, A., Trujillo-Barreto, N., & Schröger, E. (2013). Hearing Silences: Human Auditory Processing Relies on Preactivation of Sound-Specific Brain Activity Patterns. The Journal of Neuroscience, 33(20), 8633–8639. doi:10.1523/JNEUROSCI.5821-12.2013

Saul, A. B., & Cynader, M. S. (1989). Adaptation in single units in visual cortex: the tuning of aftereffects in the spatial domain. *Vision Neuroscience, 2*(6), 593-607.

Sayres, R., & Grill-Spector, K. (2006). Object-selective cortex exhibits performance-independent repetition suppression. *Journal of neurophysiology*, *95*(2), 995–1007. doi:10.1152/jn.00500.2005

Schacter, D. L., & Slotnick, S. D. (2004). The Cognitive Neuroscience of Memory Distortion. *Neuron, 44*(1), 149–160.

Schadow, J., Lenz, D., Thaerig, S., Busch, N. A., Frund, I., Rieger, J. W., & Herrmann, C. S. (2007). Stimulus intensity affects early sensory processing: visual contrast modulates evoked gamma-band activity in human EEG. *Int J Psychophysiol*, *66*(1), 28-36. doi: 10.1016/j.ijpsycho.2007.05.010

Schafer, E. W. P., & Marcus, M. M. (1973). Self-Stimulation Alters Human Sensory Brain Responses. *Science*, *181*(4095), 175–177. doi:10.1126/science.181.4095.175

Sclar, G., Ohzawa, I., & Freeman, R. D. (1985). Contrast gain control in the kitten's visual system. *Journal of Neurophysiology, 54*(3), 668-675.

Searle, J. R. (1983). *Intentionality: An Essay in the Philosophy of Mind* , Cambridge University Press, (p. 278),

Sherrington, C. S. (1906). *The Integrative Action of the Nervous System,* Charles Scribner's Sons, New-York (p. 413).

Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential? *Clinical Neurophysiology, 117*(11), 2341–2356.

Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological bulletin*, *136*(6), 943–74. doi:10.1037/a0020541

Skinner, B. F. (1937). Two Types of Conditioned Reflex: A Reply to Konorski and Miller. *The Journal of General Psychology*, *16*(1), 272–279. doi:10.1080/00221309.1937.9917951

Stevens, J.C., Hall, J.W.(1966). Brightness and loudness as functions of stimulus duration, *Perception & Psychophysics*,Vol.1, 319-327

Stock, A., & Stock, C. (2004). A short history of ideo-motor action. *Psychological Research*. Vol. 68, Issue 2-3, pp 176-188 doi:10.1007/s00426-003-0154-5

Tenpenny, P. L. (1995). Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin & Review*, *2*(3), 339–363. doi:10.3758/BF03210972

Thorndike, E. L. (1898). *Review of Animal Intelligence: An Experimental Study of the Associative Processes in Animals.*

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. Journal of Experimental Psychology: Human Perception and Performance, Vol 24(3), 830-846. doi: 10.1037/0096-1523.24.3.830

Voss, M., Ingram, J. N., Haggard, P., & Wolpert, D. M. (2006). Sensorimotor attenuation by central motor command signals in the absence of movement. *Nature neuroscience*, *9*(1), 26–7. doi:10.1038/nn1592

Waszak, F., Cardoso-Leite, P., & Hughes, G. (2012). Action effect anticipation: neurophysiological basis and functional consequences. *Neuroscience and biobehavioral reviews*, *36*(2), 943–59. doi:10.1016/j.neubiorev.2011.11.004

Waszak, F., & Herwig, A. (2007). Effect anticipation modulates deviance processing in the brain, Brain. Res., vol.11(83), 74-82. doi:10.1016/j.brainres.2007.08.082

Weiskrantz, L., Elliott, J., & Darlington, C. (1971). Preliminary Observations on Tickling Oneself. *Nature*, *230*(5296), 598–599. doi:10.1038/230598a0

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics *Acta Psychologica*, 41, 67-85.

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology*, *8*(2), 227–233.

Wolpert, D M, & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural networks : the official journal of the International Neural Network Society*, *11*(7-8), 1317–29.

Wolpert, D., Ghahramani, Z., & Jordan, M. (1995). Are arm trajectories planned in kinematic or dynamic coordinates? An adaptation study. *Experimental Brain Research*, *103*(3). doi:10.1007/BF00241505

Wolpert, Daniel M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, *1*(6), 209–216.

Wolpert, Daniel M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience* , Vol.3, 1212 - 1217.

Wolpert, D. M., & Miall, R. C. (1996). Forward Models for Physiological Motor Control. *Neural Netw*, *9*(8), 1265-1279.

Yazawa, S., Ikeda, A., Kunieda, T., Ohara, S., Mima, T., Nagamine, T., Taki W, Kimura J, Hori T, Shibasaki H. (2000). Human presupplementary motor area is active before voluntary movement: subdural recording of Bereitschaftspotential from medial frontal cortex. *Experimental Brain Research*, *131*(2), 165–177. doi:10.1007/s002219900311

Ziessler, M. (1998) Response–effect learning as a major component of implicit serial learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol 24(4), 962-978. doi: 10.1037/0278-7393.24.4.962

Ziessler, M., Nattkemper, D., & Frensch, P. a. (2004). The role of anticipation and intention in the learning of effects of self-performed actions. *Psychological research*, *68*(2-3), 163–75. doi:10.1007/s00426-003-0153-6

# Table of Content

# APPENDIX 1:

Detailed resume of Desantis, A., **Roussel**, C., & Waszak, F. (2011). On the influence of causal beliefs on the feeling of agency. Consciousness and Cognition, 20(4), 1211–1220.

