Task-Related Reinforcement Signals and Visual Plasticity

UNIVERSITÀ DEGLI STUDI
DI TRENTO
CIMeC - Center for Mind/Brain Sciences

CIMeC
Center for Mind/Brain Sciences

David Pascucci

Ph.D. Thesis

2014

Doctoral School in Cognitive and Brain Sciences

26^{th} Cycle

To my brothers

Contents

List of Figures

Tables

Chapter 1

> “**Autogenic reinforcement**: *the self-induced strengthening of a reaction so that the response to a particular stimulus is more prompt or more intense than would otherwise be exhibited by the same person to the same stimulus.”*
> *(Corsini, 1999)*

Introduction

1.1 Task-related signals of reinforcement

For complex living systems, such as human beings, the ability to identify stimuli with behavioural relevance is fundamental for survival. In many situations for example, events that signal the imminence of a danger or conversely the proximity of a reward may be crucial for survival and need to be rapidly and correctly identified. However, the environment where we live provides the organism an overwhelming and constant flow of information, and to successfully detect important stimuli, the brain must have dedicated and efficient cognitive selection mechanisms.

In experimental psychology, the ability of an individual to detect, discriminate or recognize relevant events is one of the main investigated issues, with particular focus on the mechanisms by which detection can occur in crowded scenes, where other irrelevant stimuli are exposed. While intuition may suggest that the more we are engaged in detecting specific items the less we are sensitive to other irrelevant information, a recent line of evidence has challenged this notion describing an intriguing aspect of our perceptual experience. The striking finding is that when humans are engaged in target detection or identification, they also show a temporary and non-specific increase in perceptual sensitivity to concurrent information (Seitz & Watanabe, 2009; Swallow & Jiang, 2010). In other terms, this implies that being able to recognize task-relevant stimuli not only facilitates task accomplishment, but also enhances memory and learning for paired contextual and irrelevant information.

This discovery had a groundbreaking impact on several theories of learning (Watanabe, Náñez & Sasaki, 2001; Seitz & Watanabe, 2005) and attention (Swallow & Jiang, 2013; Roelfsema, van Ooyen & Watanabe, 2010) revealing a particular property of the human brain by which information presented during important perceptual episodes attains “privileged” representations. Although the mechanisms underlying this form of learning are still unclear, it has been hypothesized that alerting or reinforcement brain signals may underlie such transitory increase in receptivity. From a scientific perspective, this phenomenon may represent a fundamental adaptive brain function that deserves further investigation.

In the present work, I will explore the impact of recognizing visual targets on short- and long-term forms of cortical plasticity, with the aim to further characterize this effect and unravel the main factors involved. In order to provide a clear picture of the theoretical and experimental background of this work, in the next sections of the present chapter I will review the literature highlighting the perceptual and memory effects of targets recognition, introducing the phenomenon of task-irrelevant perceptual learning (TIPL) and the target-related memory enhancement (TRME). I will then focus on the theoretical interpretations proposed to explain such effects, with a closer analysis of the aspects that could link targets recognition to mechanisms of endogenous reward and attentional alerting. Finally, I will describe how brain signals triggered by task-relevant items have been integrated as essential components in recent models of perceptual learning and attention.

1.2 The effect of targets recognition

The effect of target recognition has been generally observed in behavioural tasks where participants have to identify one or more targets while being exposed to other stimuli that are not relevant for the task at hand. In order to measure the influence of target detection on the processing of irrelevant information, participants are then examined in their ability to perceive, discriminate or remember such task-irrelevant information. With this procedure, it is possible to test whether information that has been paired with targets or distractors gains preferential access to our memory or sensory representation. Crucially, the majority of research reported a consistent advantage for information paired exclusively with relevant targets.

In the last decade, the research focused on this topic has mainly adopted two methods and related measures. One approach investigated long-term changes in the perceptual representation of stimuli paired with visual targets (Watanabe et al., 2001; Seitz & Watanabe, 2003, 2005; Seitz, Lefebre, Watanabe & Jolicoeur, 2005; Tsushima, Seitz & Watanabe, 2008), the other measured short- and long-term memory for images presented at the time of task-relevant events (Lin, Pype, Murray & Boynton, 2010; Swallow & Jiang, 2010; Makovski, Swallow & Jiang, 2011; Leclercq & Seitz, 2012a, 2012b, 2012c). Despite the differences between the two approaches (see later in this chapter), the experimental evidence provided converges in highlighting some important points: The detection of targets enhances the representation and memorization of irrelevant stimuli only when the two events coincide in time (Seitz & Watanabe, 2003; Swallow & Jiang, 2011) and exclusively if targets are correctly identified (Seitz et al., 2005; Swallow & Jiang, 2010).

These constrains suggest a causal link between the successful detection of an item and the ability to remember or learn temporarily paired task-unrelated information, suggesting the action of target-elicited, specific brain signals on our perceptual experience (Pascucci & Turatto, 2013; Seitz & Watanabe, 2005; Roelfsema et al., 2010; Swallow & Jiang, 2013). Interestingly, these putative signals may underlie both the long-term increase in sensitivity for irrelevant stimuli reported in TIPL and the enhanced memory for target-paired images observed in TRME.

1.3 Task-Irrelevant Perceptual Learning (TIPL)

The phenomenon of TIPL represents a new and compelling finding in the field of visual perceptual learning. While the majority of studies on this topic has focused on the extent to which attention and extensive practice on a task can lead to long-term refinement of perceptual skills (Ahissar & Hochstein, 1993; Gilbert, Sigman & Crist, 2001), TIPL revealed an intriguing form of learning without awareness, training or attention to the irrelevant information.

The notion of TIPL was introduced by Watanabe, Náñez and Sasaki (2001) in their pioneering work on perceptual learning by mere exposure. In this study, human participants were asked to perform a rapid serial visual presentation (RSVP) task for several days of training. The RSVP was displayed at fovea inside a grey disk and the task was to report the identity of two light-grey letters (targets) embedded on a stream of eight black letters (distractors). During the RSVP stream, participants were exposed to irrelevant dynamic random dot displays (DRD) that contained a constant weak signal of motion coherence (5% of the dots moving to the same direction). Crucially, participants' ability to detect and discriminate the exposed coherent motion increased significantly after training, leading the authors to conclude that the frequent presentation of the same visual feature (e.g. motion direction) can sensitize our visual system (Watanabe et al., 2001). Although this interpretation has been revised (see later in this chapter), the assumption that learning was induced by mere exposure was supported by the fact that during training participants were fully engaged on the RSVP task, and the surrounding coherent motion was too weak to be noticed and attended.

Whereas in Watanabe, Náñez and Sasaki (2001) the RSVP served mainly as a strategy to engage participants' attention, a later work by Seitz and Watanabe (2003) addressed whether the targets and distractors of the RSVP could have a differential role in TIPL. As in the previous study, the paradigm consisted of extensive training on a central RSVP combined with the exposure to background DRDs. However, in this study, different directions of motion were consistently paired with the occurrence of targets and distractors. This manipulation revealed that TIPL is not a form of pure exposure-based learning, but rather it depends strictly on the presence of specific items of the main task, and occurs only when irrelevant motion is presented at the time of appearance of the targets.

On the ground of this evidence, Seitz and Watanabe (2005) proposed a revised interpretation of TIPL according to which task-irrelevant information is learned as a consequence of the release of an internal "learning" signals (Seitz & Watanabe, 2003, 2005; Roelfsema et al., 2010; Huang & Watanabe, 2012) or "endogenous rewards" (Pascucci & Turatto, 2013; Seitz, Náñez, Holloway, Koyama & Watanabe, 2005; Seitz et al., 2005), triggered by the execution of the main task. The "endogenous reward" hypothesis, in addition, has been corroborated by a subsequent work of Seitz, Lefebvre, Watanabe and Jolicoeur (2005) where the authors showed how TIPL requires high-level processing and conscious identification of the targets to be observed, supporting the crucial role for task accomplishment in this form of learning.

In the series of studies reviewed so far, a key aspect is that irrelevant stimuli are regularly presented below the sensory threshold. This is mainly for two reasons: (1) Visible stimuli systematically paired with task-relevant elements may facilitate association between these items, making the irrelevant stimuli somehow relevant for the main task; (2) Supraliminal stimuli may capture attention (Yantis, 1993; Jonides & Yantis, 1988) and be voluntarily or involuntarily attended during training.

As for the first point, however, it may be hypothesized that even subliminal stimuli can be implicitly associated with task elements, and thus, used to predict useful features of task-related events. To rule out this possibility, Seitz and Watanabe (2008) manipulated the correlation between irrelevant motion transients and the RSVP targets in a TIPL paradigm. In this study, the authors presented participants with a central RSVP where the stream could contain a fixed pair of targets (e.g., always the same couple of number), a randomly determined pair of targets (as in Seitz & Watanabe, 2003) or only distractors, and associated each one of these conditions with different directions of subliminal motion coherence. Under the hypothesis that TIPL occurs when subliminal stimuli can be used to predict task-relevant events or responses, the prediction was that directions of motion anticipating the exact identity of the targets (and the related motor response) are more likely to be learned. Nonetheless, the authors found TIPL only for conditions of response uncertainty (where targets were randomly selected) and no beneficial or “predictive” effect of the subliminal stimuli on task performance. This result supported the interpretation that task-irrelevant learning arises from internal signals of reinforcement (or endogenous reward) and not as a consequence of implicit associations between relevant and irrelevant stimuli (Seitz & Watanabe, 2008).

A second important reason why TIPL requires subliminal stimuli is that, if the irrelevant items paired with the RSVP are supraliminal, they may capture attention and consequently they may be actively and repeatedly attended during training, limiting the interpretation of TIPL in terms of “perceptual learning without attention” (Setiz & Watanabe, 2002, 2005). As for this point, the effect of presenting supraliminal irrelevant stimuli, as well as the impact of directing attention to them, have been addressed in two separated studies. In the first one (Tsushima, Seitz & Watanabe, 2008), the intensity of the irrelevant motion was varied methodically from subliminal and para-threshold to supraliminal. The results clearly demonstrated that only subliminal and para-threshold stimuli are learned, whereas supraliminal information is not. This evidence revealed that attentional mechanisms may play a crucial role in TIPL by promoting the processing of relevant features and suppressing or inhibiting the representation of irrelevant and disturbing (supraliminal) information.

On the basis of this view, a later work by Choi, Seitz and Watanabe (2009) directly addressed how covert shifts of attention toward the irrelevant stimuli affect TIPL. In this study, five arrows replaced the central RSVP and participants were asked to report their pointing direction. Concomitant with the target arrows, two separated DRDs containing different directions of subliminal motion were exposed on both sides of the screen. Crucially, the arrows, in addition to represent the targets of the relevant task, served as exogenous orienting cues, inducing systematic shifts of attention to one of the two DRDs. TIPL was found only for the motion signal not cued by the arrows, and, even more interesting, there was a trend for a negative improvement (if compared to a control condition with non-exposed motion signals) for the motion signal that was covertly attended. This result confirmed that, whereas attention is fundamental for task-relevant learning, it might have an inhibitory impact on TIPL.

Thus, after a decade of studies on TIPL, the key aspects of this phenomenon can be summarized as follows:

- TIPL is a form or learning that occurs for unattended irrelevant stimuli presented during training on a relevant task (Watanabe et al., 2001; Seitz & Watanabe, 2003)
- TIPL occurs only for stimuli that temporally coincide with behaviourally relevant events (e.g., targets of a visual task) (Seitz & Watanabe, 2003)

- TIPL is prevented when irrelevant stimuli are visible and can be noticed (Tsushima et al., 2008) or when participants exogenously attend to them (Choi et al., 2009)
- The phenomenon of TIPL seems to result from a complex interplay between mechanisms of attentional selection and internal non-specific learning signals elicited by the successful recognition of task-relevant items (Roelfsema et al., 2010; Seitz & Watanabe, 2005; Seitz & Dinse, 2007).

1.4 Target-Related Memory Enhancement (TRME)

While studies on TIPL have focused on the effects of task-relevant items on long-term learning of irrelevant visual features, several lines of research addressed the impact of behaviourally relevant events on visual memory. This research has been carried out mainly by two distinct laboratories (Lin et al., 2010; Swallow & Jiang, 2010; Leclercq & Seitz, 2012a) using a paradigm conceptually similar to the one adopted in TIPL, but with some fundamental modifications.

In a canonical TRME experiment participants are engaged in a dual task, generally characterized by a central RSVP and a stream of background images (faces or landscapes) that have to be monitored for a subsequent memory task. In the two prevalent versions of this paradigm, the RSVP and stream of images are presented recursively for several minutes or with a single stream on each trial. Then, according to the procedure used, long- or short-term memory for images is tested with dedicated tasks (e.g., Old/New familiarity task). It is important to note that, unlike the TIPL approach, in these paradigms the "irrelevant stimuli" were instead relevant for a secondary task, and thus, presented well above sensory threshold.

The first attempt to examine how task-related items may affect memory was made by Swallow and Jiang (2010), and Lin, Pype, Murray and Boynton (2010). Although these two independent studies employed slightly different paradigms, they provided some converging results. For instance, they both demonstrated that a) memory is higher for target-paired than for distractor-paired images, b) that this effect is not modality specific (occurs also with auditory targets and distractors), and, more important, c) that target items must be actively processed in order to enhance memory for background images. These findings indicated that the recognition of targets in one task enhances the processing and encoding of information for a second task (Swallow & Jiang, 2010; Lin et al., 2010), or, to put it in other terms, the attentional selection of a relevant item temporarily boosts attention to concomitant visual information, facilitating its memorization. But how does this short-term memory effect of targets recognition relate to the phenomenon of TIPL, and how can these results be reconciled with evidence showing that paying attention to the irrelevant stimuli suppresses TIPL?

To address these questions, Leclercq and Seitz (2012a) adopted a paradigm similar to the one proposed by Lin and colleagues (Lin et al., 2010) and tested whether the presentation of a central arrow instead of the RSVP would facilitate or disrupt memory for the associated image. As in Choi and colleagues (2009), the arrow served both as the relevant target, since participants were asked to report its pointing direction, and as exogenous cue, directing attention to the background stream. The prediction of Leclercq and Seitz (2012a) was that, if attention plays a crucial role in this fast form of TIPL (defined fast-TIPL) then exogenously attending to the irrelevant stimuli would increase, rather than disrupt (as in canonical TIPL), the processing of these images, favouring their memorization. Their finding confirmed this prediction by showing that scenes temporally correlated with the arrow or presented at the location where the arrow was pointing were better remembered than images preceding the arrow, and thus, attention produced an opposite effect if compared to TIPL. The crucial role of attention in TRME has been subsequently confirmed by Leclercq and Seitz (2012d), who demonstrated that when participants are instructed to ignore the background stream, thus removing any memorization and attentional requirement for the images, TRME failed to occur.

One of the works that provided a detailed analysis of the TRME effect is the study by Swallow and Jiang (2011). In this study, the authors manipulated the timing between targets and background images to test which of several hypothetical mechanisms can account for TRME. More in details, this study evaluated which of three candidate mechanisms, attentional cuing, prediction-based reinforcement learning, and perceptual grouping, could explain the enhanced memory for target-paired images. On the basis of the specific temporal parameters of these mechanisms, one could expect memory to be higher for images that occur after (attentional orienting), before (reinforcement learning) or with (perceptual grouping) the RSVP targets. However, none of these manipulations was effective, while the only factors that seemed to be fundamental for TRME were the temporal overlap between targets and images and the active monitoring of the image stream (as in Leclercq & Seitz, 2012d). Further studies have also verified that the target-induced memory enhancement is not due to the saliency or frequency of the target events (Swallow & Jiang, 2012) and does not occur for other kinds of relevant events, such as alerting cues or oddball items (Swallow & Jiang, 2012; Leclercq & Seitz, 2012a, 2012b), which do not require overt or covert responses (Leclercq & Seitz, 2012c).

Taken together, the results of these studies defined several major aspects of the effect of targets recognition on short- and long-term memory for concomitant information:

- TRME consists of a transient increase in the ability to encode and memorize visual information presented at behaviourally relevant points in time (Lin et al., 2010; Swallow & Jiang, 2010).
- In TRME, the successful recognition of targets in one task facilitates the processing and memorization of stimuli in a secondary task, provided that these stimuli are visible, actively monitored and attended (Swallow & Jiang, 2010; Leclercq & Seitz, 2012b, 2012d).

- TRME occurs exclusively for stimuli that temporally overlap with events requiring a specific response or action (Swallow & Jiang, 2011).

1.5 Mechanisms underlying TIPL and TRME

As the interest in TIPL and TRME increased, several hypotheses have been offered to explain the underlying mechanisms. On the one hand, early studies on TIPL proposed a role for endogenous signals of reward in promoting task-relevant and irrelevant learning (Seitz & Watanabe, 2003, 2005; Seitz & Dinse, 2007). On the other hand, late work on TIPL and the majority of study on TRME suggested a critical role for brain mechanisms of attentional selection (see Swallow and Jiang, 2013 for a review) and alerting (Seitz & Watanabe, 2005; Yotsumoto & Watanabe, 2008). Whereas the first suggestion rested on the assumption that detecting task-relevant items may evoke a rewarding experience, the second hypothesis postulated that targets detection is followed by a phasic alerting signal that temporarily boosts processing and encoding abilities. Although strong evidence in support of the role of such mechanisms in TIPL and TRME is still lacking, the two hypotheses presented above are worth further description and investigation.

In the next two sections I will provide some theoretical background and the results of several studies that can clarify the rationale behind these two hypotheses.

1.6 The hypothesis of endogenous reward

One of the hypotheses that have been proposed to explain the mechanisms behind TIPL and TRME is based on the assumption that performing a task may engage reward-related processes in the brain (Satterthwaite, Ruparel, Loughead et al., 2012; Huang & Watanabe, 2012; Pascucci & Turatto, 2013). This intriguing view, although mainly speculative, postulates that the successful interaction with task-relevant items triggers an internal reward, which in turn fosters reward-related learning and memory processes (Roelfsema et al., 2010). However, this implies a particular form of reward that deviates from its classical notion and that may deserve a more formal characterization.