In this study we manipulated participants' prior belief of causality and assessed whether this manipulation modulated intentional binding. To be more precise, participants were made to believe that an auditory stimulus was either triggered by themselves or by somebody else, although, in reality, the sound was always triggered by the participants. We used the Libet-clock method to measure intentional binding (Haggard, Clark, et al., 2002; Libet et al., 1983). Contrary to other methods commonly used (e.g., interval estimations, see Engbert et al., 2008; Engbert, Wohlschläger, Thomas, & Haggard, 2007; Moore, Wegner, & Haggard, 2009), the Libet-clock method allows for the assessment of both the perception of the movement and of the sensory consequence. We predicted that binding effects would be larger when participants are convinced that their action triggered the tone, compared to when they believe that somebody else's action triggered the tone.

Notice that, as we outlined before, intentional binding has been considered to contribute to people's experience of authorship. A modulation of prior authorship belief on intentional binding would suggest that binding is also a consequence of our prior beliefs about the causal link between an action and a sensory change in the environment.

## a Methods

**Participants.** Thirty-six subjects (average age 24.1 years; sd = 4.38) participated in the experiment for a payment of € 10/h. All had normal or corrected-to-normal vision and were naïve as to the hypothesis under investigation.

**Material.** Stimulus presentation and data acquisition were conducted using Matlab 7.5.0 for Windows XP running on a PC computer connected to two 19-in. 120 Hz CRT monitors.

**Stimuli and procedure.** The experiment was run with a real and a bogus participant, hereafter referred to as the participant and the confederate. At the beginning, participant and confederate were told that the experiment consisted of three phases. Two baseline phases, one at the beginning and one at the end of the experiment, which they would carry out in two separate rooms, and a test phase (operant condition) between the two baselines which they would carry out conjointly in the same experimental room. In all trials (baselines and test) participants (and confederate) were presented with a clock-face marked with 5 'min' intervals and a clock-hand (2.5 cm of length and 0.1 cm of width) rotating with a period of 2560 ms (Haggard, Clark, et al., 2002; Libet et al., 1983). A green-shaded arc covered 120° of the clock surface (Figure 4). The position of the shaded arc was random (the arc could start at 90, 150, 180, 210, 270, 330 degrees of the clock-face). The initial clock-hand position was randomly chosen within the quarter of the clock which followed the end of the shaded arc.

The main task of the participants was to judge the onset-time of one of two events, a self-produced movement (right index finger key-press) or a tone (1000 Hz, 60 ms duration). In the baseline phases, movement/tone was produced/appeared alone. In the test phase, the participants' movements triggered tone presentation and the participants were required to judge onset-times of one or the other. To do so they used a computer keypad with their left hand to report the clock-hand position at which the given event occurred. Participants knew which event to report from the beginning of each trial (see below).

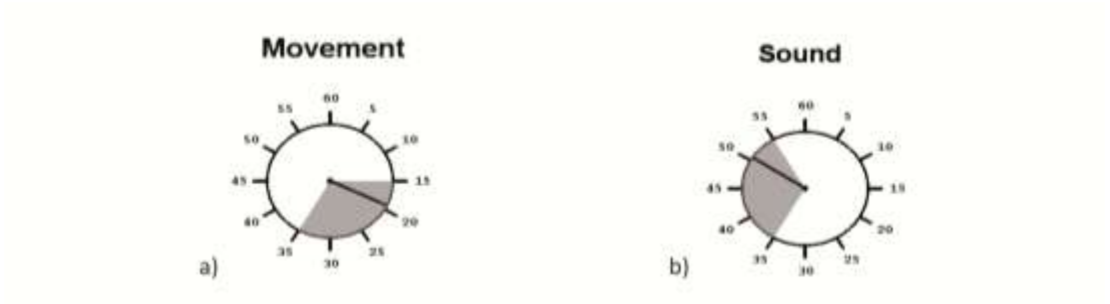

**Figure 4**. Example of the stimuli used during both baselines. In the movement condition (a) participants have to press a key within the green-shaded arc, in the sound condition (b) participants are asked to wait the occurrence of a sound.

**Baseline phases.** The baseline phases comprised two types of block. In the movement condition, participants had to make a key-press within the shaded arc and to judge the onset-time of their movement. Participants were instructed to avoid responding in a stereotyped way and to concentrate on the moment they actually pressed the key. In the sound condition, participants were asked to wait for the occurrence of a tone presented via a loudspeaker. Sound onset-times were individually yoked to the movement production times recorded in the previous block in the movement condition. To be precise, we used the recorded movement production times and added one of the three intervals (350, 550 and 750 ms), which were used during the test phase (see below). The clock-hand stopped at a random position 1.000–1.600 ms after the event and then disappeared. Thereafter participants reported the onset-time of the movement or the sound depending on the condition. They were encouraged to use the highest possible precision, and were not restricted to use the numbers marked on the clock-face. Each baseline phase was organized in four blocks; each block consisted of eight movement trials followed by eight tone trials for a total of 32 trials per event type.

**Test phase.** After the first baseline the confederate joined the participant to carry out the operant conditions. In this phase we used two monitors and a keyboard connected to the same PC. One monitor was in front of the participant, and the other in front of the confederate. The monitors and the shared keyboard were separated by a card board in order to prevent the participant from seeing what the confederate is doing. Before the test phase, participants ran two training sessions. In the first, both monitors showed the same clock as in the baseline phase (but without the shaded area). Participant and confederate were asked to carry out, alternately, a key-press. There were three conditions. In the subject condition the name of the participant was displayed on the screens (just below the clock). In this condition the participant's key-press triggered a tone; the confederate's key-press did not. In the confederate condition, the confederate's name was displayed on the screens. Here, the tone followed the confederate's key-press instead of the participant's key-press. Finally, in the ambiguous condition a sequence of rhombs (######) was displayed on the screens. In this condition the tone was triggered randomly by either the confederate's or the participant's key-press. The aim of this training session was to make the participant adopt three contextual beliefs: (a) if my name is displayed, the tone follows my action; (b) if the other's name is displayed, the tone follows his action; (c) if no name is displayed, the tone may follow my action or his action. This training ended when both subjects pressed their key nine times per condition. During the second training, the shaded arc was displayed. Both subjects had to press their key, alternately, within the shaded arc. As before the sound followed the key-press of the person whose name was displayed below the clock (or randomly one or the other in the ambiguous condition). The sound occurred randomly after one of three possible inter-onset intervals (350, 550, 750 ms). The aim of the second training was to reinforce the belief of the participants that, depending on the condition, either the confederate or they themselves trigger the tone. Moreover, it was meant to familiarize participants with the three different SOA used during the operant conditions of the test phase (see below). This training ended when both subjects pressed their key six times per condition.