With the term “reward” we tend to refer to objects or events that “make us come back for more” (Schultz, 2007). Generally, rewards are stimuli or experiences that increase the likelihood of the behaviour they follow or are associated with, mediate our motivation and provide a sense of accomplishment and satisfaction. More specifically, an effective reward is a positive reinforcer, a stimulus that “stamps in” specific behavioural patterns (Schultz, 2002) or cognitive processes (Della Libera & Chelazzi, 2006) and that can alter the value and saliency of stimuli and contexts that are paired with it (Hickey, Chelazzi & Theeuwes, 2010). While this definition holds true for many forms of reward, the distinction that is required by the present framework concerns more the nature of a rewarding experience, and particularly, its exogenous or endogenous origin.

Exogenous rewards are reinforcers that mediate motivation, satisfaction and sense of accomplishment and that are provided to the organism by the external world. Typically, they are environmental stimuli that satisfy primary survival needs, like food or liquids, or secondary and social needs, like money or verbal reinforcements. These stimuli are defined to be exogenous since they are located and originate from outside the individual.. Under some circumstances, however, a reward can be experienced even in the absence of external factors. Feeling gratified by particular thoughts, memories or by performing our preferred activity are all examples of events that originate internally and may evoke endogenous signals of reward (Ainslie, 2013).

If we refer to this meaning of endogenous reward, it is possible to argue that the recognition of a visual target may be, *per se*, rewarding, since a target represents a behaviourally relevant event. On a lower level perspective, in addition, and without summoning high-level processes of self-accomplishment and gratification for task execution, it can be also hypothesized that the attainment of sufficient sensorial evidence about the occurrence of a relevant stimulus may automatically trigger an internal (neuromodulatory) signal from reward-related areas, which action, similar to reward in reinforcement learning, would be to reinforce and stamp-in the actual neural representation, favouring memory and perceptual processing (Roelfsema et al., 2010).

However, in order to represent a strong account for TIPL and TRME, the endogenous reward hypothesis requires at least some evidence that the successful detection of a target evokes activity in brain areas known to respond to exogenous reward, and that once reward-related structures are triggered (exogenously or endogenously) their activity correlates with changes in sensory areas and in perceptual episodes.

With respect to the first point, recent studies with functional imaging have provided some results that support the involvement of reward-related areas during the correct execution of a task. In particular, Daniel and Pollmann (2012) examined the neural bases of neural signals of being correct and confident in a category-learning task without external feedback or reward. They found that activity in the humans striatum was overall higher during correct than incorrect performances, and that this modulation was similar to the one observed when humans are provided with external feedback or reward (Daniel & Pollmann, 2012). In a subsequent study, Satterthwaite and colleagues (2012) reported a similar increased activity in the ventral striatum for correct versus incorrect responses in a large sample of human subjects engaged on a visual memory task. They presented participants with a visual memory task under three different levels of difficulty and reported a larger activation of the ventral striatum for correct than incorrect trials that was also scaled as a function of the task difficulty (Satterthwaite et al., 2012). These findings, in addition to supporting the hypothesis of endogenous reward, indicated the existence of a complex internal mechanism that detects task completion and may be sensitive to task difficulty and confidence.

But how can these internal signals of reinforcement (or reward) affect perceptual processing of task-relevant and irrelevant information? One possibility is that when reward-related signals (e.g., dopamine) are globally released, they have a neuromodulatory impact on the activity of other brain areas, strengthening the neural patterns elicited by the current, relevant and irrelevant, sensory information (Roelfsema et al., 2010).

Although there is reliable evidence of reward-related modulation in sensory areas (Bao, Chan & Merzenich, 2001; Reader, 1978; Arsenault, Nelissen, Jarraya & Vanduffel, 2013) a more direct indication of the effect of reward on the processing of irrelevant stimuli has been presented by Seitz, Kim and Watanabe (2009). In this work, food and water deprived human participants were extensively exposed to two subliminal gratings of distinct orientation. Crucially, only one of the two gratings was exposed at the time of liquid (water) delivery. Participants improved their ability to detect the oriented grating that was associated with the liquid reward, whereas their sensitivity to the other stimulus was unaltered. This finding showed that the simple pairing of a subliminal stimulus with reward evokes long-lasting changes in perceptual representations, a result that has been confirmed by physiological studies in awake macaques (Frankó, Seitz & Vogels, 2010).

Hence, the hypothesis of endogenous reward seems to represent a plausible account of the target-induced perceptual enhancements, based on the rationale that the successful targets recognition triggers reward signals, which in turn, reinforce and stamp in the concurrent sensory representation.

1.7 Alerting and temporal selection

A different account of how targets recognition may affect general perceptual processing and memory has been proposed in later research on TIPL (Yotsumoto & Watanabe, 2008) and in the majority of works on TRME (Swallow & Jiang, 2013). These studies referred to two main, and not necessarily separated, mechanisms, namely, attentional alerting and temporal selection. Contrary to the endogenous reward explanation, which focuses on the rewarding aspect of target detection, the alerting account relies on the fact that temporal cues, such as the appearance of a relevant stimulus, induce a phasic increase of alertness and arousal (Aston-Jones & Cohen, 2005).

This supposition is based on the definition of attention proposed by Posner and colleagues (Posner & Petersen, 1990), according to which, orienting, alerting and executive control represent three dissociable attentional subsystems. The role of orienting is to channel resources to specific locations or features, the executive control subsystem is supposed to monitor and solve task conflicts, and the alerting system is activated by the occurrence of important stimuli and leading to a transient and non-specific increase in attentional and processing abilities (Seitz & Watanabe, 2005). This latter mechanism, therefore, appears a likely candidate to explain the perceptual and memory enhancements evoked by targets. However, as pointed out by Swallow and Jiang (2013) the idea that TIPL and TRME may be due to a global increase of attention caused by alerting cues is not fully justified. The criticism relies on the fact that the facilitatory effect of targets is limited to stimuli that co-occur with it (Swallow & Jiang, 2011) whereas there is no evidence of increased memory or learning for stimuli presented briefly after the target, when the effects of alerting and arousal should be to be greater (Swallow & Jiang, 2013). Furthermore, it has been demonstrated that alerting cues *per se* (stimuli that indicate the arrival of a target but do not require a response) have no effect on the processing of irrelevant stimuli (Leclercq & Seitz, 2012c).

An alternative account based on attentional mechanisms has been proposed by Swallow and Jiang (2013) and would involve temporal selection processes. According to this view, a specific attentive mechanism, selective for time but not for space or modality, would maximize the processing of relevant stimuli by enhancing their representation and by protecting such processing from interference by subsequent irrelevant information (Swallow & Jiang, 2013). In line with theories that account for the attentional blink phenomenon (Olivers & Meeter, 2008), the temporal selection interpretation postulated a sort of temporal "gating" by which an attentional gate is opened by the detection of relevant items and switched off instantaneously after, in order to reduce interference by distractors. Crucially, when temporal selection takes place, attentional resources are at their maximum and can spillover from the main task to contextual information, leading to TIPL and TRME.

Whereas the hypothesis of endogenous reward suggested a crucial role in TIPL for neuromodulatory signals within the mesocortical/mesolimbic reward circuit, the temporal selection account assumes that the transient attentional and perceptual boost reported in TIPL and TRME may be caused by the phasic activity of the locus coeruleus-norephinephrine (LC-NE) system. This proposal is supported by the fact that of LC neurons activity is known to be modulated by the occurrence of behaviourally relevant events, such as items that require a response or a change in behaviour (Aston-Jones & Cohen, 2005) and , furthermore, there is consistent evidence that the transient firing of LC neurons results in a release of NE to afferent brain regions, which role is to boost neural activity and promote the process by which the network settle into a final representational state (Swallow & Jiang, 2013). In other terms, the detection of a relevant change in the environment (e.g., the appearance of a target) would trigger a transient increase in LC activity, which in turn, would reinforce the current attentional- and perceptual-related activity in other brain areas.

1.8 Target-elicited signals in models of learning and attention

On the ground of the evidence presented so far, recent models of perceptual learning and attention have implemented target-elicited brain signals as key components (Roelfsema et al., 2010; Swallow & Jiang, 2013; Seitz & Watanabe, 2005; Yotsumoto & Watanabe, 2008). For the aim of the present work, it is useful to introduce the main aspects of at least two models that may represent valid approximations of the TIPL and TRME phenomenon.

According to the attention-gated reinforcement learning (AGREL) account proposed by Roelfsema, van Ooyen and Watanabe (2010), learning and brain plasticity are determined by the joint action of attentional and reinforcement signals. AGREL postulates that when reward (endogenous or exogenous) follows task accomplishment, specific neuromodulators (e.g. dopamine) are released in the brain, strengthening the active synaptic connections and reinforcing stimulus representations and stimulus-response associations. The key role of attention, thus, is to restrict and drive the action of these non-specific signals to neural patterns that were relevant for the task execution, avoiding unwanted plasticity. In other words, whereas reward evokes a global reinforcing signal, attention acts as a gate allowing relevant sensory representations to be reinforced and inhibiting plasticity for irrelevant information. Interestingly, this interplay between attention and reward seems to represent a reliable account of both TIPL and TRME. On the one hand, it would explain why supraliminal stimuli paired with targets need not be completely irrelevant (and thus, require active attentional processing) in order to be memorized in TRME; on the other hand, it would justify why subliminal irrelevant stimuli that are too weak to be noticed and to be suppressed by the attentional control, become learned in TIPL.

A different account relying on the attempt to reconcile the TRME effect with theories of attentional load (Lavie, 2005) and dual-task interference (Pashler, 1994), comes from Swallow and Jiang (2013). Their theoretical model, defined "dual-task interaction", is suited for conditions of divided attention, when information related to two independent tasks (as the RSVP and image stream in TRME) has to be constantly monitored and stimuli may compete for perceptual processing. The authors suggested that in dual tasks, a key role is played by a central executive mechanism (EC) (Baddeley, 1996) and by temporal selective attention (Desimone & Duncan, 1995). Whereas the EC maintains control over the goal of the two tasks and guarantees early processing of the whole relevant information, temporal attention is a temporally specific mechanism triggered by the occurrence of items that require a response. Once temporal selection occurs, it causes the release of phasic boosting signals (e.g. norepinephrine from the locus coeruleus) that enhance the processing of all perceptual information, turning the competition between simultaneous stimuli into a positive interaction. More precisely, the dual task interaction model suggests the existence of limited temporal windows where the correct detection of a target facilitates the processing and encoding of information of a second task, ensuring that both stimuli are maintained in memory (Swallow & Jiang, 2013).

The AGREL and the dual task interaction model, although not mutually exclusive, describe similar mechanisms of target-elicited processing enhancement but embed different assumptions about the origins of this phenomenon. In the AGREL model, perceptual enhancement is driven by (implicit or explicit) feedbacks on correctness or reward regulated by top-down attentional settings; in the dual task interaction account, instead, phasic boosts of sensory representations are evoked by the attentional selection of task-relevant episodes, resembling a more mechanistic process that may not require internal reward associated with task accomplishment. Therefore, despite the concrete difficulty to disentangle the pure effect of reward from that of attention (Maunsell, 2004), it may be interesting to test the predictions that can be derived from these models, in order to provide a clearer understanding of the impact of target stimuli on our perception, learning and memory.

1.9 Summary of the present work

The aim of this manuscript is to further investigate perceptual changes induced by the detection of visual targets. This thesis contains four experimental studies designed to address the impact of relevant stimuli on short-term forms of cortical plasticity and to evaluate the role played by other factors, such as task-difficulty, monetary reward and task precision in this phenomenon. The manuscript is organized according to the following structure:

- In *Chapter 2*, I will present the results of three experiments showing that the detection of RSVP targets enhances the tilt aftereffect (TAE), a brief visual adaptation reflecting a form of short-term visual plasticity. This study represents first evidence that behaviourally relevant events can alter our immediate perceptual experience.

- *Chapter 3* reports a series of experiments that further explore this phenomenon, investigating whether the difficulty of targets recognition or the number of correct recognitions are crucial.
- In *Chapter 4* it is described a recent research where I manipulated the reward value associated with a target to demonstrate the involvement of reinforcement mechanisms in sensory plasticity for invisible inputs (TIPL).
- Finally, *Chapter 5* will present preliminary results of a functional magnetic resonance imaging study aimed at identifying the neural substrate of internal signals of reinforcement.
- The experimental results reported in this work will be summarized in *Chapter 6*.

Each study is introduced by a brief summary.

Chapter 2

Immediate effect of internal reward on visual adaptation

The present study has been published under the above title by Pascucci, D., and Turatto, M. in *Psychological Science* (2013).

The following chapter reports the results of one study where we (Pascucci and Turatto, 2013) investigated the effect of targets recognition on a short-term form of visual adaptation, namely, the tilt-aftereffect illusion (TAE). The idea behind this work came after a thorough discussion about the findings of Watanabe and colleagues (Seitz & Watanabe, 2003) and relies on the intriguing hypothesis that target-elicited signals of reinforcement may temporally alter visual plasticity. We wandered whether a similar phenomenon could be isolated in the frame of a single trial. To address this question, I implemented the defying aspects of TIPL (the RSVP and a surrounding irrelevant stimulus) in a paradigm of visual adaptation.

The particular measure of adaptation that we adopted (the TAE) reflects a short-term form of cortical plasticity that is induced by the prolonged exposure to tilted stimuli (Dragoi, Sharma & Sur, 2000). As a consequence of adaptation, the processing of new oriented stimuli is momentarily corrupted, inducing perceptual biases that can be measured within few seconds. Crucially, the degree of these biases depends on several parameters that regulate the strength of the adapter (e.g., the duration, intensity or size). Our hypothesis was that, if detecting relevant stimuli boosts our perceptual experience, the recognition of targets during adaptation would alter the strength of the adapter and thus, would modulate the magnitude of the resulting visual illusion.

Abstract

In the last decade there has been an increasing interest in the effects of rewards on visual perception. Exogenous rewards have been shown to increase visual sensitivity and to affect attentional selection. Human beings, however, feel rewarded also by the correct accomplishment of a task. It has been proposed that this form of endogenous reward triggers reinforcement signals in the brain, making the sensory system more sensitive to those stimuli extensively and repeatedly paired with the rewarding experiences and modulating long-term cortical plasticity. Here, we report the striking observation that a well-known visual illusion, the tilt aftereffect, which is due to a form of short-term cortical plasticity, is immediately enhanced by a concurrent and independent target recognition process. Our results show that endogenous rewards can alter our visual experience with virtually no delay.

2.1 Introduction

Rewards not only modify behaviour, but also our attentional priorities and perceptual analysis (Della Libera & Chelazzi, 2009; Kiss, Driver & Eimer, 2009; Raymond & O'Brien, 2009). Furthermore, the visual system can develop an increased sensitivity to stimuli that, albeit unattended and unconsciously processed, are repeatedly and extensively paired with exogenous rewards (Seitz, Kim & Watanabe, 2009). Human beings, however, usually feel rewarded also by the sense of achievement evoked by the correct execution of a task, such as, for example, the identification of a target. It has been suggested that like exogenous rewards, endogenous ones trigger reinforcement signals in the brain (Karni & Bertini, 1997; Seitz & Watanabe, 2003). These signals, in turn, lead to enhanced processing of incoming sensory information, making the sensory system more sensitive to those stimuli repeatedly paired, over several days of training, with the rewarding experience (Sasaki, Náñez & Watanabe, 2010; Seitz & Watanabe, 2005). In addition, in rats learning-related cortical reorganization can also be induced by direct stimulation of reward-related neural structures (Bao, Chao, Merzenich, 2001). Hence, rewards can increase visual sensitivity through a slow incremental learning mechanism.

However, rewards can exert much more rapid effects on selective attention (Della Libera & Chelazzi, 2006). For example, if a given visual feature such as the color of an item is paired with a monetary reward on the current trial, attention will tend to be deployed to items possessing the same feature on the next trial

(Hickey, Chelazzi & Theeuwes, 2010). Furthermore, it has recently been shown that endogenous rewards can also have rapid effects on visual memory, as documented by fast task-irrelevant perceptual learning (Leclercq & Seitz, 2012a, 2012b).

But can reward alter also our immediate visual experience? To address this question we investigated whether the rewarding experience associated with targets recognition can affect the short-term form of plasticity that underlies a phenomenon of visual adaptation, the direct tilt aftereffect (TAE). This is a well-known visual illusion (Dragoi, Sharma & Sur, 2000), in which prolonged exposure to an oriented stimulus affects subsequent perceived orientations (Gibson & Radner, 1937). For example, if a first grating (the adapter) is presented for some seconds tilted to the left, then a subsequent vertical grating (the test) appears tilted to the opposite direction. Therefore we predict that if endogenous rewards can increase the strength of the adapter during its exposure, then a more pronounced TAE should be observed.

2.2 Experiment 1

The present experiment investigated whether the successful execution of a discrimination task during the adaptation period induces a boost in the TAE. To this purpose, at the end of each trial observers reported the orientation of a briefly presented low-contrast test grating tilted at various orientations to the left or to the right. The test grating appeared after a period of exposure to an adapter grating, at the centre of which a series of characters was presented in quick succession (rapid serial visual presentation; RSVP) for approximately 5 s (Figure 1*a*).

2.3 Methods

2.3.1 Participants

Twenty students (14 females; mean age = 21.3±2.77) from the University of Trento, Italy, participated in the experiment. In this and the following experiments all participants had normal or corrected to normal vision and were naïve as to the purpose of the experiment. Written informed consent was obtained from all participants, and the experiment was carried out in accordance with the Declaration of Helsinki.

2.3.2 Apparatus

In this and the following experiments stimuli were presented on a ViewSonic G90fB CRT 19” monitor (1280 x 1024, 85 Hz). Generation and presentation of the stimuli was made using Matlab and the Psychophysics Toolbox 3.8 (Pelli, 1997) on a Windows 7 based machine.

2.3.3 Stimuli

In all experiments, the RSVP was presented on top of a grey disk (44.7 cd/m2, 1° visual angle in diameter) centred on the adapter grating at the centre of the computer screen. In Experiment 1, the targets, presented on half of the trials, were three different numbers embedded within a RSVP sequence of distracter letters superimposed on a flickering (45 Hz) full-contrast adapter grating. The distracters had random values of luminance drawn from a Uniform Discrete Distribution

(±20% of contrast from the grey background, 44.7 cd/m2). The targets were randomly chosen from 2 to 9, and their luminance changed according to a staircase procedure pointing at 75% of correct recognition (of three targets). The contrast of the adapter and test gratings were fixed at 100% and 20% respectively, and both covered 3° of visual angle, with a spatial frequency of 1.5 c/d. The former was tilted -10° off vertical, whereas the latter could have one of five possible orientations: -4°, -2°, 0°, +2°, or +4° off vertical.

The RSVP sequence consisted of 50 frames (4400 ms), and each frame of 88 ms duration contained a single character. In the “Target-present” condition, the targets could appear at any point between the second frame from the beginning to the second-to-last frame in the RSVP sequence. In the “Target-absent” condition no targets were presented in the RSVP sequence. In this and the following experiments, two consecutive targets were separated by a minimum of 5 frames (440 ms).

The adapter was presented for the whole duration of the RSVP sequence.

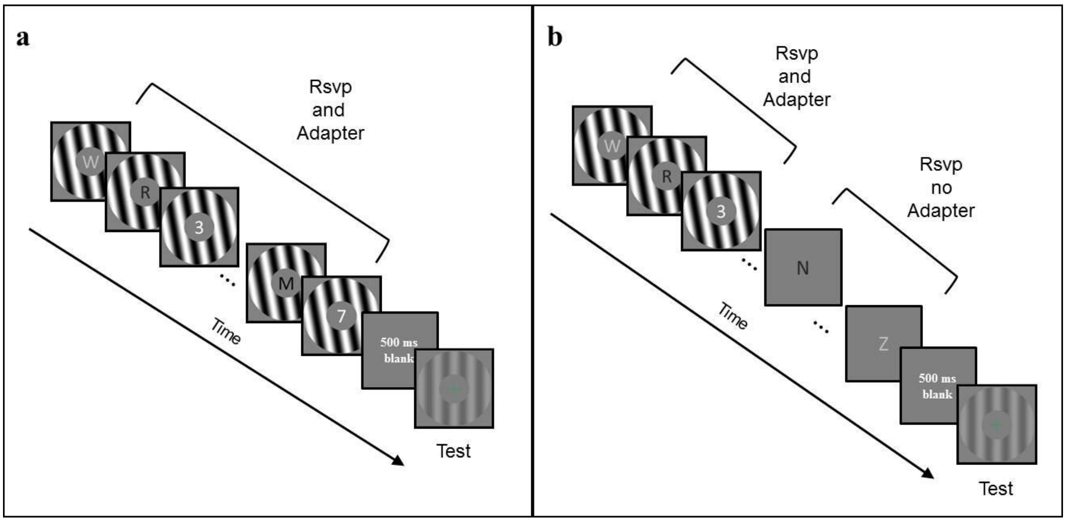

Figure 1. Panel *a* depicts the sequence of events in Experiments 1 and 3. In the former, participants decided whether the test grating was either vertical or tilted. In the latter, participants were to discriminate whether the test grating was rightward or leftward tilted. In both experiments the adapter was always present during the RSVP sequence. Panel *b* depicts the sequence of events in Experiment 2, in which the targets were always presented in the RSVP sequence, but the adapter was removed during the last part (approximately 1 sec) of the sequence. Participants were asked to decide whether the test grating was either vertical or tilted.

2.3.4 Procedure

Participants sat at approximately 60 cm from the monitor in a dark room. Each trial started with a fixation cross lasting 500 ms, followed by the adapter and the RSVP. After a 500-ms blank interval, the test grating was presented for 50 ms, and participants had to report its orientation (Vertical vs. Tilted) by pressing one of two different keys on the computer keyboard. Then a blue question mark appeared at the centre of the screen, and participants were asked to report the identity of the targets using the corresponding keys (they pressed the space-bar to indicate the absence of targets in the RSVP sequence). Visual feedback was given after each response.

The experiment consisted of 200 trials. On 80 trials (40%) the orientation of the test gratings was vertical, and on the remaining 120 trials (60%) it was tilted from vertical (30 trials for each of the four inclinations).

2.4 Results

Data (proportions of “vertical” responses) were entered into a two-way repeated-measure analysis of variance (ANOVA), with Target presence and Test orientation as factors. Both main effects were significant, Target presence F(1,19), 7.550, p = .013; Test orientation F(4,76), 30.128, p = .0001. Inspection of Figure

2a suggests two important considerations. First, we observed a robust TAE, attested by the fact that the proportion of vertical responses was maximum when the test was leftward tilted rather than vertical (as instead happened when, in a control condition depicted by the grey line, the adapter was not presented). Second, pairwise comparisons (t-tests), confirmed that the visual illusion was significantly stronger when, during adaptation, the target digits were successfully identified as compared to when no targets were presented (p = .004, at -4° ; n.s., at -2°; Bonferroni corrected). Data were also analysed with best-fitting procedures, by comparing (t-tests) the parameters of the psychometric functions derived for the two conditions. The results were consistent with those obtained with the ANOVA. The same consistency emerged also in the next experiments.

This result is consistent with the idea that target recognition, which occurred on the majority of trials (~80% of correct responses), may have been accompanied by a sense of fulfilment which acted as an internal reward (Karni & Bertini, 1997), thus enhancing the representation of the concomitant stimuli (Nishina, Seitz, Kawako & Watanabe, 2007; Seitz & Watanabe, 2003), namely the adapter, and leading to a stronger visual illusion.

2.5 Experiment 2

Although the results of Experiment 1 are consistent with the hypothesis that target recognition triggered reinforcement signals that strengthened the representation of the adapter, an alternative interpretation is possible. Since there is evidence that attention potentiates the TAE (Spivey & Sprin, 2000), one may hypothesize that target-present and target-absent trials involved different attentional loads. Specifically, participants could have paid more attention to the RSVP sequence when they detected the occurrence of the targets. If this were the case, and assuming that paying more attention to the RSVP entails also a larger amount of attention deployed to the underlying adapter, then the stronger TAE could be attributed to different attentional loads between target-present and target-absent trials.

To rule out this possibility as an explanation of the results of Experiment 1, in Experiment 2 the targets were always presented in the RSVP sequence. However, they could occur while the adapter was visible or after it was removed (approximately 1 s before the end of the RSVP sequence; see Figure 1b). If the strength of the illusion is increased by the endogenous reinforcement signals, then this should occur only when the targets are presented (and recognized) while the adapter is visible, but not after its removal. By contrast, since the targets were always presented and their time of occurrence was unpredictable, the RSVP sequence should be continuously and uniformly attended. Therefore the attentional explanation predicts no difference in the TAE between the two conditions.

2.6 Methods

2.6.1 Participants

Eighteen students (15 females; mean age = 21.3±2.91) from the University of Trento, Italy, participated in the experiment.

2.6.2 Apparatus, stimuli and procedure

As in Experiment 1 but with the following exceptions. The targets were always present during the RSVP sequence, which consisted of 60 frames (5280 ms), but the adapter was presented only for the first 40 frames. Hence, in the "During adaptation" condition the targets appeared between the 2nd and 40th frame, whereas in the "After adaptation" condition they were presented during the last 15 frames. The experiment consisted of 240 trials. On 96 trials (40%) the orientation of the test gratings was vertical, and on the remaining 144 trials (60%) it was tilted from vertical (36 trials for each of the four inclinations).

2.7 Results

Data (proportions of "vertical" responses) were entered into a two-way repeated-measure ANOVA, with Target time and Test orientation as factors. The main effect of Test orientation and the interaction were significant, Test orientation F(4,68), 242.182, p = .0001; Target time x Test orientation F(4,68), 2.865, p = .03. The results (see Figure 2b) replicated those emerged in Experiment 1. Pairwise comparisons (t-tests) confirmed that the visual illusion was indeed significantly stronger in the “During adaptation” than in the “After adaptation” condition (p = .009, at -4°; p = .024, at -2°; Bonferroni corrected).

The present pattern of results is entirely consistent with the reward interpretation, whereas it does not appear to be compatible with the “different attentional load” explanation.

2.8 Experiment 3

In the previous experiments we interpreted the increased probability of the test to be reported as being vertical, when tilted at -4°, as evidence that the rewarding effect of target discrimination enhanced the strength of the adapter. However, we cannot a priori exclude the possibility that, for some reason, the result was instead due to a stronger decisional bias toward the “vertical” response in the target-present than in the target-absent condition, rather than to a genuine change in the appearance of the test grating orientation. To control for this possibility, we asked participants to report whether the test was left or right tilted. In addition, this allowed us to control how the TAE was modified by targets recognition. If in the present paradigm endogenous rewards increase the strength of the adapter, then on target-present trials the test should be more often perceived as being tilted to the right then to the left, at least for some values of the tested orientations.

2.9 Methods

2.9.1 Participants

Fourteen students (12 females; mean age = 21.3±1.34) from the University of Trento, Italy, participated in the experiment.

2.9.2 Apparatus, stimuli and procedure

As in Experiment 1 but with the following exceptions. The RSVP sequence consisted of 20 frames (1760 ms). We extended the possible orientations of the test grating, which were the following: -5°, -3°, -2°, +2°, +3° or +5° off vertical. Participants had to report whether the test was leftward or rightward tilted. The experiment consisted of 480 trials, with 80 trials for each of the six orientations (40 trials for "Target-present" condition and 40 trials for "Target-absent" condition).

2.10 Results

Data (proportions of "right" responses) were entered into a two-way repeated-measure ANOVA, with Target presence and Test orientation as factors. The main effect of Test orientation, as well as the interaction, were significant: Test orientation F(5,65), 51.928, p = .0001; Target presence x Test orientation F(5,65), 4.565, p = .001. The results (see Figure 2c) confirmed and extended those emerged in Experiments 1 and 2, showing that the TAE was enhanced if the targets appeared during the period of adaptation. Pairwise comparisons (t-tests) confirmed that the visual illusion was enhanced by targets recognition, as attested by a larger proportion of "right" responses in the "Target present" than in the "Target absent" condition (p = .005, at -2°; Bonferroni corrected).

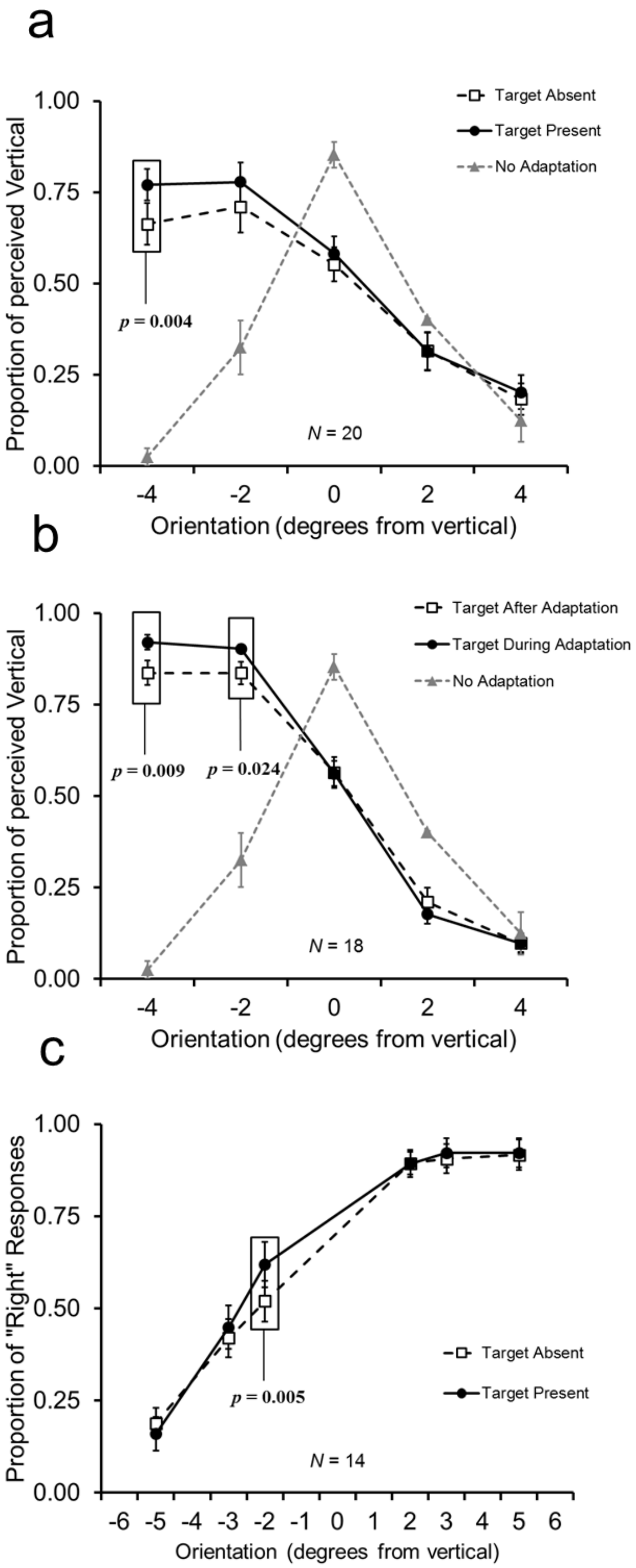
a
Target Absent
Target Present
No Adaptation
Proportion of perceived Vertical
1.00
0.75
0.50
0.25
0.00
-4
-2
0
2
4
Orientation (degrees from vertical)
p = 0.004
N = 20
b
Target After Adaptation
Target During Adaptation
No Adaptation
Proportion of perceived Vertical
1.00
0.75
0.50
0.25
0.00
-4
-2
0
2
4
Orientation (degrees from vertical)
p = 0.009
p = 0.024
N = 18
c
Proportion of "Right" Responses
1.00
0.75
0.50
0.25
0.00
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
Orientation (degrees from vertical)
p = 0.005
Target Absent
Target Present
N = 14

Figure 2. Proportions of “vertical” (Panels *a* and *b*) or “right” (Panel *c*) responses to the test grating (for the orientations tested) as a function of the target condition in Experiments 1, 2 and 3 respectively. The grey line depicted in Experiments 1 and 2 represents data from a control group of 5 participants tested with no adaptation.

2.11 General Discussion

In three experiments, we found that the TAE was magnified by targets recognition during the adaptation period. We used a central discrimination task similar to that adopted in task-irrelevant perceptual learning studies. This form of unintentional learning, which typically involves unseen and unattended stimuli, is explained by proposing that the identification of the targets in the RSVP sequence triggers endogenous reinforcement signals which modulate cortical plasticity and induce long-lasting learning (Sasaki, Náñez & Watanabe, 2010; Seitz & Watanabe, 2003). Here, we showed that the same mechanism can also alter the TAE, a visual illusion due to a form of short-term plasticity whose neural bases are located in the primary visual cortex (Dragoi, Sharma & Sur, 2000). This seems to suggest that endogenous reinforcement signals, like exogenous ones, can reach early stages of cortical visual processing (Pilly, Grossberg & Seitz, 2010; Shuler & Bear, 2006; Watanabe et al., 2002).

However, the effect of reward on visual perception we documented is different from that previously reported in task-irrelevant perceptual learning studies (e.g., Seitz & Watanabe, 2003). First, it resulted in a rapid phenomenological change of perception (the perceived inclination of the test grating), not in a more accurate representation of the reward-paired stimulus. An analogous phenomenological change in perception, likely caused by endogenous rewards and consisting in a perceptual bias in motion detection, was already reported by Seitz et al. (2005), but this emerged only after extensive training. Second, because the effect emerged for the same adapter and only on target-present trials, which were intermingled with target-absent trials, the perceptual change must have occurred at the single-trial level (also see Leclercq & Seitz, 2012a, 2012b). Such fast alteration of visual perception induced by endogenous rewards is quite different from the previously reported increment in visual sensitivity (e.g., Seitz, Kim & Watanabe, 2009), which instead relies on a slow cumulative learning process that gradually strengthens or refines the neural representation of the stimulus repeatedly paired with the reinforcement signals.

To conclude, expanding upon previous findings showing that monetary rewards can immediately influence attentional priority, and that endogenous rewards can rapidly alter visual memory, the present study indicates that endogenous rewards can also induce a fast change in visual sensory experience by modulating the form of short-term cortical plasticity underlying the TAE.

Chapter 3

Task Accomplishment, Endogenous Reward and short-term Visual Adaptation

The present study has been submitted under the above title by Pascucci, D., Infanti, E., Mastropasqua, T., and Turatto, M. to *Visual Cognition*.

The previous study revealed an interesting phenomenon, that is, the recognition of visual targets during adaptation increases the magnitude of the TAE. This finding, in addition to provide first evidence of the impact of target-elicited signals on short-term plasticity, suggested us two important questions. The first question relies on the role of task difficulty. Under the assumption that recognizing visual targets evokes endogenous signals of reward, we hypothesized that the successful detection of difficult targets may trigger higher internal reward than the recognition of easy targets, and thus, difficulty could differentially impact adaptation to irrelevant stimuli. As for the second question, we took advantage of our particular paradigm, where the effect of multiple targets can be tested on the same irrelevant stimulus, to investigate whether the number of targets presented (and recognized) during adaptation is crucial. In the following study, we addressed these two questions in three dedicated experiments where we manipulated task-difficulty and the number of relevant targets during adaptation to tilted gratings.

Abstract

An increasing number of studies on task-irrelevant perceptual learning has shown that the perception of behaviourally relevant events can lead to improved sensory analysis of a concomitant-but-irrelevant stimulus' features. It has been suggested that the correct identification of the task-relevant target works as an endogenous reward for the brain, which would release reinforcement signals that gate cortical plasticity for both relevant and irrelevant information. Recently, we have shown that the identification of targets embedded in a rapid serial visual presentation enhanced the tilt aftereffect (TAE), a brief visual adaptation reflecting a form of short-term visual plasticity. Here, in three experiments, we further explored this phenomenon to test whether the putative reinforcement-signal mechanism is more sensitive to the difficulty of target recognition, or to the number of such recognitions. Our results show that the main factor affecting the magnitude of the TAE was the latter, whereas task difficulty had little or no effect. This pattern of results is in agreement with the possibility that the detection/recognition of a behaviourally relevant event acts as an endogenous reward, automatically releasing reinforcement signals that affect long- and short-term cortical plasticity.