Thereafter, participant and confederate carried out the three operant conditions (subject, confederate, ambiguous). In the subject condition, both screens displayed the participant's name (in order to make the participant believe that the tone is triggered by his/her key-press). In the confederate condition both screens displayed the confederate's name (in order to make the participant believe that the effect is caused by the confederate). In the ambiguous condition both screens displayed a sequence of rhombs '#####' (in order to make participant believe that s/he may or may not trigger the tone). We told the participant that, in the latter condition, the probability that the tone followed her/his action or the confederate's action was equally distributed. Importantly, contrary to the training sessions, during the test phase the tone was always triggered by the participant's key-press and never by the confederate's key-press, regardless of whether the participant's or the confederate's name (or a sequence of rhombs) was shown on the screen. In order to prevent participants from understanding that they always triggered the tone (on the basis of the temporal contiguity between their action and the sound), in all three operant conditions the participant and the confederate were asked to execute a key-press within the shaded arc (Figure 5). In particular, they were required to execute the movement when the clock-hand passed for the first time through the shaded area such that participant and confederate pressed the key at about the same time.

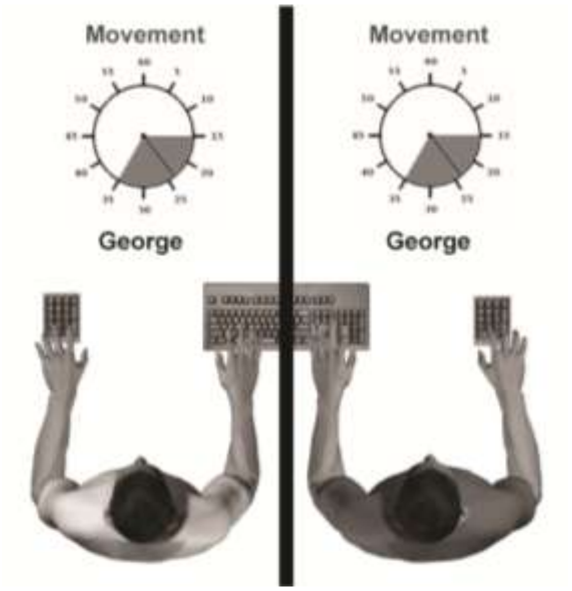

**Figure 5**. Illustration of the experimental set-up of the test phase. Our confederate sat on the right of the real participants. They were separated by a card board and shared the same keyboard to trigger the sounds. Both saw the same stimuli on the screens. The event they had to report was displayed above the clock. The name of the participant/confederate/a sequence of rhombs was shown blow the clock (see text for details). Both used a keypad to indicate the moment at which the event of interest appeared (as a matter of course, our confederate feigned to answer).

Since we used three different SOAs, participants were unable to tell that they actually triggered the sound in all conditions. Moreover, in order to strengthen the participants' causal beliefs in the three conditions, the confederate deactivated the participants' key once or twice per confederate and ambiguous block. In these trials the confederate triggered the tone himself, from time to time even during the second rotation of the clock-hand. As in the second training session, the tone was delivered randomly 350, 550 or 750 ms after the key-press. Notice that the probability that the tone is presented at the next possible SOA increases with increasing SOA (given that it has not been presented at the previous SOA(s)). At least this would be the case if the probability of tone presentation was the same across the SOAs. This phenomenon, known as the hazard function (Luce, 1991), entails foreperiod effects due to expectancy (Drazin, 1961; Keller & Van Der Schoot, 1978; Näätänen, 1970). To avoid these effects we kept the conditional probability of the occurrence of the tone after each of the three intervals constant. The corrected frequencies were such that the probability of the presentation of the tone at each of the three possible moments was 0.5. The formulas we used to calculate corrected frequencies were as follows (Näätänen, 1970):

(1) $n1 = {}^{1}/_{2} \cdot n$

(2) $n2 = {}^{1}/_{2} \cdot (n - n1)$

(3) $n3 = {}^{1}/_{2} \cdot (n - n1 - n2)$

(4) $NoTone = n - n1 - n2 - n3$

where $n$ is the number of trials for each belief condition, $n1$ is the number of trials in which the sound appears with a SOA of 350 ms, $n2$ is the number of trials in which the sound appears with a SOA of 550 ms and so forth. Keeping constant conditional probabilities necessitated including very few trials in which no tone was delivered (*NoTone*).

When the clock-hand disappeared participants were asked to report the clock-hand position at which the given event occurred. The confederate's screen displayed the participant's answers, which allowed him to monitor her/his performance. Moreover, participants were asked to answer the question ''Who triggered the tone?'' by pressing one of two keys ('ME'/'HIM'). During the subject and the confederate belief condition (where participants 'knew' who triggered the sound) participants were instructed to answer this question according to the name displayed on the screen, otherwise, they could not pass to the next trial. This permitted us to strengthen participants' association between name and causal role. However, in the ambiguous condition participants had to tell who was the origin of the sound.

The test phase consisted of 8 blocks of movement judgment and eight blocks of sound judgment (in which participants had to judge the movement and the tone, respectively.) The 16 blocks were presented alternately with block presentation being counterbalanced between subjects. Belief conditions (subject, confederate, ambiguous) were randomly distributed within each judgment type block. Likewise, SOAs were randomly distributed within each belief condition. Each block consisted of 24 trials (eight trials per belief condition). Before starting the test phase, participants ran one block of training per judgment type.

At the end of the experiment participants were asked whether they actually believed that: (i) they always triggered the sound when their name was displayed on the screen; (ii) the confederate triggered the sound when his name was displayed on the screen. They were also asked whether they believed that, in the ambiguous condition, the probability that the tone followed their action or the confederate's action was equally distributed.

## b Data analysis

The mean estimation error (the difference between estimated and actual onset of the given event) was calculated for each condition. Anticipatory estimates are represented as negative values. In order to calculate binding scores we subtracted each subject's mean estimation error in the baseline conditions from the mean estimation error for the same event in the operant conditions. For example, the mean estimation error of a voluntary action occurring alone (baseline) was subtracted from the mean estimation error of the voluntary action in a given operant condition (e.g., confederate condition, 350 ms). Likewise, the mean estimation error of the tone occurring alone was subtracted from mean estimation error of the tone triggered by a key-press in a given operant condition. The resulting perceptual shifts measure binding between actions and effects. We ran a separate analysis for the ambiguous condition, since, in this condition, participants had to carry out a double task. Here participants were not only asked to determine the time at which the event in question occurred but also to determine who triggered the tone. In the subject and confederate condition, by contrast, participants carried out a simple task. Here they only had to report the time of the event in question.

## c Results

Two participants were excluded from the analysis, because they realized that they triggered the tone in all conditions. The rest of the participants believed that they triggered the sound when their name was displayed on the screen and the confederate triggered the sound when his name was displayed. However, all of them had the strong impression of causing the tone during almost all the ambiguous trials. The trials in which no tone was delivered and values ±2 standard deviations from the mean were excluded from the analysis. A repeated-measures ANOVA, on sound estimates (defined as the difference between the actual and the perceived onset of the sound) in the baselines with SOA as factor, was conducted. As was to be expected, there was no significant effect of SOA. Thus, we collapsed across SOAs to calculate the sound baseline.

### *Subject vs. confederate belief condition*

**Sound estimations.** A repeated-measures ANOVA on sound binding scores (the difference between the estimation error for the sound in the baseline condition and the estimation error for the sound in the operant conditions) with belief (subject, confederate) and SOA (350, 550, 750 ms) as factors was conducted. The main effect of belief was highly significant F(1, 33) = 9.2989, p = .0045. As predicted, we observed stronger anticipation in the subject belief condition than in the confederate belief condition (Figure 6).

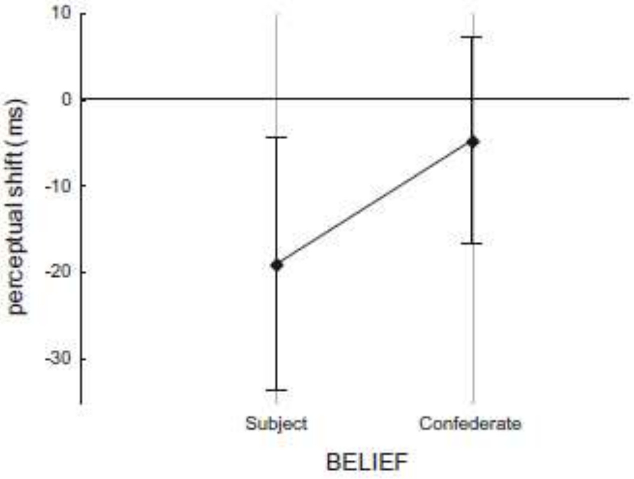

**Figure 6**. Mean sound binding scores in ms for the subject and confederate belief conditions.

The interaction between Belief and SOA was not significant F(2, 66) = .56853, p = .56911. To examine absolute binding effects for the sound, paired t-tests comparing the baseline with the subject belief condition and the confederate belief condition, respectively, were carried out. One-tailed t-Tests showed that the sound was significantly anticipated in the subject belief condition t(33) = 2,01, p = .0026, but not in the confederate belief condition t(33) = 0,18, p = .426. All in all, the pattern of results, thus, corroborates that prior causal belief influences the temporal perception of sensory action consequences, with the subject belief condition resulting in stronger binding than the confederate belief condition.

**Movement estimations.** A repeated-measures ANOVA on movement binding scores (the difference between the estimation error for the movement in the baseline condition and the estimation error for the movement in the operant conditions) with Belief (subject, confederate) and SOA (350, 550, 750 ms) as factors was conducted. Neither the main effect of Belief F(1, 33) = .83128, p = .3685; (Figure 7), nor the interaction between Belief and SOA was significant F(2, 66) = .13939, p = .87015. To assess absolute binding scores, paired t-tests comparing the baseline with the subject belief and the confederate belief condition (collapsed across SOAs) were carried out. We found that the movement was significantly anticipated in both the subject t(33) = 2.62 p = .006 and the confederate belief condition t(33) = 2.07, p = .026.