3.1 Introduction

Exogenous rewards are events or objects that we may encounter in the external world and that can shape our motivation, satisfaction and sense of accomplishment, and hence, ultimately, our behaviour. Primary rewards, like food, liquids or sex, are stimuli of interest because they satisfy basic survival and biological needs. However, we are also familiar with secondary or conditioned rewards like money (by means of which primary rewards can be obtained), or positive verbal reinforcements (that satisfy social needs like the sense of approval). All these forms of reward have been termed “exogenous”, essentially because they are provided to the individual by the surrounding environment. Human beings, however, are peculiar animal creatures capable of fabricating an “endogenous” rewarding experience in the absence of any obvious exogenous reward. Indeed, we are all familiar with the distinct sentiment of satisfaction felt when we accomplish a task, even when there is no obvious immediate reward at stake, like when we finally understand a difficult notion or score a goal during a soccer match among friends. These informal observations seem to suggest that, in the absence of explicit external feedback or reward, the brain might possess a mechanism for evaluating whether a given task has been accomplished, and when this is the case a feeling of gratification is generated that acts as endogenous reward (Ainslie, 2013).

While in the recent years the interaction between exogenous reward, perception and attention has attracted the interest of many researchers (e.g., Anderson, Laurent & Yantis, 2011; Della Libera & Chelazzi, 2006, 2009; Hickey, Chelazzi & Theeuwes, 2010; Raymond & O'Brien, 2009), the role of endogenous reward in many perceptual and cognitive processes remains much less explored. One of the firsts attempt to understand how endogenous reward could operate in humans dates to the beginning of the last century, when the psychologist Raymond Dodge introduced the term "autogenic reinforcement" (Dodge, 1917) to describe the process of internal reinforcement that is active when a task is completed. Similarly the term "endogenous reward" has been coined to indicate the release of a brain reinforcement signal that, when no external feedback or reward is available, follows the self-evaluation that a correct (overt or covert) response has been provided (Holroyd & Coles, 2002; Karni & Bertini, 1997). This idea of an internal reinforcement signal triggered by task completion has been proposed to regulate sensory improvements and cortical plasticity (Sasaki, Náñez & Watanabe, 2012). For example, in one of their early studies on the task-irrelevant perceptual learning (TIPL) phenomenon, Seitz and Watanabe (2003) proposed that endogenous reward signals would play a pivotal role in perceptual learning, favouring long-term plasticity for relevant and irrelevant visual stimuli. In their experiments (Seitz & Watanabe, 2003; Seitz, Lefebre, Watanabe & Jolicoeur, 2005), participants were trained on a central rapid serial visual presentation (RSVP) task for several days without feedback, while a cloud of task-irrelevant dots randomly moved in the periphery. Crucially, on each trial a small proportion of dots subliminally but coherently moved in a given direction, which was different for the target and distractors of the RSVP. After extensive training with this procedure, participants improved not only in the execution of the RSVP task, but also in their ability to detect or discriminate the subliminal motion direction that was exclusively paired with the targets occurrence (Seitz & Watanabe, 2003). Importantly, the correct identification of the target was a crucial condition for TIPL to emerge (Seitz et al., 2005). On the ground of these results, it has been suggested that the successful recognition of behaviourally relevant events acts as endogenous reward that generates a signal of reinforcement, which temporally increases the brain's sensitivity to the concurrent task-relevant and task-irrelevant sensory information (Seitz & Watanabe, 2005). According to this view, long-term perceptual learning and brain plasticity would arise from the extensive reiteration of this process during training (Roelfsema, van Ooyen & Watanabe, 2010).

Moving from this theoretical framework, in a previous study we investigated whether the recognition of relevant events could affect other forms of visual processing, occurring on a shorter time scale (Pascucci & Turatto, 2013). To address this question we measured whether the magnitude of the tilt aftereffect (TAE) could be affected by the identification of targets in an independent task. The TAE occurs after a period of exposure to an oriented adapting stimulus, and consists in a misperceived orientation of a subsequent test stimulus toward the direction opposite to that of the adapter. This illusion is conceived as a manifestation of short-term cortical plasticity in the early visual cortex (Dragoi, Sharma, & Sur, 2000). We hypothesized that targets recognitions, during visual adaptation, should activate internal signals of reinforcement, which in turn would enhance the representation of the adapter, leading to a more pronounced TAE. Our results supported this prediction, as we found that the TAE increased in magnitude when the adapter overlapped with targets recognition, suggesting that the hypothesized mechanism of endogenous reward affected visual perception and the corresponding short-term plasticity.

The putative role of the endogenous-reward mechanism is also supported by the results of recent neuroimaging studies. Daniel and Pollmann (2012) examined the neural activations reflecting accuracy and confidence during the execution of a category-learning task without external feedback. The authors found greater activation in the striatum for correct than incorrect responses, with patterns of activities similar to those observed in the presence of external rewards. A subsequent study by Satterthwaite, Ruparel, Loughead et al. (2012) reported converging evidence showing increased activity in the ventral striatum for correct versus incorrect responses in a visual memory task. The authors used a “n-back” memory task with different conditions. In one condition participants had to detect the occurrence of a target image during a series of trials, whereas in the other two conditions they were asked to memorize the images presented in two and three consecutive trials, and to report possible repetitions of the same image. The three conditions differed in terms of memory load and task difficulty, and were presented in separate blocks of trials. The results showed a larger activation of the ventral striatum for correct trials that was scaled as a function of the task difficulty. This is in agreement with the folk notion that in real-life situations the successful accomplishment of hard tasks feels more rewarding than the attainment of easy tasks. Hence, is seems that the magnitude of an endogenous reward is, at least partially, proportional to the difficulty of the task at hand.

A study by Huang and Watanabe (2012) further explored this issue. The authors measured, in separate experiments, both memory for background images and TIPL for irrelevant motion direction, presented during an easy or hard RSVP task. The result showed that memory was better for background images paired with easy than with hard detectable targets, whereas the opposite was true for TIPL, which was larger with hard targets. To explain this pattern of results, the authors argued that the memorization of the background images required attention, and therefore the more attention was dedicated to the targets the less remained available for processing the background image. By contrast, TIPL for subliminal motion direction did not require attention (Tsushima, Seitz & Watanabe, 2008), but instead depended on reinforcement signals elicited by the target recognition (Seitz & Watanabe, 2003; Seitz et al., 2005). In addition, because of the low expectation of a successful task execution in the hard condition, difficult targets recognitions were more rewarding than easy targets recognitions, thus triggering stronger reinforcement signals (Huang & Watanabe, 2012)

If the brain has a mechanism that detects task completion, and that releases a reinforcement signal influencing long- and short-term cortical plasticity, one may also hypothesize that the degree of modulation of such plasticity could vary not only as a function of the strength of the reinforcement signal (proportional to task difficulty), but also as a function of the number of reinforcement signals operating on a given perceptual representation. Hence, the goal of the present study was to investigate the effect of task difficulty and number of correctly identified targets (i.e. the number of reinforcement signals) on the TAE. Task difficulty is a complex psychological construct, which is not easy to define. Indeed, how difficult a task is subjectively perceived may depend on the level of motivation, the emotional state, or the type of cognitive and motor processes involved in the task at hand. However, as in most psychological studies, we operationally defined task difficulty in terms of the accuracy with which the task was performed. Therefore, two tasks, or task conditions, were considered equivalent in terms of difficulty if participants provided a comparable level of response accuracy.

Here, we adopted the paradigm we devised in our previous study (Pascucci & Turatto, 2013), and in which, by means of an adaptive procedure, we could manipulate both task difficulty and the number of targets presented and correctly identified. During the exposure to a background adapter, we presented RSVP trials in which either the same number of targets were recognized with two different levels of difficulty, or a different number of targets were recognized with the same difficulty. To anticipate our results, we found that the number of successfully recognized targets induced a consistent modulation of the TAE, whereas, at least in the present paradigm, the TAE was unaffected by task difficulty.

3.2 Experiment 1

In Experiment 1 we tested whether the difficulty of a central RSVP task modulated the degree of adaptation to an oriented grating (the adapter) presented in the background (Figure 3). Participants were asked to recognize the identity of three numbers embedded in a central RSVP of letters. The two levels of difficulty of the RSVP task (defined as easy and hard), were controlled by adjusting the luminance of the targets. At the end of the RSVP the adapter grating was removed, and a test grating tilted to the left or to the right (with various orientations) was briefly shown. The TAE was measured by asking participants to report whether the test grating was perceived as being tilted or vertical.

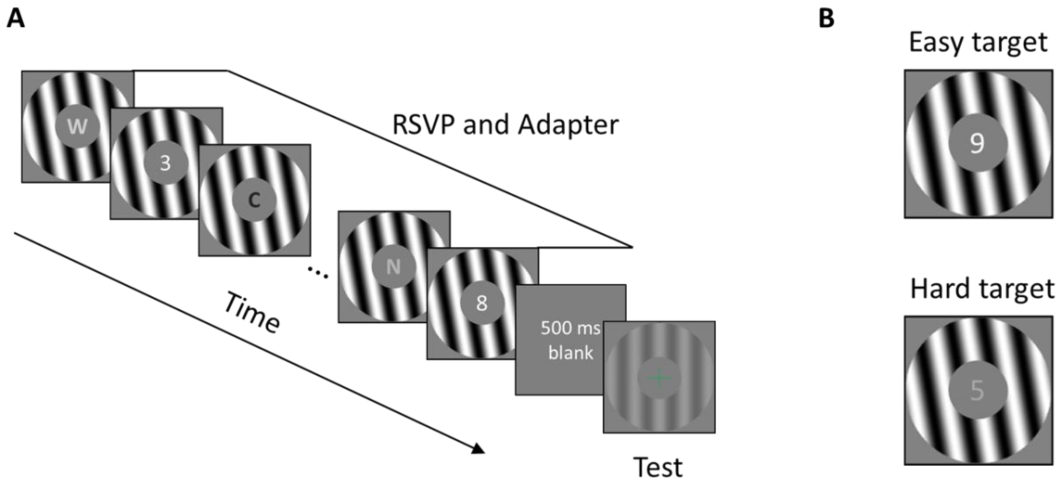

Figure 3. Methods. A) In all experiments, each trial started with the appearance of a blue fixation cross for 500 ms (not shown) followed by the presentation of the RSVP and adapter for 6000 ms. After 500 ms of blank, the test grating was presented for 50 ms. In Experiments 1 and 3, there were always three targets in the RSVP, and task difficulty was manipulated by varying the targets' luminance in two conditions (easy and hard condition). In Experiment 2, the targets were either one or three, and their luminance was adjusted in order to keep the overall accuracy at ~80% of correct responses. B) Examples of easy and hard targets presented during the RSVP (Experiments 1 and 3).

3.3 Methods

3.3.1 Participants

Twenty students (16 females; mean age = 22.7 years, SD = 7.8) from the University of Trento participated in the experiment. All participants had normal or corrected-to-normal vision and were naive as to the purpose of the research. Informed consent was obtained from all participants, and all experiments were carried out in accordance with the Declaration of Helsinki. Unless noted, this applies also to the following experiments.

3.3.2 Apparatus

Stimuli were presented on a 19-in. ViewSonic G90fB CRT monitor (1,024 × 1,768; 100 Hz). Generation and presentation of the stimuli was made using MATLAB (The MathWorks, Natick, MA) and the Psychophysics Toolbox 3.8 (Pelli, 1997) on a computer with a Windows Vista operating system.

3.3.3 Stimuli

The RSVP of digits was presented at fovea, inside a gray disk (19.8 cd/m^2, 1° visual angle in diameter) centred on a flickering (16.6 Hz) full-contrast adapter grating. In Experiment 1, the targets were three different numbers (from 2 to 9) embedded on a stream of distractor letters. The luminance of the distractors was varied randomly according to a discreet uniform distribution (± 20% of contrast from the gray background, 19.8 cd/m^2). The difficulty of the RSVP was manipulated by adjusting the luminance of the three targets with two separate non-parametric staircases, converging respectively at ~80% (easy condition) and ~55% (hard condition) of correct responses (correct recognition of three targets). The entire sequence of the RSVP lasted 50 frames (6000 ms) with a single digit presented on each frame (120 ms). The three targets could appear at any frame after the second and before the penultimate frame of the RSVP, with the only constrain that two consecutive targets were separated by a minimum of 5 frames (600 ms).

To induce the TAE we used an adapter consisting of a grating of 3° of visual angle, tilted to -10 off vertical and with spatial frequencies of 1.5 cycles per degree. The adapter was presented for the whole duration of the RSVP. After the end of the RSVP and the removal of the adapter, a low-contrast (20%) test grating of the same size and spatial frequency was shown at the same position for a brief period of time. The test grating could have one of the following orientations: −4°, −2°, 0°, 2°, or 4° from the vertical orientation.

3.3.4 Procedure

Participants sat at approximately 60 cm from the monitor, with their head positioned on a chinrest. At the beginning of each trial, a blue fixation cross was presented at the centre of the screen for 500 ms, followed by the RSVP and the adapter (lasting for 6000 ms). After a blank interval of 500 ms the test grating was shown for 50 ms, anticipated by a brief green fixation cross (100 ms). Participants had to report the whether or not the test grating appeared vertical by pressing one of two keys on the computer keyboard (yes/no task). Then a blue question mark appeared at the centre of the screen, prompting participants to report the identity of the RSVP targets by using the corresponding numbers on the keyboard. In trials where participants were not able to identify all the three targets, they were instructed to report the identity of the targets they recognized, if any. Each reported number was displayed on the screen as visual feedback: numbers shown in green indicated correct recognitions, whereas numbers shown in black wrong recognitions. The experiment consisted of 240 trials. On 96 trials (40%), the orientation of the test gratings was vertical, and on the remaining 144 trials (60%), it was tilted from the vertical orientation (36 trials for each of the four inclinations).

3.4 Results

Participants' performance in the RSVP task was better in the easy condition (78 ± 6% of correct responses) than in the hard condition (53 ± 2%, $p < 0.001$, paired t-test). The proportions of vertical responses to the test grating were arc-sin transformed and submitted to repeated measures analysis of variance (ANOVA) with task difficulty and test-grating orientation as factors. The ANOVA revealed a significant effect of task difficulty, $F(1,19) = 5.39$, $p < 0.05$, $\eta p^2 = 0.22$, test-grating orientation, $F(4,76) = 57.56$, $p < 0.001$, $\eta p^2 = 0.75$, and a significant interaction between the two factors, $F(4,76) = 3.81$, $p < 0.01$, $\eta p^2 = 0.16$. The results indicated that the TAE was modulated by the difficulty of the RSVP task performed during adaptation. However, before accepting this conclusion, one must consider an alternative explanation for the difference in the TAE between the easy and hard condition. It is possible that such difference could be accounted for by the higher number of correctly identified targets in the easy condition (with ~80% of trials where participants succeeded in recognizing three targets) with respect to the hard condition (~55%). To address this alternative interpretation we ran a separate analysis in which we collapsed the data of the easy and hard conditions, and used as factor the number of correct recognitions (3 vs. < 3). The transformed proportions of vertical responses were submitted to ANOVA with the factors number of recognitions and orientation. The results showed a significant effect of the number of recognitions, $F(1,19) = 16.78$, $p < 0.01$, $\eta p^2 = 0.47$, orientation, $F(4,76) = 56.74$, $p < 0.01$, $\eta p^2 = 0.75$, and a significant interaction, $F(4,76) = 5.43$, $p < 0.01$, $\eta p^2 = 0.22$, suggesting that the number of targets correctly identified during adaptation had a robust influence on the TAE (Figure 4). To substantiate this observation, we compared the effect size of the two factors, showing that it was larger for number of recognitions ($\eta p^2 = 0.47$) than for task difficulty ($\eta p^2 = 0.22$).

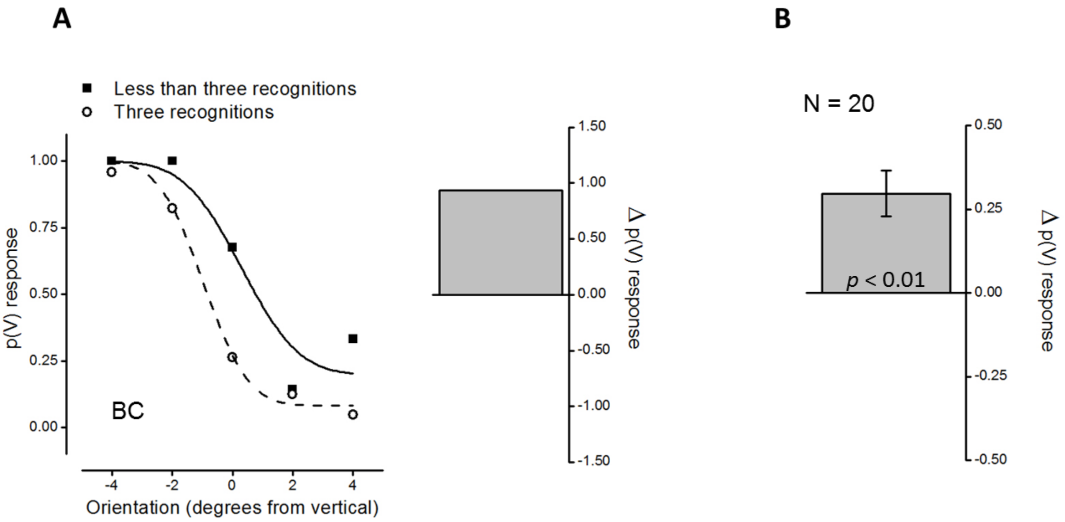

Figure 4. Results from Experiment 1. A) Data from one representative participant (BC). The proportions of vertical responses (p(V) response) are plotted as a function of the test-grating orientations and number of targets recognitions (3 vs. < 3) during adaptation. The TAE is reflected by the fact that the perceived-vertical function is shifted toward the adapter's orientation (-10° off vertical). The TAE was stronger for three targets recognitions, as indicated by the more pronounced leftward shift in this condition (open circles and dotted line) than when the recognitions were less than three (filled squares, black line). The overall magnitude of this effect is summarized by the index Δ p(V) response (gray bar histogram), computed as the sum of the differences in the p(V) response between the two conditions for each of the test-grating orientations. Positive values of this index indicate a higher leftward shift of the perceived-vertical function (i.e. stronger TAE) in the case of three recognitions. Curves are descriptive Gaussian fits of the data. B) The averaged Δ p(V) response of all participants, showing that the TAE was stronger when three targets were correctly recognized during adaptation.