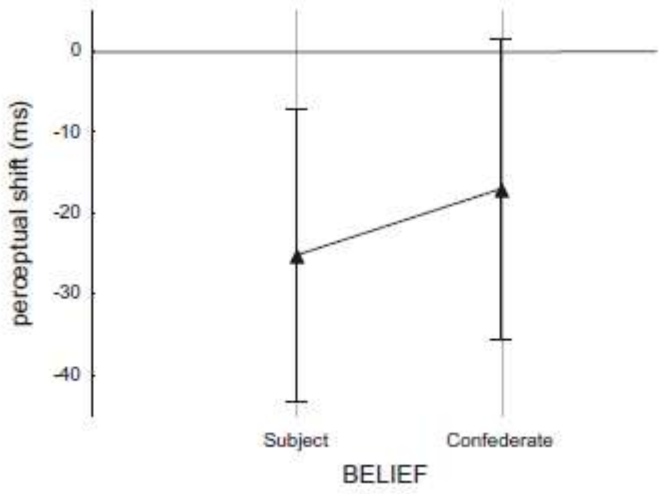

**Figure 7**. Mean movement binding scores in ms for the subject and confederate belief conditions.

### ***Ambiguous condition***

All the participants told us that they had the strong impression of causing the tone during almost all the ambiguous trials. This made the ambiguous condition more similar to subject belief condition than to the confederate condition. Nonetheless, participants tried to equally distribute their causality responses on ‘me’ and ‘him’ answers. Accordingly, analyzing this condition separately for the two answers is impossible. So, we calculate movement estimations and sound estimations irrespective of the distinction ‘me’ and ‘him’. We carried out paired t-Tests comparing the baseline estimations with the movement and the sound estimations, respectively, of the ambiguous condition (collapsed across SOAs). The t-tests showed that the sound t(33) = 1.83, p = .0037 and the movement t(33) = 4.93, p = .0000 were significantly anticipated in the ambiguous condition compared to the baseline (Figure 8).

Hence, the pattern of results of the ambiguous condition strongly resembles the results obtained in the subject belief condition. The finding that, in the ambiguous condition, participants believed that they were the origin of the sound is in agreement with previous studies which showed an egocentric bias in attribution of action in ambiguous context (Daprati et al., 1997; Franck et al., 2001; van den Bos & Jeannerod, 2002).

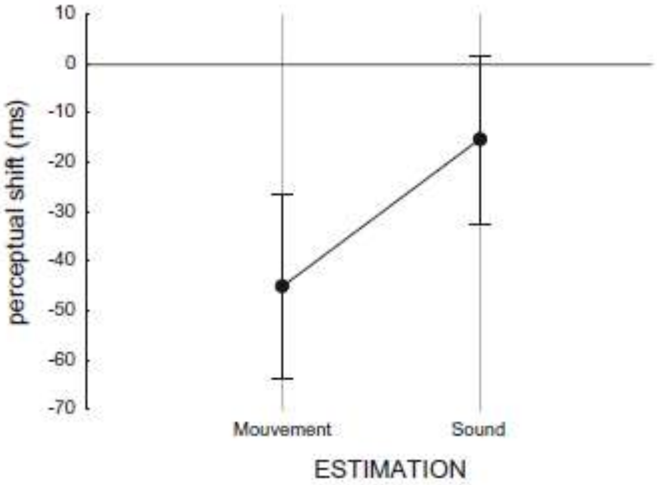

**Figure 8**. Mean binding scores in ms for both movement estimation and sound estimation in the ambiguous condition.

## d Discussion

The present study tested the hypothesis that intentional binding is modulated by prior causal beliefs. Our study provides support for this hypothesis. We found that binding effects for a tone were significantly larger when participants were convinced that their actions triggered the tone, compared to when they were convinced that the confederate's action triggered the tone. Notice that both belief conditions matched in terms of temporal expectancy and voluntary action control, in fact the tone was always triggered by the participants' actions.

Our findings are in general agreement with previous data which showed that both causality and voluntary action are necessary for intentional binding to emerge (Buehner & Humphreys, 2009; Cravo et al., 2009). For instance, in a recent study Cravo et al., (2009) used a variation of the launching effect proposed by Michotte, (1963). In some blocks, participants controlled the launch stimulus; in other blocks the launch stimulus was externally controlled. In addition, in two conditions (collision vs. non-collision), they varied the causal connection between the stimuli they used in their launching setting in order to induce the participants to get a high or low impression of causality. Temporal binding was found in the presence of both voluntary action and high causality only. One possible explanation for our results is that participants allocated more attention to the sensory consequence when they believed that they were its origin. However, this simple explanation seems to us to be unlikely since we did not find participants' estimates to show more variability in the confederate belief condition compared to the subject belief condition. Instead, we suggest that to understand how prior causal beliefs influence intentional binding we need to distinguish two signals that the brain can use to determine the onset-time of a sensory action effect. First, it can use internal prediction provided by the forward model (Wolpert, 1997). Second, it can use the ‘external’ signal, i.e., the actual sensory feedback. In the case of passively presented stimuli, the brain has to rely on the ‘external’ signal only. In the case of stimuli triggered by a voluntarily movement, on the other side, it can use both internal prediction and external signal, thus yielding the perceived shift in time. According to the comparator model of motor control, internal prediction is used to provide internal feedback of the predicted outcome of an action which can be used before sensory feedback is available, thereby shifting earlier in time the perceived occurrence of action effects (Wolpert, 1997).

We suggest that prior causal belief influences how reliable the brain considers the information provided by the internal forward model to be. When people believe that they are the author of an upcoming sensory event, predictive signals are considered to be reliable, if they believe that someone else is the origin of the upcoming stimulus event, predictive signals are considered to be unreliable. If, as Moore and colleagues' (2008; 2009a; 2009b) suggest, the sense of agency depends on a Bayesian cue integration process that is based on a weighted combination of different cues, then our results can be explained by assuming that in the subject and confederate belief conditions the internally anticipated effect get high and low weighs, respectively. Accordingly, the results of the ambiguous condition could be taken to show that under uncertainty, participants use sensorimotor information to guide their experience of action. It is only if external cues to agency are sufficiently compelling that sensorimotor information is overridden. Notice that, in the confederate condition, there was no significant shift in perceived stimulus onset-times at all (compared to the baseline). This might be taken to indicate that internal prediction did not play any role in this condition. Thus, a more radical explanation of our findings would be that prior causal belief influences whether or not the action’s consequence is predicted in the first place. If the system believes not to trigger the upcoming stimulus why should it bother to try to predict the event?