This pattern of results provided a first hint of the possibility that the main factor affecting the TAE was the number of correctly identified targets rather than task difficulty. In the next experiment we addressed this hypothesis by varying the number of RSVP targets that can be recognized while maintaining the overall task difficulty constant.

3.5 Experiment 2

In Experiment 2 we presented participants with two RSVP task conditions, one target and three targets, while the difficulty of the task was kept constant for the two conditions. If the number of targets recognitions is the crucial factor that modulates the TAE, we expected a more pronounced TAE when three targets were present in the RSVP.

3.6 Methods

3.6.1 Participants

Twenty-two students (17 females; mean age = 22.7 years, SD = 6.5) from the University of Trento participated in the experiment.

3.6.2 Apparatus, stimuli and procedure

As in Experiment 1, with the following exceptions. Half of the trials contained three targets (as in Experiment 1) and the other half one target. The difficulty in the two conditions was regulated by two separate staircase converging both at ~80% of accuracy. The targets luminance was regulated according to the correct identification of either three targets or one target in the two conditions respectively. The two conditions were randomly intermixed across trials.

3.7 Results

Performance in the RSVP task did not differ between conditions (three targets, 82 ± 7%; one target, 84 ± 7%, $p > 0.05$), confirming that task difficulty was comparable with three targets and one target. To evaluate the TAE, the transformed proportions of vertical responses were submitted to an ANOVA with number of targets (three vs. one) and test-grating orientation as factors. The ANOVA revealed a significant main effect of orientation, $F(4,84) = 58.01$, $p < 0.01$, $\eta p^2 = 0.73$, which attested the overall TAE, and number of targets, $F(1,21) = 4.95$, $p < 0.05$, $\eta p^2 = 0.19$, showing an overall leftward shift of the perceived vertical orientation in the three-target condition (Figure 5).

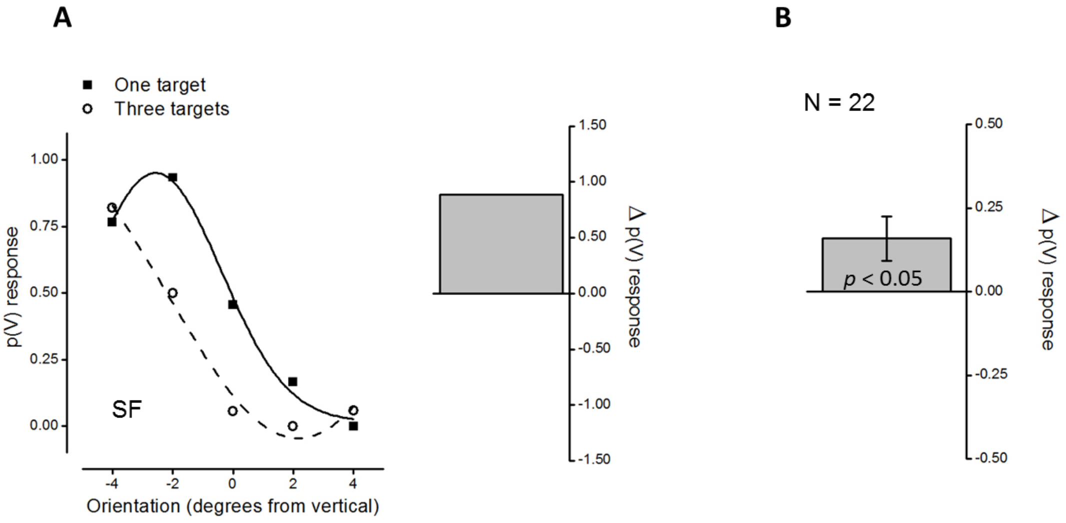

Figure 5. Results from Experiment 2. A) Data from one representative participant (SF). The p(V) response is plotted as a function of the test-grating orientations and number of targets presented in the RSVP (one vs. three). The TAE was stronger in the three-target condition (open circles, dotted line) than in the one-target condition (filled squares, black line), as indicated by the high positive value of Δ p(V) response (gray bar histogram). Curves are descriptive Gaussian fits of the data. B) The averaged Δ p(V) response of all participants showing that the TAE was stronger when three targets were presented during adaptation.

The results of Experiment 2 confirmed and extended those observed in Experiment 1, showing that the number of correctly identified targets during adaptation was the crucial factor affecting the TAE. Because the one-target condition and the three-target condition had the same level of difficulty, this suggests that it is the process target recognition that, irrespective of its difficulty, has an effect on the adapter representation. In addition, the more of this process is implemented during the RSVP, the stronger its effect on the TAE. This supports the conclusion that the results of Experiment 1 were due to the different number of correct recognitions occurred in the easy and hard conditions, rather than a difference in task difficulty between the two conditions.

However, before concluding that task difficulty does not affect the TAE, we decided to confront the TAE elicited by the same number (three) of easy- or hard-target recognitions, with a comparable number of trials in the two conditions. If task difficulty has a role in the modulation of the TAE, this comparison should be able to revel it. This analysis, however, cannot be applied to the data of Experiment 1, as the number of trials in which three targets were recognized was not comparable in the easy and hard condition. Hence, in the next and final experiment, we tested a new group of participants with the same paradigm of Experiment 1, but with a larger proportion of trials in the hard condition, to obtain a comparable number of three correct recognitions in the two conditions.

3.8 Experiment 3

3.9 Methods

3.9.1 Participants

Sixteen students (12 females; mean age = 21.8 years, SD = 1.5) from the University of Trento participated in the experiment.

3.9.2 Apparatus, stimuli and procedure

As in Experiment 1, with the following exceptions. The experiment consisted in two separate sessions, performed in two consecutive days. There were 520 trials in total, 320 for the Hard condition and 200 for the Easy condition. This resulted in 48.3 ± 7.1 and 30.1 ± 5.4 data points for the non-vertical test orientations, and 126.5 ± 9.7 and 79.3 ± 5.6 for the vertical test orientation paired with the easy and hard conditions respectively. The two conditions were randomly intermixed across trials. Data of four participants were excluded from the final analysis because they performed at chance level in the RSVP task and in the TAE test.

3.10 Results

As in Experiment 1, participants' performance in the RSVP task was better in the easy condition (83 ± 2%) than in the hard condition (55 ± 1%, $p < 0.001$). To analyse the TAE, the transformed proportions of vertical responses were submitted to a first ANOVA with task difficulty and test-grating orientation as factors. The results revealed a significant effect of task difficulty, $F(1,15) = 16.78$, $p < 0.01$, $\eta p^2 = 0.52$, and orientation, $F(4,60) = 278.62$, $p < 0.001$, $\eta p^2 = 0.95$. In the subsequent analysis, the proportions of vertical responses were calculated according to the number of (correct target) recognitions, independently of the task-difficulty condition. These data were submitted to a second ANOVA with number of recognitions (3 vs. < 3) and test-grating orientation as factors. This analysis revealed a significant effect of the number of recognitions, $F(1,15) = 32.60$, $p < 0.001$, $\eta p^2 = 0.68$, and orientation, $F(4,60) = 268.64$, $p < 0.001$, $\eta p^2 = 0.95$.

This pattern of results (Figure 6A,B) was identical to that observed in Experiment 1, and confirmed that the effect size for the factor number of recognitions ($\eta p^2 = 0.68$) in the second ANOVA, was larger than the effect size for the factor task difficulty ($\eta p^2 = 0.52$) in the first ANOVA. However, the purpose of the present experiment was to directly compare the effect of task difficulty on the TAE, with a comparable number of recognitions (3) in the two difficulty conditions. Thus, we selected only those trials in which participants provided three correct recognitions, and split these data according to the easy and hard condition. This resulted in 26.9 ± 4.3 and 25.4 ± 5.2 data points for the non-vertical test orientations, and in 69.2 ± 7.8 and 64.8 ± 6.3 for the vertical test orientation, for the easy and hard condition respectively. The transformed proportions of vertical responses for this subset of data were submitted to an ANOVA, with task difficulty and test-grating orientation as factors. The analysis revealed a significant effect of orientation, $F(4,60) = 237.37$, $p < 0.001$, $\eta p^2 = 0.94$, but neither a significant effect of task difficulty, $F(1,15) = 2.28$, $p = 0.15$, $\eta p^2 = 0.13$, nor the interaction between task difficulty and orientation, $F(4,60) = 0.14$, $p = 0.96$, $\eta p^2 = 0.01$. This confirmed that when the number of correct target recognitions was no longer a factor, task difficulty had no impact on the TAE, as originally suggested by the analysis of Experiment 1.

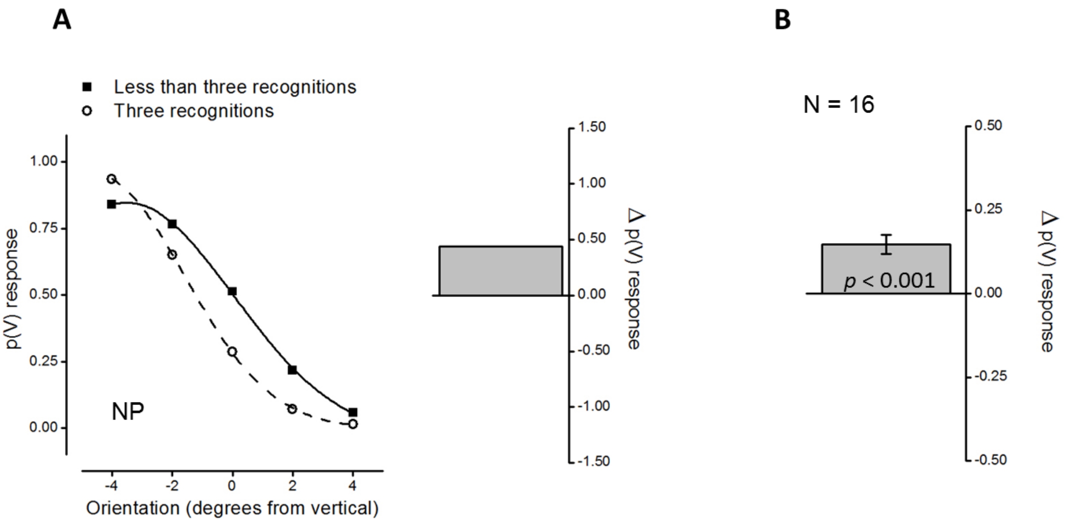

Figure 6. Results from Experiment 3. A) Data from one representative participant (NP). The p(V) response are plotted as a function of the test-grating orientations and number of targets recognitions (3 vs. < 3). As in Experiment 1, the TAE was stronger for three targets recognitions, as indicated by the high positive value of Δ p(V) response (gray bar histogram). B) The averaged Δ p(V) response of all participants showing an increase in the TAE when three targets were correctly identified.

Although the results of the present experiment confirmed that the crucial factor modulating the TAE is the number of targets correctly recognized, the proportion of trials with three recognitions was greater than the proportion of trials with less than 3 recognitions, as the former proportion results from the sum of ~80% of trials in the easy condition, plus ~55% of trials in the hard condition. By contrast, a more unbiased analysis would require to have a similar proportion of trials in the 3-target recognition and in the less-than-3-target recognitions. This condition is naturally provided by the hard task (55% accuracy), and therefore we ran a further and concluding ANOVA on this subset of data, separating the trials on the basis of the number of recognitions (3 vs. < 3). The results confirmed a significant effect of the number of targets on the TAE, $F(1,15) = 12.44$, $p < 0.01$, $\eta p^2 = 0.45$ (Figure 6C).

3.11 General Discussion

In the last decade, it has been shown that top-down attention is not the only factor regulating perceptual learning. Subliminal unattended features can also be learned, suggesting a role for a putative endogenous reinforcement signal triggered by task accomplishment (Sasaki, Náñez & Watanabe, 2012). The basic idea behind this view, stemming from TIPL studies, is that the detection or recognition of a task relevant target is, *per se*, a rewarding episode for the brain, an event that triggers the release of a reinforcement signal that consolidates, after extensive training, the sensory representations concomitantly active, thus gating long-term cortical plasticity.

We have recently shown that target recognition can also enhance a transitory visual illusion, the TAE, based on a form of short-term cortical plasticity (Pascucci & Turatto, 2013). In agreement with the theoretical view outlined above, we interpreted our previous findings as evidence that the representation of the adapter, and the following TAE, was enhanced by an endogenous-reward mechanism activated by target recognition. Because both common sense and recent evidence (Huang & Watanabe, 2012; Satterthwaite et al., 2012) suggest that it might be more rewarding to accomplish a difficult rather than easy task, here we addressed whether the magnitude of the TAE could be modulated by the difficulty of target recognition. Alternatively, we also explored the possibility that the crucial factor modulating the TAE could be the number of target recognitions. The first possibility rests on the assumption that difficult target recognitions would trigger stronger reinforcement signals than those triggered by easy target recognitions. The second possibility assumes that the more the target recognitions, the stronger the summed reinforcement signal across all instances. The results were clear in showing that the number of target recognitions was the crucial factor modulating the TAE, which instead was insensitive to task difficulty.

While we interpret of our findings within the framework of endogenous reward, one may wonder whether our results could be accounted for by the attentional-boost hypothesis, which holds that the processing resources are automatically boosted whenever the target is recognized, with the beneficial effect of the boost spilling over other sensory/memory representation paired with the target (Swallow & Jiang, 2010, 2011, 2013). Therefore, in our paradigm multiple target recognitions would have caused distinct attentional boosts, which concurred in enhancing the representation of the adapter, and thus the magnitude of the resulting TAE. Although our results are compatible with this possibility, one consideration seems to undermine the attentional-boost explanation. More precisely, the attentional-boost phenomenon has been reported in dual-task paradigms, where the background stimuli presented during the RSVP task were relevant for a secondary memory task (Swallow & Jiang, 2010, 2011). Crucially, when the same stimuli were made irrelevant by asking participants to disregard them, the attentional-boost effect disappeared (Leclerq & Seitz, 2012). This is however the scenario present in our paradigm, where attention was focused exclusively on the central RSVP, with no dual-task conditions. Indeed, the adapter was completely task irrelevant, and did not require any attentional processing during the RSVP. Yet results show a significant effect of target recognition on the TAE. For this reason, in the interpretation of our results we are more sympathetic with the reinforcement-signal hypothesis.

In following the reward-based explanation, though, we must acknowledge that our findings showing that target recognition difficulty did not affect the TAE, suggest that the putative reinforcement signal was not related to task difficulty, a result that is in contrast with other recent behavioral and neuroimaging studies providing evidence in support of such relation (Huang & Watanabe, 2012; Satterthwaite et al., 2012). At present, it is not entirely clear how this discrepancy can be reconciled, but contrasting results might not be so surprising given that very different paradigms have been adopted in the three studies, which addressed distinct visual processes (visual perceptual learning, visual adaptation, and visual short-term memory). Yet, we would like to stress a final consideration that might help explaining why, in the present paradigm, we found the TAE, and the putative reinforcement signal, to be insensitive to task difficulty.

Exogenous rewards are objects that stimulate approaching and consummatory behaviors, and that are characterized by hedonic values, as they usually elicit feelings of liking and pleasure. Obviously, different rewards, or the same reward in different moments, can be more or less salient and reinforcing for the individual’s behavior, and neurophysiological evidence suggests that stronger rewards generate greater reinforcement signals in the brain (Delgado, 2007). It is usually a common experience that accomplishing a difficult task is more rewarding than accomplishing an easier task, one would predict stronger reinforcement signals when a difficult task is successfully completed. Hence, in the present experiments one would have expected stronger TAE for difficult target recognitions (55% staircase) than for easy target recognitions (80% staircase), as the former should have been more challenging, and therefore rewarding, than the latter, but this is not what we found.

Pivotal to this view, however, is the psychological hedonic aspect of reward, and the assumption that the reinforcement signal would be proportional to the feeling of gratification for having succeeded in target recognition. However, in the RSVP we used the targets presentation, and eventually their identifications, was extremely rapid, very likely too fast to leave time to the feeling of gratification to emerge, triggering a proportional reinforcement signal. Hence, in our paradigm it seems more parsimonious to drop the issue of the reward’s hedonic value, and to hypothesize a more straightforward mechanism, operating without the mediation of positive feelings, whereby the brain directly activates the reinforcement system whenever the target is recognized. The reinforcement signal then would enhance the representation of the task-relevant target, and, incidentally, that of other sensory representations (here the adapter) paired with the target (Pascucci & Turatto, 2013; Roelfsema, van Ooyen & Watanabe, 2010; Seitz & Watanabe, 2005).

We do not deny that in some circumstances the positive feeling associated with task achievement might be an important component of endogenous reward and reinforcement signal. Indeed, one can imagine that the feeling of having been successful could have a more prominent role in paradigms where the observer' expectation of being correct is manipulated in advance. This could explain the discrepancy between our results and those of previous studies reporting that task difficulty affected TIPL (Huang & Watanabe, 2012) or the activation of the mesolimbic circuitry of reward (Satterthwaite et al., 2012). It should be noted that in both studies the easy and hard task conditions were presented on different blocks of trials, and this might have reduced participants' expectation of being correct in the "difficult" block of trials. When such expectation is low, as predicted by the reward-prediction error theory (Niv, Duff & Dayan, 2005), the internal reinforcing signal triggered by correct trials would be stronger (D'Ardenne, McClure, Nystrom & Cohen, 2008; Satterthwaite et al., 2012). By contrast, a modulation of reward expectation on the reinforcement signal was not possible in our paradigm, given that easy and hard trials were intermixed.