Our results can also be explained in terms of the ‘pre-activation model’ of sensory attenuation of self-generated sensory consequences that has recently been suggested by Cardoso-Leite et al. (2010). Cardoso-Leite and colleagues suggest that the preparation or execution of a voluntary action results in pre-activity of the sensory network that represent the action’s expected perceptual consequence, thus increasing the mean level of activity in this network to some pedestal level. This assumption is corroborated by an fMRI study from Kühn, Seurinck, Fias, & Waszak (2010) who found an increase in activity in the parahippocampal place area (PPA) and fusiform face area (FFA) for actions that in a previous training phase have triggered houses and faces, respectively. PPA and FFA are considered to represented houses and face perceptually (Epstein & Kanwisher, 1998; Kanwisher, McDermott, & Chun, 1997). Cardoso-Leite and colleagues suggest sensory attenuation to be due to the fact that discrimination of the signal activation (of the real stimulus) from the pedestal level is more difficult than discrimination of the signal activation from the baseline level (i.e., in a condition without effect anticipation, in which baseline activity is not raised to pedestal levels). The same model can also explain a shift of the perceived onset-time of anticipated sensory action consequences: an action effect might reach the threshold of awareness faster, if the activity in its neural representation starts from the pedestal level than when it starts from the baseline level. Thus, signal accumulation for passively presented stimuli takes longer to reach the threshold of awareness than signal accumulation for action effects, since the perceptual representation of the latter has already been pre-activated by internal motor signals.

To explain our results by means of the pre-activation account, we just need to assume, very similar to what we suggested above, that the effect is only predicted, or more efficiently, if the agent believes that s/he is the origin of the upcoming sensory event. This would bring about that accumulation of sensory evidence starts from a higher level in the subject than in the confederate belief condition, resulting in a shift of the perceived onset in the former condition compared to the latter.

Contrary to the perception of the tone, we did not find any influence of causal belief on the temporal perception of the movement. This finding suggests that binding effects of the movement and of the sensory consequence that have been reported in previous studies (Haggard, Clark, et al., 2002; Tsakiris & Haggard, 2003; Wenke, Waszak, & Haggard, 2009) may be based on different mechanisms. This notion is in agreement with a recent study by Moore et al., (2010). These authors demonstrated that a disruption of pre-SMA function by transcranial magnetic stimulation (TMS) affected the perception of the time of the sensory effect but not of the movement. Importantly, some studies point out that pre-SMA has a pivotal role in the prediction of sensory consequences. Ikeda et al. (1999), for example, emphasize its role in both action preparation and anticipatory processing of warning signals prior to action. Likewise, disruption of pre-SMA activity has been found to abolish sensory suppression during voluntary movement (Patrick Haggard & Whitford, 2004), a phenomenon that has been linked to predictive mechanisms (Blakemore et al., 2002). Taken together, one might, thus, argue that the binding effect of the sensory consequence (but not of the movement) depends on internal predictive information. This would be in agreement with our hypothesis that prior causal belief influences the way in which internal predictive signals are processed.

We would like to point out two caveats though. First, Moore & Haggard (2008) showed that both predictive and postdictive mechanisms contribute to the temporal perception of one's movement. Thus, we cannot exclude the possibility that postdictive mechanisms are also involved on the temporal perception of the sound and that it is these postdictive processes that prior causal belief influences.

Second, we did not find the usual binding effect for the movement estimation. This unexpected finding cannot be easily explained by existing accounts of intentional binding. We believe that it is due to the fact that we assessed intentional binding in a social setting. This supposition is corroborated by a recent study of Strother, House, & Obhi (2010) who observed the movement effect to be absent in a social setting. They suggest that their results are due to paradigm-specific factors resulting from the fact that participants performed the task in duos. The social setting, thus, seems to influence movement estimations in particular. According to the paradigm of social facilitation subjects' performances may be affected in various ways by the presence of a co-actor (Bond & Titus, 1983; Zajonc, 1965).

Furthermore motor responses seem to be specially affected by social settings (Zajonc, 1965). In our study participants freely decided when to execute their movement which makes their own movement highly attended. Several studies point out that when subjects are asked to estimate the time at which they initiated a finger movement, the estimated time of awareness consistently anticipated the actual starting time of the movement by 50–80 ms (Libet et al., 1983; McCloskey, Colebatch, Potter, & Burke, 1983; Obhi, Planetta, & Scantlebury, 2009). This suggests that participants tend to report the time at which they intended to move and not the actual movement itself. We speculate that, in the presence of a co-actor, participants attend more on the moment at which they intended to execute the movement, thus reporting even earlier movement onset-times. More research is needed to clarify these issues.

In conclusion, it has been assumed that an action producing intentional binding needs to be voluntary and that the effect needs to be predictable. However, our results show that these two factors are not sufficient for the emergence of intentional binding: causal belief seems to be another essential precondition for the effect to emerge. At the same time, our results seem

to suggest that high-level contextual information influences sensorimotor processes responsible for generating intentional binding.