To conclude, our results give support to the hypothesis according to which an endogenous-reward mechanism is activated by the successful detection or recognition of behaviorally relevant events. However, when there is no expectation about the difficulty of the task at hand, this endogenous-reward mechanism would trigger a reinforcement signal in a rather stereotyped fashion whenever the target is recognized, independently from any hedonic aspects of reward associated with task accomplishment.

Chapter 4

Monetary reward modulates task-irrelevant perceptual learning for invisible stimuli

The present study has been submitted under the above title by Pascucci, D., Mastropasqua, T., and Turatto, M. to *PloS one*.

The two studies I have reported so far provided evidence of fast phenomenological changes in perceptual experience due to the recognition of behaviourally relevant events. This phenomenon is in line with TIPL and it is plausible to assume that the mechanisms underlying these forms of short- and long-term plasticity may be the same. Under this view, we discussed our results embracing the theoretical framework proposed by early literature on TIPL, according to which plasticity for irrelevant information is the consequence of internal signals of reward. However, this interpretation remained mainly speculative, as clear evidence of the involvement of reward-related processes in TIPL was still lacking.

In the two experiments presented next, we directly address the role of reward-related information in TIPL by manipulating the incentive value of visual targets recognition. We hypothesized that, if TIPL is mediated by mechanisms of reward processing, then it should be modulated by the level of reward associated with a given target.

Abstract

Task Irrelevant Perceptual Learning (TIPL) shows that the brain's discriminative capacity can improve also for invisible and unattended visual stimuli. It has been hypothesized that this form of "unconscious" neural plasticity is mediated by an endogenous reward mechanism triggered by the correct task performance. Although this result has challenged the mandatory role of attention in perceptual learning, no direct evidence exists of the hypothesized link between target recognition, reward and TIPL. Here, we manipulated the reward value associated with a target to demonstrate the involvement of reinforcement mechanisms in sensory plasticity for invisible inputs. Participants were trained in a central task associated with either high or low monetary incentives, provided only at the end of the experiment, while subliminal stimuli were presented peripherally. Our results showed that high incentive-value targets induced a greater degree of perceptual improvement for the subliminal stimuli, supporting the role of reinforcement mechanisms in TIPL.

4.1 Introduction

Perceptual learning refers to a relative long-term improvement in perceptual tasks following practice. Selective attention and active training on the relevant stimuli have always been considered two fundamental ingredients for perceptual learning to occur (Ahissar & Hochstein, 1993). In the last decade, however, the phenomenon of task-irrelevant perceptual learning (TIPL) has challenged this view, showing perceptual improvement for stimuli that are outside the current goal of the observer and that are presented below the sensory threshold (Sasaki, Náñez & Watanabe 2010). TIPL, however, does not ensue from mere exposure, but is specific for the subliminal stimuli repeatedly paired with the occurrence of task-relevant events (Seitz & Watanabe, 2003). For example, in a typical TIPL paradigm the observer must discriminate a target letter embedded in a rapid stream of distracter letters shown at fixation, while subliminal stimuli are presented peripherally. Crucially, TIPL is found only for the irrelevant stimuli paired with the targets but not for those paired with the distracters (Seitz & Watanabe, 2003). Hence, target processing appears to be fundamental for this form of perceptual learning, and indeed TIPL does not take place if the target remains unnoticed (Seitz, Lefebre, Watanabe & Jolicoeur, 2005). On the ground of this pattern of results, it has been hypothesized that TIPL occurs because the successful recognition of a task-relevant stimulus triggers an endogenous reward, which functions as an exogenous reward in reinforcement learning (Seitz & Watanabe, 2005; Roelfsema, van Ooyen & Watanabe, 2010; Pascucci & Turatto, 2013). Specifically, the reinforcement signal evoked by target detection (or identification) gives rise to the learning of the irrelevant subliminal stimuli that are synchronous with the endogenous reward.

However, although the effects of reward on learning and attention have been widely explored (Schultz, 2002; Della Libera & Chelazzi, 2006, 2009; Raymond & O'Brien, 2009; Rutherford, O'Brien, & Raymond, 2010; Anderson, Laurent, & Yantis, 2011; Dayan & Balleine, 2002), clear evidence that TIPL is driven by reward-related reinforcement mechanisms associated with target occurrence is still lacking (Seitz & Dinse, 2007).

To address this issue, we used monetary incentives to manipulate the reinforcing value of target recognition. Participants were given three days of training on a central task associated with either a high or low monetary reward, while subliminal and irrelevant stimuli were presented peripherally. We reasoned that if perceptual improvement is regulated by reinforcement mechanisms, then targets with different reinforcing values should differentially modulate the amount of task-relevant learning and, crucially, of TIPL. Although we manipulated the amount of exogenous reward associated with the target, the reward was not delivered trial by trial on the basis of the correct target identification; rather it was provided only at the end of the experiment as a function of the overall participant's performance. The reward value (high vs. low) of a given trial was indicated either by the color of the target itself (Experiment 1), or by the color of a cue preceding the target (Experiment 2). This procedure allowed us to achieve two important goals: first, to adopt a paradigm in line with previous TIPL studies in which perceptual improvement for subliminal stimuli has been shown in auto-supervised learning conditions (e.g., Watanabe et al., 2001); second, and related to the first point, by omitting the reward on a trial-by-trial basis the process of target identification was made crucial for the activation of the reinforcement system (Seitz & Watanabe, 2003), which should be modulated by the level of monetary reward associated with a given target.

4.2 General Methods

4.2.1 Participants

Fourteen (Experiment 1) and twenty (Experiment 2) participants (aged 19-28) with normal or corrected-to-normal vision were recruited from the university population and paid at the end of the of experiment according to their performance. The study was approved by the local institutional ethics committee (Comitato Etico per la Sperimentazione con l'Essere Umano, University of Trento, Italy). Written informed consent was obtained from all participants, and all the experiments were carried out in accordance with the Declaration of Helsinki.

4.2.2 Apparatus

Participants sat at approximately 60 cm in front of the monitor in a dimly illuminated room. Stimuli for the test and training phases were generated with Matlab and the Psychophysics Toolbox 3.8 (Pelli, 1997) and presented on a gamma calibrated monitor DiamondTron V2-CRT 19'' (1024 x 768, 100 Hz). Eye movements were monitored during pre- and post-test sessions with an Eyelink 1000 Tower Mount system (sampling rate: 1000 Hz; SR Research, Ontario, Canada). The apparatus was identical in the two experiments.

4.2.3 Stimuli and procedure

Each experiment is divided in three phases: a pre-test phase, a training phase, and a post-test phase.

4.2.3.1 Test phase

The pre- and post-test phases were used to estimate the participant's sensitivity to the appearance of a Gabor patch in four possible positions. The Gabor consisted of a sinusoidal grating of 5 cycles/° windowed by a Gaussian envelope of 1°.

Sensitivity in each position was tested separately in different blocks of 200 trials (three blocks in Experiment 1, and four blocks in Experiment 2) with the method of constant stimuli. In each position the Gabor had a different orientation (-60°, -30°, 30° and 60° from vertical), identical to that of the Gabor presented subliminally in the same position during the training phase (see below). The four possible positions corresponded to the corners of an imaginary square centered on the fixation point, with each corner at 3° of retinal eccentricity. The Gabor appeared inside a rectangular (14° x 20°) matrix of dynamic visual noise obtained by redrawing, every 100 ms, 80% of the pixels of the background (mean luminance of 84 cd/m^2) with random values of luminance (drawn from a uniform discrete distribution). The Gabor was embedded on the noisy background by replacing 80% of its image with the dynamic-noise matrix (see Nishina, Seitz, Kawato & Watanabe, 2007 for a similar procedure).

Each trial started with the presentation of a fixation point and the dynamic-noise matrix for 200 ms, followed by two consecutive spatial cues (outlined blue circles of 2°) presented on the position of the Gabor for 100 ms, and separated by a 400-ms interval. The Gabor was presented inside either the 1st or the 2nd spatial cue, with one of five possible contrast levels: 10%, 16%, 26%, 43%, or 70%. Participants were asked to report, in a two-interval forced choice task (2IFC), whether the Gabor appeared inside the 1st or 2nd cue.

4.2.3.2 Training phase

In the training phase, the background and the dynamic-noise matrix remained the same as in the pre- and post-test phases. However, a small squared region (2° x 2°) centred at fixation was dynamic-noise free and contained, when presented, the training target and the corresponding cue (see Figure 7).

The training target could be of two types. In the line-bisection task, three lines of 1° x 0.1° were presented inside the central noise-free region, with the two flanking lines positioned at 0.7° of eccentricity. The horizontal displacement of the central line was regulated by a 3-down/1-up staircase procedure designed to keep participants accuracy at ~80%. In the three-dot hyperacuity task, three dots (0.14° in diameter) were presented at fixation, and the vertical displacement of the central dot was regulated by the same adaptive procedure.

In Experiment 1, the training target could be either green or red, whereas in Experiment 2 it was always black and anticipated by an incentive cue consisting in an outlined-coloured (green or red) square (2° side) (see Figure 7). The training phase was conducted for three consecutive days, with daily sessions of 800 trials divided in 2 blocks. In the first block (line-bisection task) participants were asked to report the lateral offset (Left vs Right) of the central line, while in the second block (three-dot hyperacuity task) they had to report the vertical offset (Up vs Down) of the central dot.

In Experiment 1, each trial started with the presentation of the dynamic-noise area for 500 ms, followed by the appearance of the training target for 100 ms. The crucial manipulation was that together with the onset of the training target a below-threshold Gabor appeared on the screen for 500 ms. The subliminal Gabor appeared in two different positions as a function of the incentive value of the training target, and indicated by the target's colour (green or red). The contrast (12%) of the subliminal Gabor was determined through a 2IFC task in a pilot experiment.

In Experiment 2, each trial started with the presentation of the dynamic-noise area together with a green or red incentive cue, which indicated the value of the upcoming target (if presented) and remained onscreen for the whole trial duration (1400 ms). 500 ms after cue onset, the subliminal Gabor appeared in one of the four positions, and remained on screen for 900 ms. Each position was systematically paired with one of the four experimental conditions, defined by the combination of target presence and target value: High Value Target Present (HVTP), Low Value Target Present (LVTP), High Value Target Absent (HVTA) and Low Value Target Absent (LVTA). In the target-present trials, the target was presented for 100 ms and after 400 ms from the onset of the subliminal Gabor.

At the end of each block of trials participants were reminded of the final monetary reward paired with high- and low-value targets.

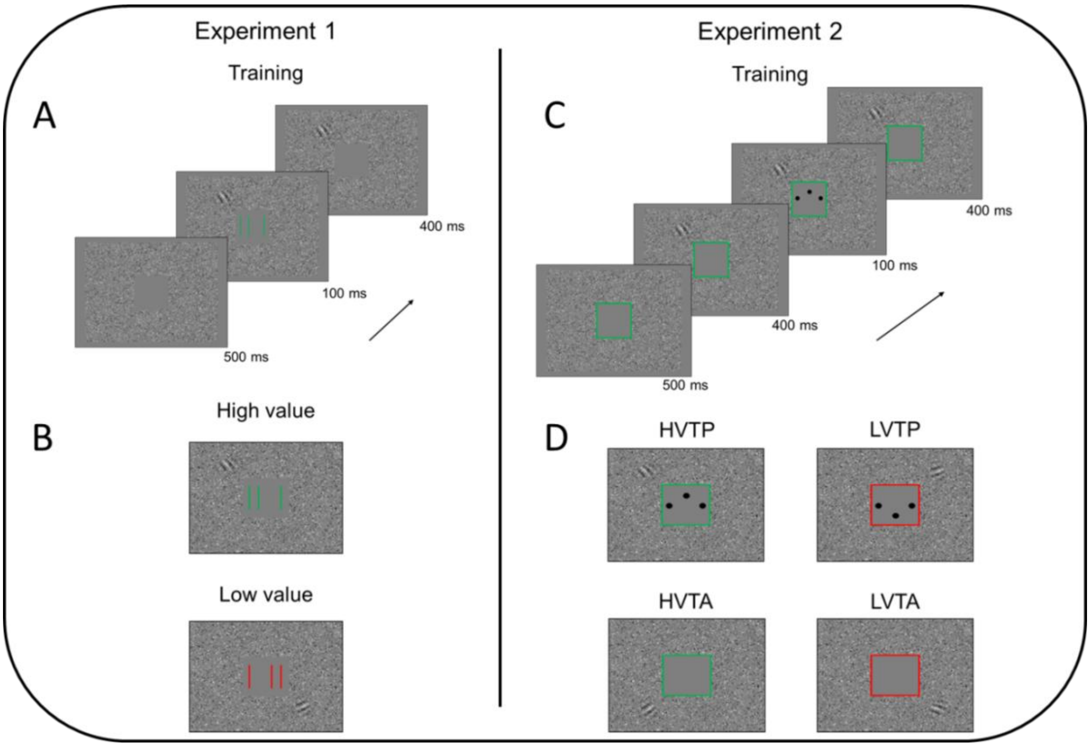

Figure 7. Schematic illustration of a training trial in Experiments 1 and 2. (*A*) Sequence of events in the line-bisection task of Experiment 1: after 500 ms of the dynamic-noise matrix the bisection stimuli and the paired Gabor appeared at the same time; bisection stimuli were presented for 100 ms, while the paired Gabor remained on screen for other 400 ms. (*B*) Example of high and low-value bisection stimuli and the paired irrelevant Gabor presented with different spatial location and orientation. (*C*) Sequence of events in the three-dot hyperacuity task of Experiment 2: after 500 ms of the dynamic-noise matrix the paired Gabor appeared for 900 ms; in the target-present trials, the three dots were presented 400 ms after the onset of the Gabor. The outlined square indicating the high or low-value condition was presented for the entire duration of the trial.

(*D*) Example of the four conditions in Experiment 2 and the corresponding irrelevant Gabor presented at four different spatial locations and with a different orientation. The peripheral high-contrast Gabor is depicted in the figures only for representation purpose; the Gabor had a below-threshold contrast (12%) during training.

4.2.4 Data analysis

Data from the central task were analysed by computing the mean of the last five reversals of each staircase procedure and then comparing the averaged means across days and tasks between the two conditions. For the analysis of the pre- and post-test data, the proportion of correct responses of each participant in each condition was fitted to a logistic function of the form:

$$P(Y=1) = \gamma + (1 - \gamma - \lambda)\frac{1}{1 + \exp\left(-\beta(x - \alpha)\right)}$$

where $P(Y=1)$is the probability of a correct answer, γ and λ are the guess and lapse rate, α is the threshold (or location parameter), β is the slope and x the contrast level of the Gabor. Parameters of the psychometric function were estimated via maximum likelihood allowing threshold and slope to vary freely and with the only constrains that the guess rate was set at 0.5 and the lapse rate was free to vary within a restricted range of values (0 – 0.08) (Wichmann & Hill, 2001).

4.3 Experiment 1

In Experiment 1 we measured TIPL for stimuli that were paired with high- and low-value targets during three days of training (Figure 7*A*). On each daily session of training, participants were asked to perform two visual discrimination tasks at fixation (see General Method). The manipulation of the incentive value of the target in each task was obtained using two different target colors (Figure 7*B*): correct identification of green target stimuli was valued 40 cents (high-value), whereas correct identification of red target stimuli was valued 4 cents (low-value). High- and low-value targets had the same probability (.5). At the beginning of training, participants were informed that their final payment was estimated on the weighted sum of the total high- and low-value trials in which they responded correctly. During training, one of two irrelevant and below-threshold Gabors, with different orientations and spatial positions (see General Method), was paired with high- and low-value targets respectively. Perceptual improvement in the detection of the irrelevant Gabor paired with either the high- or low-value target was evaluated, before and after training, with a 2IFC (see General Method). In addition, as a control condition we also tested perceptual improvement to a new Gabor occurring in a new location and with a different orientation, that was not presented during training.

4.4 Results

In the training task threshold was lower for target stimuli associated with high rather than low reward (high reward, M = 0.047° ± 0.001; low reward, M = 0.065° ± 0.001; $p < 0.001$, paired t test). To calculate the degree of TIPL for the irrelevant Gabors we calculated the participant psychometric functions (pre- and post-test, Figure 8*A*,*B*,*C*) and analyzed these results with a two-level analysis also defined as Parameter-As-Outcome Model (PAOM; Moscatelli, Mezzetti & Laquaniti, 2012) (see General Method). In the first level of the analysis we fitted the psychometric functions to the proportions of correct responses separately for each participant and condition. In the second level of analysis the individual threshold estimates (Figure 8*E*) were submitted to a repeated-measure ANOVA with Target value (High, Low and Control) and Test (Pre-Test vs. Post-Test) as factors. The ANOVA revealed a significant two-way Target value x Test interaction $F(2,26) = 5.78$, $p = 0.008$, $\eta_p^2 = 0.30$. Post-hoc analyses showed a contrast threshold reduction between pre- and post-test for the detection of the Gabors paired with high-value targets (Mean reduction = 7% ± 0.02; $p = 0.006$, paired *t* test) and low-value targets (Mean reduction = 5% ± 0.01; $p = 0.017$, paired *t* test), but this effect was not reliable in the control condition ($p = 0.11$, paired *t* test) (Figure 8*A,B,C,E*). Importantly, threshold reduction was larger in the high-value than in the low-value condition ($p = 0.04$, paired *t* test), as expected if the internal reinforcement signal that induces TIPL is proportional to the targets incentive value (Figure 8*D*).

The results of Experiment 1 were in line with those reported by previous TIPL studies showing significant learning effects for subliminal stimuli temporally paired, across multiple days of training, with relevant targets at fixation (Seitz & Watanabe, 2003). To be sure that our Gabors were below threshold during training, in the present and next experiment we used a contrast level (12%) that was smaller than the one (15%) used to estimate the absolute threshold (50% of correct response in a 2IFC task) in the pilot experiment. In addition, at the end the experiment each participant was debriefed as to whether he/she noticed any Gabor in the dynamic noise during training. None of the participants reported to have noticed the Gabor.