# APPENDIX 2:

Supplementarry result for Study 2

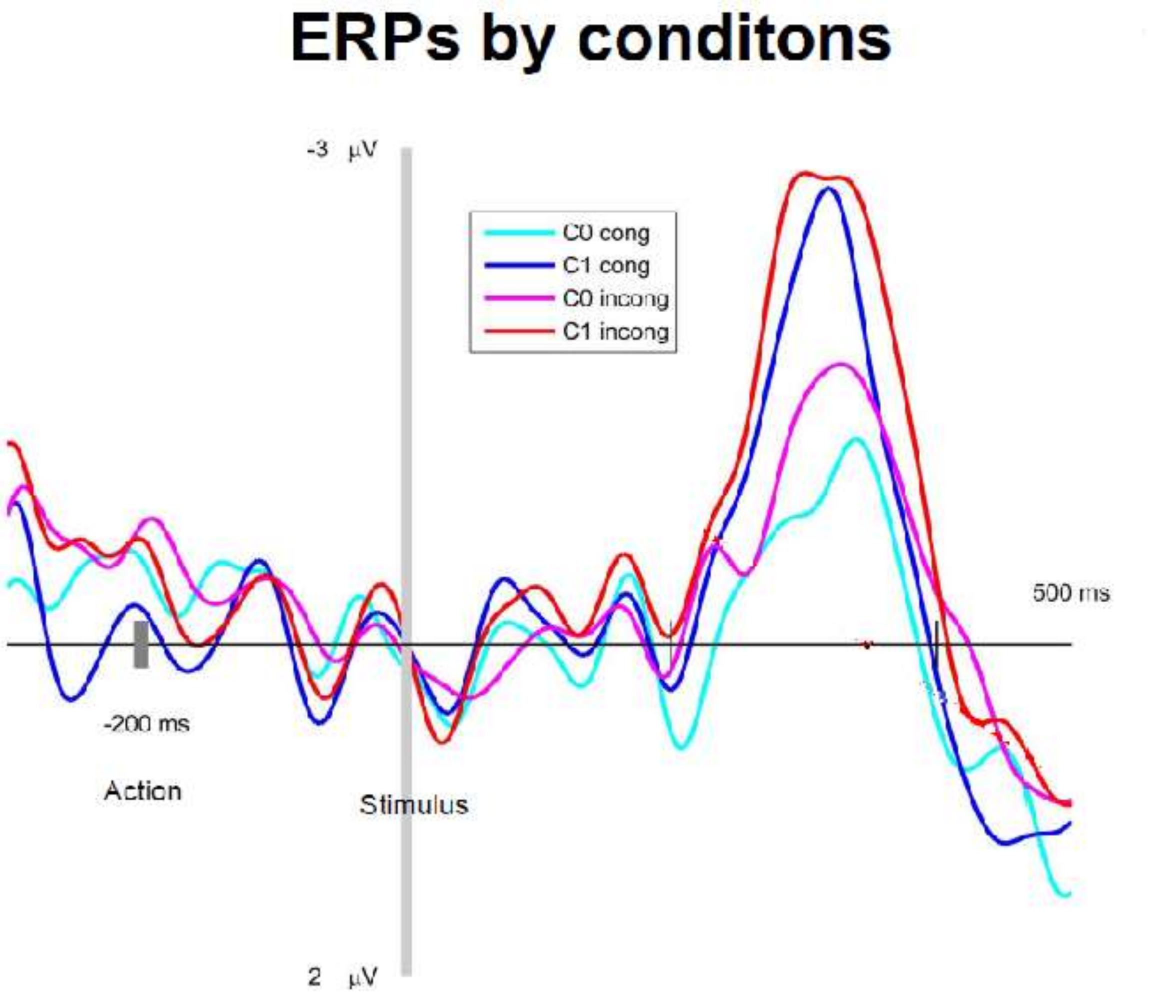

**Figure 1 Appendix 2:** On this figure we present the pattern of the interaction between Contrast (C0 VS C1) and Congruency (Congruent VS Incongruent). No significant interaction was found between these two factor.