The novel finding was to show that the level of delayed-monetary reward associated with the target occurrence modulated the degree of TIPL. This gives strong support to the hypothesis according to which reinforcement mechanisms are responsible for the type of cortical plasticity documented by TIPL.

However, if on the one hand the results of Experiment 1 are in agreement with a role of reward in TIPL, on the other hand it is less clear what type of process associated with target occurrence triggered the reinforcing mechanism leading to TIPL. One possibility is that the strength of the reinforcement signal was determined by the mere recognition of the target feature (here, color) indicating the expected delayed-reward value in the trial, regardless of whether the target task-relevant information was correctly discriminated or not. In other words, targets with higher values may have triggered stronger reinforcement signals only by virtue of their color, which could have acted as a conditioned reward. Alternatively, the reinforcement mechanisms leading to TIPL could have been triggered only by the correct task accomplishment (Seitz & Watanabe, 2005).

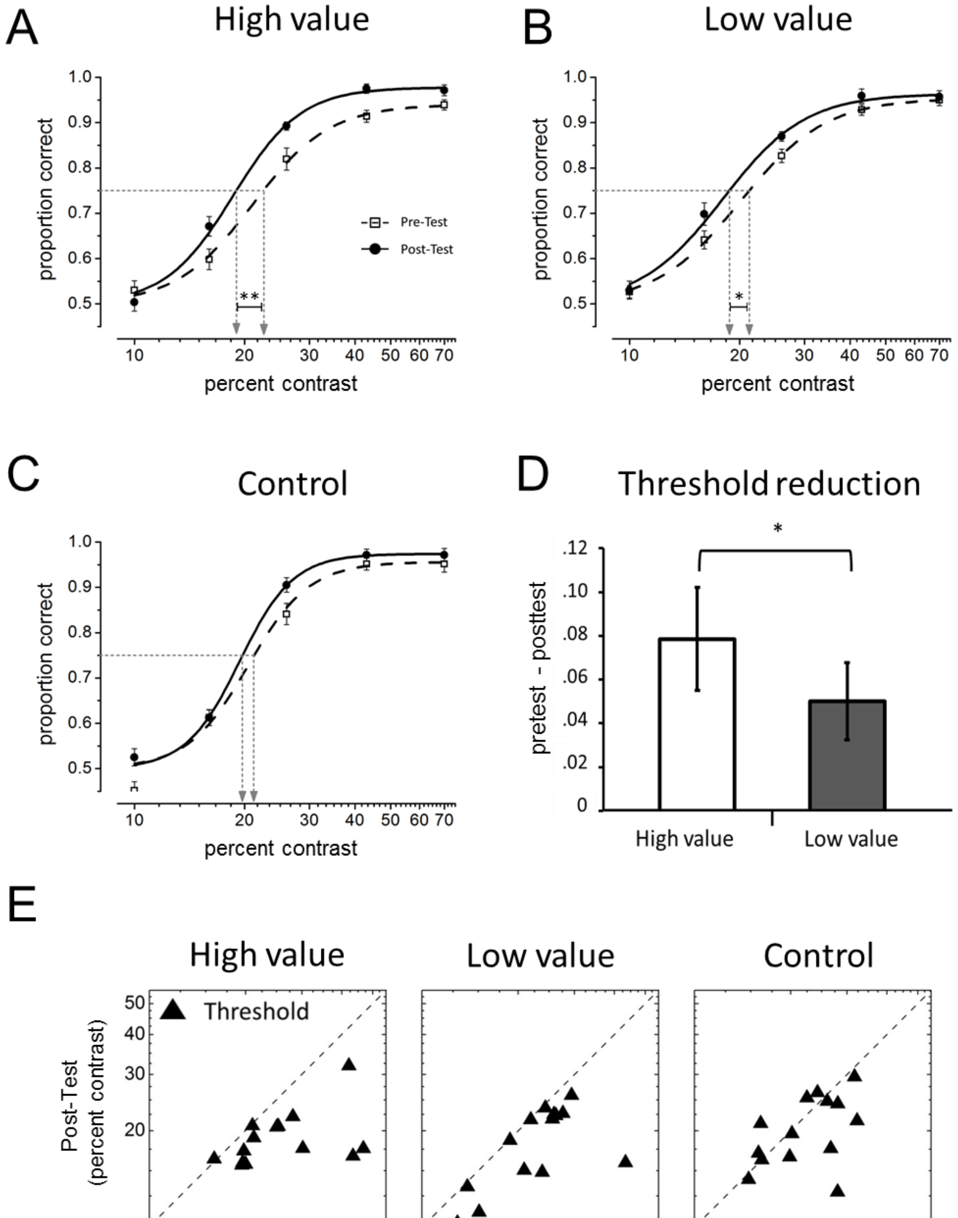

Figure 8. Results from Experiment 1. (*A*,*B*,*C*) Psychometric functions (maximum likelihood logistic function fit) obtained in the 2IFC tasks, plotting the probability of correct responses for the five levels of Gabors contrast (data averaged across participants). The points where the horizontal-dotted line intersects the curves indicate the estimated threshold (at 75% of correct responses). (*D*) Threshold reduction between pre- and post-test for the high- and low-value conditions. (*E*) Pre- and post-test distribution of the participants' thresholds in the three conditions. The results indicate that the degree of TIPL varied as a function of the target incentive value. * statistical significance at $p < .05$; ** statistical significance at $p < .01$; Error bars represent SEM.

4.5 Experiment 2

To disentangle between the two possible explanations of the results of Experiment 1, we separated the feature indicating the value of the target from the process of target discrimination. The incentive value of the upcoming target was indicated by a visual cue preceding the appearance of the target. A green-outlined square surrounding the target indicated a high-value trial (40 cents), whereas a red-outlined square indicated a low-value trial (4 cents, Figure 7*C* and *D*). The target appeared always in black and, crucially, was presented only on half of the trials. The inclusion of target-absent trials allowed us to evaluate whether the appearance of the cued value alone was sufficient to induce TIPL without the contribution of target recognition.

4.6 Results

As in Experiment 1, participants performed the relevant task with higher precision (i.e. lower thresholds) when the target was presented and was associated with high (HVTP; M = 0.054° ± 0.001) rather than low (LVTP; M = 0.070° ± 0.001) levels of reward ($p < 0.001$, paired t-test). The thresholds estimated with the PAOM procedure (Figure 9) were submitted to a repeated-measure ANOVA with Target (Present vs. Absent), Value (High vs. Low) and Test (Pre-test vs. Post-test) as factors. The ANOVA revealed a significant three-way Target x Value x Test interaction $F(1, 19) = 4.71$, $p = 0.04$, $\eta_p^2 = 0.20$. Post-hoc analyses showed a significant threshold reduction between the pre- and post-test (Mean reduction = 3% ± 0.007) in the HVTP condition ($p = 0.01$, paired t test), but no effect of training in the other conditions (Figure 9*A,B,C,D*). Hence, in Experiment 2 TIPL occurred only when the irrelevant Gabor was paired with high-value targets. Crucially, no TIPL was found in the HVTA condition, when the cue predicted a high-value target that instead was not presented. This suggests that the mere appearance of a cue indicating a possible high-value target was not sufficient to trigger the reinforcement signal. Therefore, some process related to target detection (or discrimination) was required to the release of reinforcement learning signals, which in turn fostered TIPL.

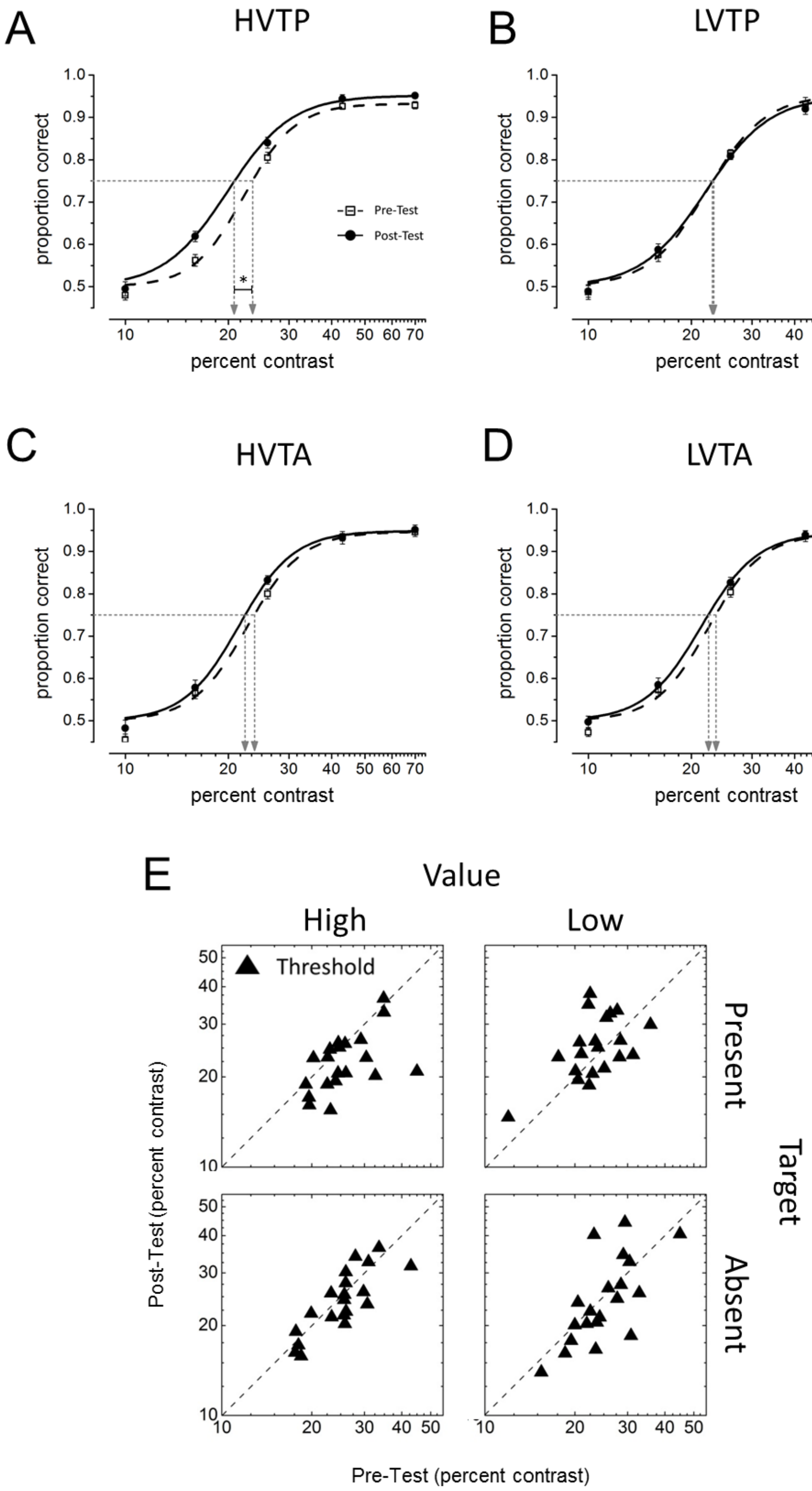
A
HVTP
B
LVTP
C
HVTA
D
LVTA
proportion correct
percent contrast
1.0
0.9
0.8
0.7
0.6
0.5
10
20
30
40
50
60
70
*
Pre-Test
Post-Test
E
Value
High
Low
Threshold
Present
Absent
Target
Post-Test (percent contrast)
Pre-Test (percent contrast)
10
20
30
40
50

Figure 9. Results from Experiment 2. (*A*,*B*,*C*,*D*) Psychometric functions (maximum likelihood Logistic function fit) obtained in the 2IFC tasks, plotting the probability of a correct response as a function of Gabor contrast (data averaged across participants). The points where the horizontal-dotted line intersects the curves indicate the estimated threshold (at 75% of correct responses). (*E*) Pre- and post-test distribution of the participants' thresholds in the four conditions. The results indicate that TIPL took place only for high-value targets, but only if the target was presented. * statistical significance at p < .02;. Error bars represent SEM.

4.7 General Discussion

Models of TIPL assume that it takes place because the detection of relevant events, such as the target appearance, works as an endogenous reward in the brain, which reacts by releasing diffuse learning signals that strengthen the representation of the subliminal stimuli (Seitz & Watanabe, 2005; Roelfsema et al., 2010).

In Experiment 1, we found that the amount of learning for task-relevant stimuli was higher when they were paired with high-value targets than with low-value targets. This is in line with theoretical models and previous findings showing that larger reward produces stronger learning (Rose, Schmidt, Grabemann & Güntürkün, 2009). Importantly, the same pattern of results emerged also for the subliminal task-irrelevant stimuli, showing that TIPL varied as a function of the reward magnitude associated with the target. This result provides experimental support to the hypothesized role of reinforcement mechanisms for task-irrelevant subliminal stimuli in auto-supervised tasks.

The idea that reinforcement mechanisms are involved in TIPL has been investigated by Seitz, Kim and Watanabe (2009). In their study, participants were presented with one of two oriented Gabor patches that were below visual threshold, and thus served as subliminal stimuli. One Gabor acted as a conditioned reinforcer for a subsequent rewarding stimulus (a drop of water delivered to the participant's mouth). Participants were asked to passively observe the stimuli without task. After several days of passive participation the authors found significant perceptual learning for the Gabor paired with the reward. They concluded that reinforcement enhanced the perceptual representation of the subliminal Gabor without the necessity of task-related attention to the stimuli.

While the presence of the final monetary reward makes our study somehow comparable to that of Seitz, Kim and Watanabe (2009), we believe the two studies to differ in some important aspects. In Seitz, Kim and Watanabe's (2009) study participants were passively exposed to the conditioned and unconditioned stimuli, whereas we asked participants to perform a central task to receive monetary reward at the end of the experiment. Therefore, in the study of Seitz and colleagues the reinforcement signals leading to TIPL was generated by a pure Pavlovian conditioning, whereas in our study reinforcement was contingent on correct task execution, bringing the procedure closer to operant or Skinnerian conditioning. In relation to this procedural difference, one could argue that despite the fact that the reward was delivered only at the end of the current experiment, the cue acted as a conditioned reinforcer because participants were informed about the contingency between the cue and outcome. However, Experiment 2 showed that TIPL did not appear when there was no central task, thus showing that in our study the reinforcement mechanisms leading to TIPL were yoked to task execution. In other words, TIPL was affected by the target value only if the target was presented and recognized (at least in the majority of trials); when the target did not appear - a condition of passive exposure more similar to the study of Seitz and colleagues - we failed to find evidence of TIPL.

We did not measure reward-related activity in the brain, and because of this one may argue that the current results do not strongly relate to investigations of neural signals evoked by the processing of reward. However the same critique could also be applied to the study of Seitz, Kim and Watanabe (2009), and, for that matter, to all the psychophysical studies on TIPL explicitly invoking reinforcement signals triggered by target recognition as the mechanism leading to learning. In addition, even electrophysiological studies directly addressing the response of dopamine neurons to reward acknowledge that the response is driven by salient stimuli (rewards) that alert the organism and attract attention (Schultz, 1998), thus showing how difficult is to disentangle the pure effect of reward-induced reinforcement from that of strategic attention (Hickey, Chelazzi & Theeuwes, 2010; Maunsell, 2004).

Overall our results are in agreement with the ‘attention-gated reinforcement learning’ model (AGREL) (Roelfsema et al., 2010), which proposes that learning occurs by trials and errors, and is modulated by the interplay between a global neuromodulatory reward-related signal and selective attention. Attention would gate learning by selectively marking the task-relevant neural representations that need to be potentiated and those that are irrelevant and must be inhibited. The AGREL model describes a form of supervised learning based on external feedback (or reward) given on a trial-by-trial basis, but reinforcement learning can take place also in unsupervised or auto-supervised manner (Karni & Bertini, 1997; Herzog & Fahle, 1998). In the latter case, the reinforcement signal inducing brain plasticity could be triggered on the basis of the observer’s self-estimation that the current task has been accomplished (Sasaki, Náñez & Watanabe, 2010), or more in general upon detection of task-relevant events (e.g., the target). A similar mechanism may have operated in our experiments, given that no feedback was provided for the central task; yet, the strength of the reinforcement signal was clearly modulated by the amount of the delayed-monetary reward.

Although the present findings are in agreement and substantiate the hypothesis that TIPL is controlled by reward-related reinforcement mechanisms triggered by target detection (or discrimination), alternative possibilities must be considered and evaluated. For example, one may wonder to what extent attention, rather than reinforcement signals, was the main determinant of TIPL in our experiments. The answer to this question is not straightforward, but depends on the type of attention one refers to. Attention, indeed, is not a unitary process, but consists of at least 3 subsystems: the alerting system, the orienting system, and the executive system (Posner & Petersen, 1990). A role of the latter form of attention can reasonably be excluded, as TIPL refers to learning of stimuli that do not require any type of response. As for a possible contribution of spatial or feature-based attention, it could be argued that on high-incentive value trials, when more attention was paid to the central target, more attention was also deployed to the peripheral subliminal stimuli, thus explaining the pattern of TIPL we observed. However, this explanation seems untenable because recent evidence has shown that directing attention to the irrelevant stimuli inhibits, rather than facilitates, TIPL (Choi, Seitz & Watanabe, 2009). Thus, orienting and executive attention subsystems do not seem to be involved in inducing TIPL (Seitz & Watanabe, 2009).

It is instead entirely possible that attention would be involved in TIPL if one considers the alerting system. According to this view, the appearance of a task-relevant stimulus activates the observer's alerting system, whose transient activation is accompanied by a diffuse reinforcement signals, possibly mediated by the norepinephrine system, that lead to TIPL (Yotsumoto & Watanabe, 2008). Hence, the non-specific reinforcement signals that control cortical plasticity for irrelevant stimuli can be activated upon target occurrence either because target detection (or discrimination) works as a reward, or because of a boost in the alerting system.