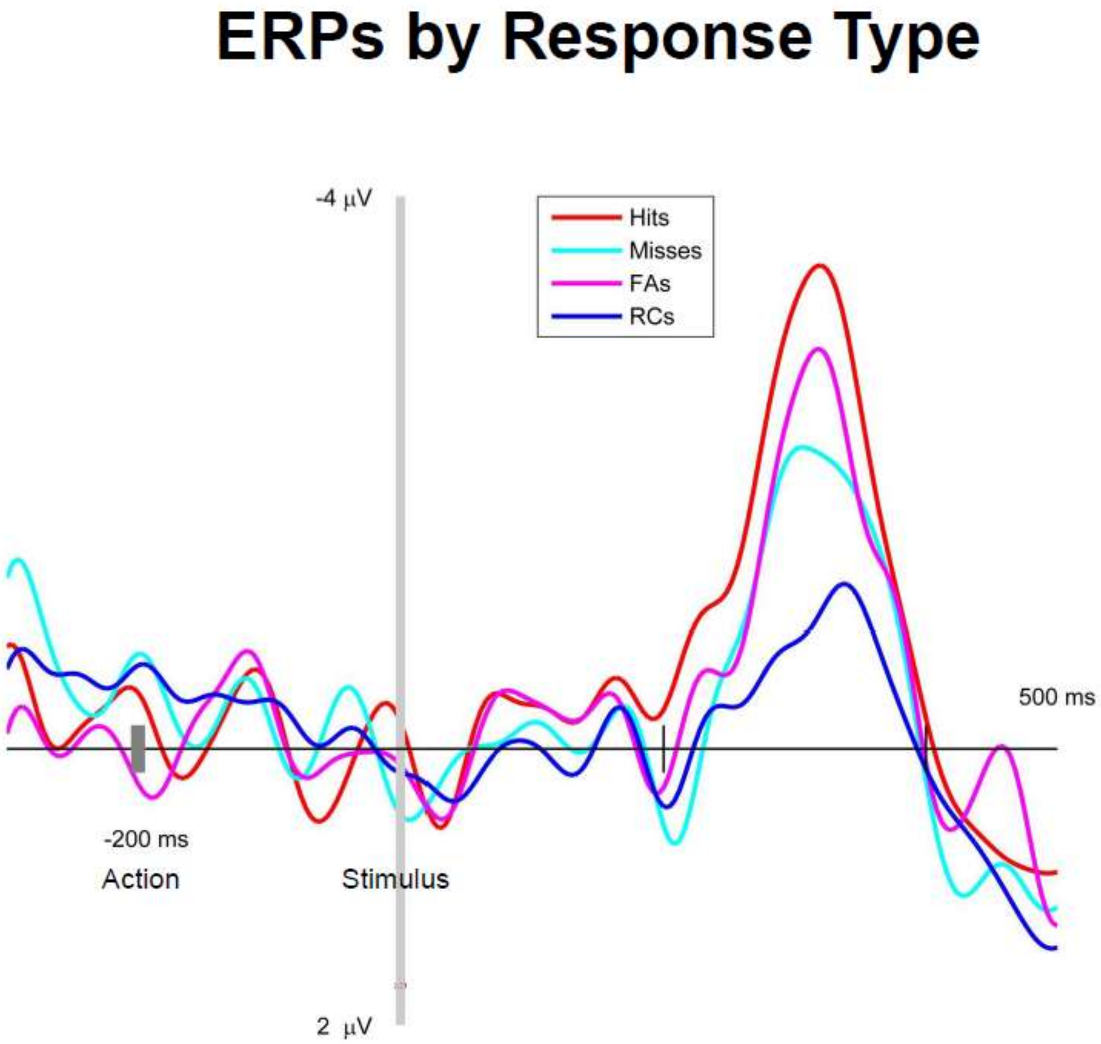

**Figure 2 Appendix 2:** On this figure we present the interaction between the Contrast (Hit and misses (C1) vs RC and FA (C0)) and the accuracy (Hit and RC vs Fa and misses). No significant interaction was found. The pattern seems to indicate that the isolated component contributes to the perception of intensity and its judgement.

Taken together this two results suggest that more investigation of the effect of action effect prediction on the stimulus perception function of the stimulus intensity could lead to a better understanding of the mechanisms underlying the prediction of the action’s consequences.

# APPENDIX 3

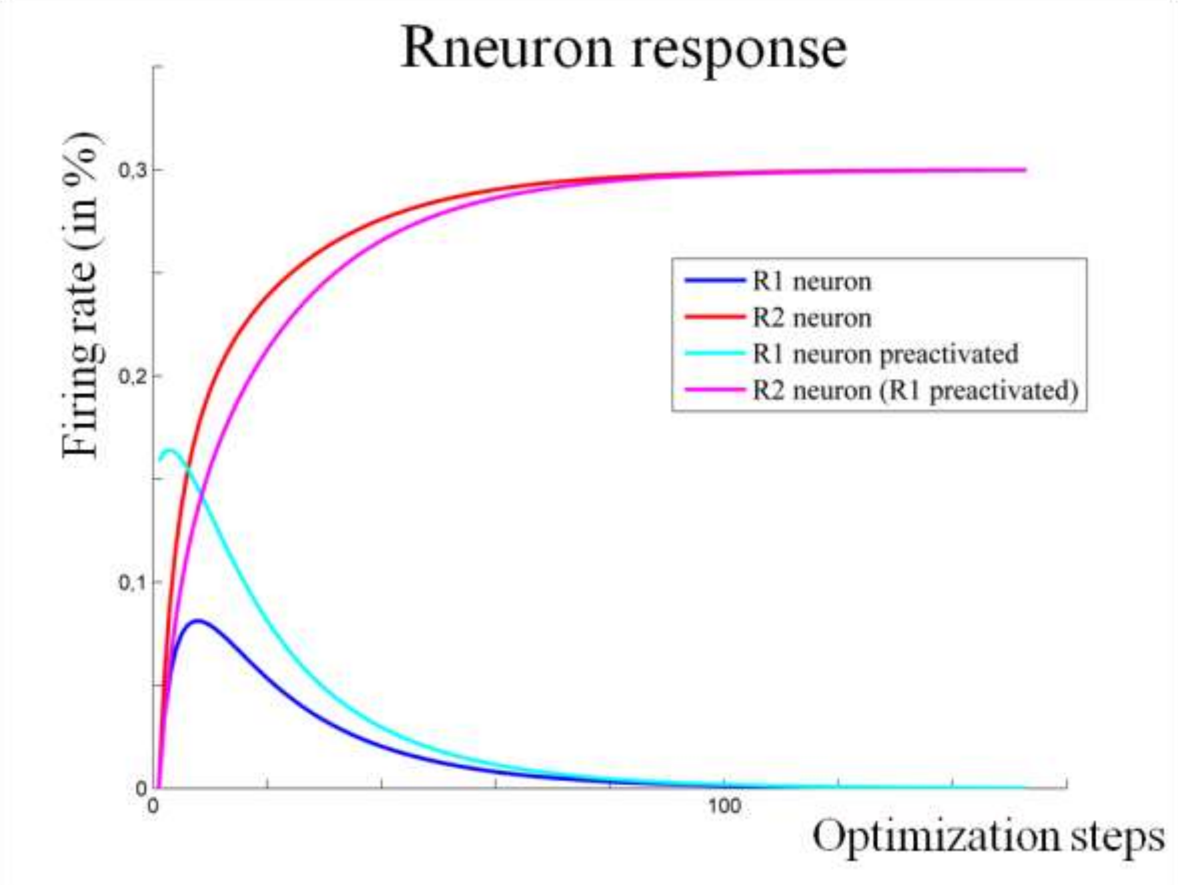

Figure 1 Appendix 3: On this figure are presented the responses of R1 and R2 neurons to a left tilted stimulus when R1 is preactivated or not. This is a match for the condition incongruent (Left tilted gabord is presented while a right tilted gabord is predicted) vs the neutral condition (leftward (or rightward ) tilted grating is presented while no prediction is made).

This suggests that the identification of an incongruent stimulus is worse than a neutral stimulus. Thus the preactivation hypothesis predicts an increase in identification performance for congruent trial and a decrease in identification performance for an incongruent trial both compared to the identification performance in a neutral condition (e.g when no predictions are made).