With respect to the latter possibility, the enhanced processing of information paired with the detection of relevant targets has been also reported by recent studies investigating the “attentional boost” effect (Swallow & Jiang, 2010, 2013). In these studies, participants must detect the occasional appearance of a target presented among a stream of distractors, while they simultaneously encode background images for a subsequent memory test. The results show that memory for images presented concurrently with the targets is better than memory for images paired with the distractors. The attentional boost effect assumes that a transient non-specific increase of attention due to target detection leads to enhanced processing of information also in the secondary task, without however postulating any involvement of reward-related mechanisms. Despite the similarities between the effects reported by studies on TIPL and those on attentional boost, it is still unclear to what extent the two phenomena are the same or different, and if they rely on similar brain mechanisms (Swallow & Jiang, 2013).

With regard to our findings, we acknowledge that they are compatible both with the "reinforcement-mechanism" hypothesis, and with the "attention-boost" hypothesis. However, we tend to favor the former because several features of the paradigm we used are typical of TIPL studies rather than of attentional boost studies: first, our participants performed a single task and were not invited to memorize the irrelevant stimuli; second, the irrelevant stimuli were below threshold; and third, the emergence of TIPL required several hours of training.

Finally, one may ask how the degree of motivation could explain the present results. For example, it could be hypothesized that the participants' degree of motivation in the task positively correlated with the task-incentive value, and that this modulated the level of learning observed for both the central and the peripheral stimuli. A strong argument against this account can be made on the basis of the results of Experiment 2, where TIPL was not observed when no target was presented, indicating that the mere motivational state induced by the cue was not sufficient to induce TIPL at all. Hence, although in our paradigm the motivational account can explain the task-relevant learning, it does not seem to be adequate to explain the differing degree of TIPL. The idea that reinforcement signals that lead to TIPL can also explain the perceptual learning for the relevant task strikes us as parsimonious and sufficient.

To conclude, we have shown that during auto-supervised perceptual learning task TIPL can be modulated by the incentive value of the target, a result that supports the role of reinforcement mechanisms sensitive to reward magnitude in gating visual plasticity for both visible and invisible sensory inputs.

Chapter 5

Internal reinforcement signals modulated by the “goodness” of the action that lead to reward

Pascucci, D., Jorge, J., Clayton, H., and Turatto, M.

So far, I have presented studies addressing the impact of targets recognition on visual perception. The theoretical background that we employed to interpret these results relied mainly on the hypothesis of internal signals of reinforcement or “endogenous reward” triggered by the correct detection of visual targets. We had extensively discussed this hypothesis and, particularly, we have reasoned about the concrete possibility to disentangle between endogenous and exogenous reinforcement signals. Within this framework, a recent work by Satterthwaite and colleagues (2012) suggested us that, in the absence of external reinforcers, the same neural circuit that processes exogenous reward may be activated endogenously, on the basis of a self-evaluation of task performance.

In the next chapter, I will present one study with functional magnetic resonance imaging where we try to address an interesting question: When the same external reward is obtained, is there any endogenous signal reflecting the “goodness” (precision) of the action that lead to reward?

Although the chapter is presented in the form of a research report, the data and results are preliminary, as they pertain to the first sample of an ongoing project and the analyses reported are mainly exploratory. As for this exploratory purpose, we adopted a particular method to localize reward-sensitive areas and then inspect for endogenous signals that may not represent the definitive analytical approach.

5.1 Introduction

Human and non-human animals are provided with a specific neural circuit, whose predominant role is to mediate the processing of external reward and reward-related stimuli (Schultz, 2002; Dayan & Balleine, 2002; Wise & Rompré, 1989). This network, generally defined as the “reward” or “natural reinforcement” system, recruits dopaminergic neurons that originate in the ventral tegmental area (VTA), projecting toward the nucleus accumbens (NAcc) and other regions of the ventral striatum (Haber & Knutson, 2009). Activity patterns within this complex circuit have been consistently linked to processes of external reward detection (Schultz, 2000; Apicella 2007), reward prediction error (Schultz, 2002), stimulus-reward association (Harune & Kawato, 2006; Cador, Robbins & Everitt, 1989) and reinforcement learning (Dayan & Balleine, 2002).

In recent years, neuroimaging studies have revealed a further potential property of this system that, crucially, does not depend on the availability of physical reward or reward-permeated information. The key finding is that some of the regions within the reward network show a remarkable sensitivity to the degree of correctness, self-confidence (Daniel & Pollman; 2012) and difficulty (Satterthwaite et al., 2012) with which humans can execute a task. This task-related responsiveness has been mainly described as an increase in the ventral striatum activity for correct versus incorrect responses during the execution of tasks when no external feedback or reward is provided. It has been hypothesized that, under such circumstances, “reward” may be released endogenously, on the basis of a self-evaluation of task performance (Pascucci & Turatto, 2013; Huang & Watanabe, 2012; Satterthwaite et al., 2012). In simplistic terms, these studies provided neural evidence supporting the notion that satisfying internal goals may be a rewarding event.

In the present work, we used event-related, functional magnetic resonance imaging (fMRI) to further investigate this issue. Our principal aim was to unravel brain regions, mainly within the reward circuit, whose response could resemble a more qualitative aspect of task performance, that is, the precision of task execution. To this aim, we engaged human participants in a task similar to old arcade videogames where the goal was to “shoot” and hit a particular target. Each trial could have a binary outcome (win/lose) with only positive outcomes leading to monetary reward (10 cents). Crucially, we manipulated the degree of precision with which successes or failures were obtained during the experiment. We reasoned that this particular paradigm, where the attainment of reward was regulated by a binary function (reward obtained/reward not obtained) but the precision of task execution was varied linearly, could help to shed light on endogenous reward mechanisms, unraveling reward-related areas where activity is not merely driven by the attainment of valuable stimuli, but also by a self-evaluation of the “goodness” of the action that lead to reward.

5.2 Methods

5.2.1 Participants

Participants were fourteen healthy volunteers (mean age = 25 ± 2.9, 7 female) recruited from the university of Trento and paid at the end of the experiment. All participants gave written informed consent, and the study was approved by the local institutional ethics committee (Comitato Etico per la Sperimentazione con l'Essere Umano, University of Trento, Italy).

5.2.2 Visual stimulation

Stimuli were back-projected onto a screen by a liquid-crystal projector at a frame rate of 60 Hz and a screen resolution of 1,280 × 1,024 pixels (mean luminance: 109 cd/m^2). Participants viewed the stimuli binocularly through a mirror above the head coil. Stimuli were generated with Matlab and the Psychophysics Toolbox 3.8 (Pelli, 1997).

5.2.3 Task

Participants were engaged on a visual task that resembled a classic arcade videogame. The goal was to hit the central white region (2.3 x 1°) of a black rectangle (7 x 1°) presented in the upper part of the display (at 8.5° from the center). To hit the target region, participants had to "shoot" a bullet (a red oval shape, 0.4° of diameter) from the top of a pointer (a small black rectangle, 2 x 0.5°) by pressing the key of a response box. The bullet and the pointer were non-stationary elements presented in the lower part of the display (at 10° from the center) with a continuous horizontal jitter (within ± 4° from the monitor's midline). The direction and speed of the pointer was varied randomly and the bullet was constantly sliding. After a shoot started, the trajectory of the bullet was partially occluded by the presence of a light gray rectangle (10 x 16°) that was moving on the central part of the display. Behind the occlusion, the bullet's trajectory was also artificially deviated in order to converge to a final, pre-determined position. We selected a set of five possible ending positions relative to the distance from the target's center (0±0.15°; 0.5±0.05°; 1±0.05°; 1.5±0.05°; 2.5±0.30°; see Figure 10). These positions were used to define the five levels of task precision (Center, Near Center, Near Miss, Near Win, Fail) and the two levels of reward (Reward Present [Center, Near Center, Near Miss] = 10 cents; Reward Absent [Near Win, Fail] = 0 cents).

Participants were instructed to focus on the position and speed of the pointer and to select the right moment to shoot. To maximize their involvement and motivation on the task, as well as to avoid participants to notice the pre-determined trajectory of each shoot, we informed participants that, despite the interference induced by the jittering pointer and the occluding rectangle, there was an exact moment on each trial when launching the bullet would have led to a successful outcome. At the end of the experiment, none of the participants reported of being aware of the pre-determined nature of the game.

Each trial lasted 6.6 seconds and started with 800 ms of a green fixation spot (.5°) followed by the appearance of the task-related elements. At their onset, the pointer and a dark red bullet started jittering. After one second, the bullet turned to red and the gray central rectangle started moving. Participants had one second of time from the change of the bullet’s color to shoot it toward the target. After the key press, the bullet moved toward the target and reached its final position in 1500 ms. The trajectory was visible only in its initial and final parts, while it was completely hidden in the central part by the gray rectangle. When the bullet reached the target, the three elements became stationary and the outcome of the shoot was shown for the rest of the trial. There were 50 trials on each run for a total of 250 trials. When participants failed to press the response button in time, the bullet fell from the pointer and the trial was discarded from the analysis (less than 10% of trials discarded in total). The experimental session lasted approximatively 30 minutes.

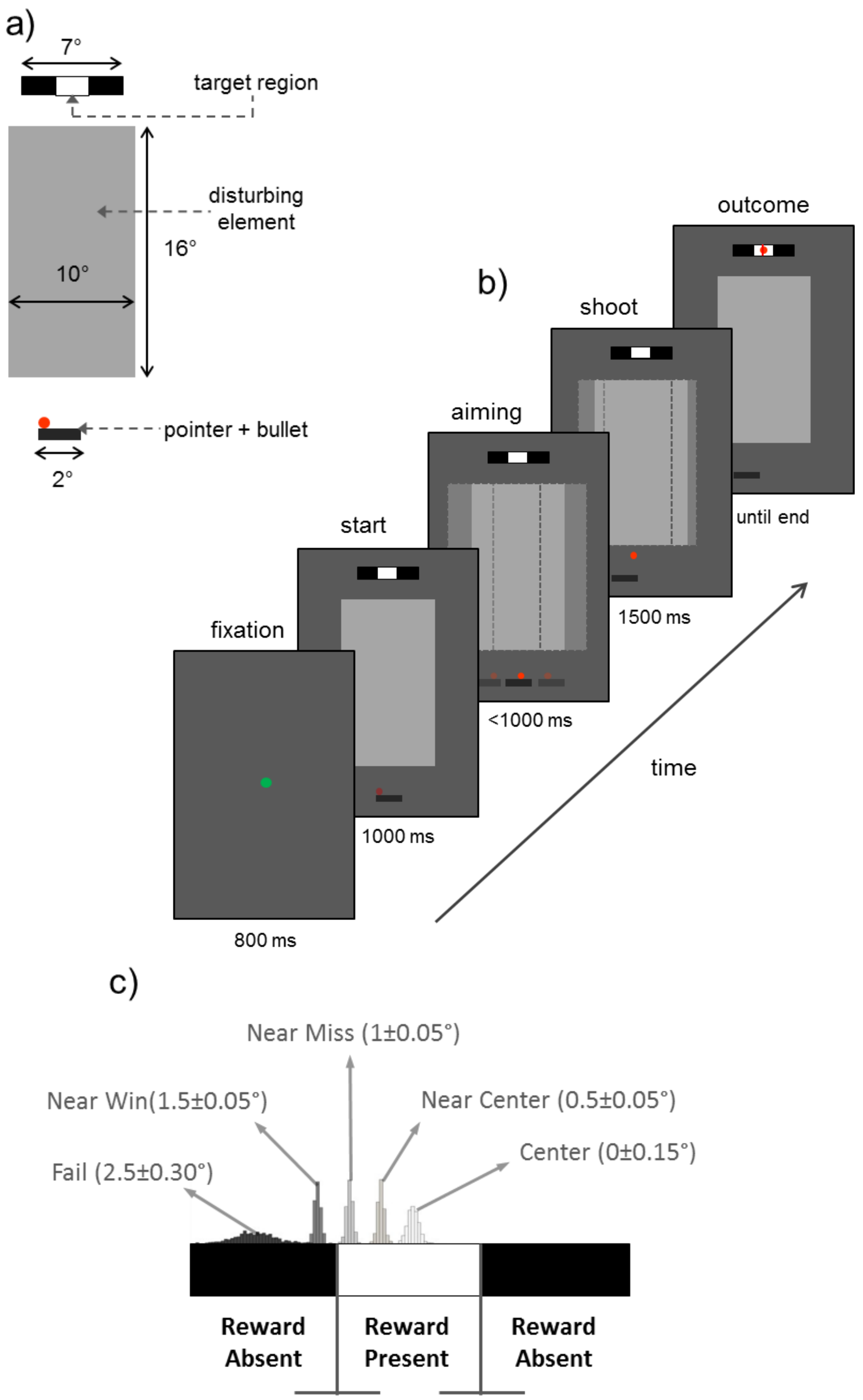
a)
7°
target region
disturbing element
16°
10°
pointer + bullet
2°
b)
fixation
800 ms
start
1000 ms
aiming
<1000 ms
shoot
1500 ms
outcome
until end
time
c)
Near Miss (1±0.05°)
Near Win(1.5±0.05°)
Near Center (0.5±0.05°)
Center (0±0.15°)
Fail (2.5±0.30°)
Reward Absent
Reward Present
Reward Absent

Figure 10. Example of the stimuli and events used in the experiment. a) The three elements that characterized the visual game: the jittering pointer and bullet were presented in the lower part of the screen (at 10° from the center) and separated from the target element (located in the upper side, at 8° from center) by a gray moving rectangle which role was to interfere and deviate the bullet's trajectory. b) A trial started with 800 ms of fixation followed by 1000 ms where the three elements were presented on screen and only the pointer (and bullet) were jittering; then central rectangle started moving and after the bullet turned to light red, participants had one second of time to "shoot" and launch the bullet toward the target item. The bullet reached the target item in 1500 ms, during which, its trajectory was obscured by the gray rectangle and artificially deviated toward a pre-selected ending position. c) Graphical representation of the pre-selected set of outcomes that we chose to achieve five levels of precision during the visual game (the outcomes outside the center where equally presented on the left or right side of the item). Note that only three outcomes lead to the monetary reward. This choice was made in order to maintain overall performances above chance (60% of successful shoots) and to obtain three Reward Present conditions where to test for a linear trend of neural activity.

5.2.4 fMRI data acquisition

fMRI images were acquired using a 4T Bruker MedSpec Biospin MR scanner and an 8-channel birdcage head coil. Each functional run consisted of 154 volumes with 32 T2*weighted echo planar slices (EPIs; repetition time (TR) = 2200 ms; time to echo (TE) = 30 ms; flip angle (FA) = 76°; field of view (FOV) = 192 x 192 mm; voxel size = 3 x 3 x 3 mm). Before the experimental session, for each participant we acquired a structural whole-head image (MP-RAGE; TR = 2700 ms; TE = 4.18 ms; FA = 7°; FOV = 256 x 224 mm; inversion time (TI) = 1020 ms; voxel size = 1 x 1 x 1 mm; slices = 176) that was used for co-registration with the functional images.

In order to increase the signal-to-noise ratio and to reduce geometric distortions in voxels located near or within the boundaries of the midbrain, we used the posterior-anterior direction of the phase-encoding gradient in the fMRI EPI sequence.

5.2.5 fMRI data preprocessing

Functional images were preprocessed and analyzed using Statistical Parametric Mapping (SPM8; University College, London). In the preprocessing stage, anatomical and functional volumes were re-oriented to the sagittal canonical plane. Functional images were then corrected for acquisition delay, realigned to the first image of the first run (for between and within runs motion correction) and co-registered to the high-resolution anatomical scan. Next, the aligned anatomical and functional images were normalized relative to the standard Montréal Neurological Institute (MNI) space using trilinear interpolation. In a final step, images were smoothed with an isotropic 6-mm full-width half-maximum Gaussian kernel and filtered with a high-pass temporal filter of 32 seconds. For the statistical analysis the first four volumes of each run were acquired in the absence of experimental task, and then discarded to allow for T1-equilibration effects.

5.2.6 fMRI data analysis

Functional images were submitted to a two-stage mixed-effects model (Friston, Holmes, Worsley, Poline, Frith & Frackowiak, 1994) with five regressors of interest, defined on the basis of the pre-determined levels of precision. The raw BOLD signal was modeled with a separate delta function for each condition (duration = 0 seconds), convolved with the canonical hemodynamic response function (HRF) and linked to the onsets of each event of interest (the outcome of the shoot on each trial). In addition, six nuisance regressors derived from the estimated motion parameters and five constant (one for each experimental run) were included in the model. Statistical *t*-maps were estimated for each participant and used to extract beta values and to compute contrasts at the single-subject level. Based on the analytic approach adopted, the estimated betas and contrasts were then submitted to second-level group analyses that consisted of voxel-wise comparisons across subjects (one-sample *t*-test) treating each subject as a random effect.

5.2.7 Definition of reward-related regions of interest (ROI)

To explore the impact of task precision on reward-related areas, we first delimited ROIs that showed the main effect of reward (clusters of voxels where the activity was higher during Reward Present than Reward Absent outcomes) and then, we locally investigate activity modulation in these ROIs as a function of task precision.

In the first step, we localized regions of interest (ROIs) where, during our particular task, the BOLD signal reflected the binary function of reward (see Figure 11). Analytically, this was obtained by contrasting (*t*-contrast analysis, Friston et al., 1994) the combined betas for the three Reward Present conditions (Center, Near Center and Near Miss) with betas of the two outcomes where reward was absent (Near Win, Fail). This subtraction procedure produced a statistical map for each participant where the main effect of reward (Reward Present>Reward Absent) was isolated. This set of individual images was then moved up to the second-level group analysis.

The group analysis revealed three separated clusters (see Figure 11) that were significantly activated (p < 0.0001, uncorrected; cluster threshold = 5 voxels) across participants, during Reward Present outcomes compared to Reward Absent outcomes. These clusters covered a portion of the right caudate nucleus and a region near and within the bilateral NAcc (as identified by the Wake Forest University PickAtlas software, http://www.fmri.wfubmc.edu/download.htm). The two regions defining the bilateral NAcc included portion of the ventral striatum and putamen, but presented a peak of intensity (see Table 1) that was close or overlapping the NAcc centre of mass localized in other studies (Demos, Heatherton & Kelley, 2012; Carter, Maclnnes, Huettel & Adcock, 2009), and we will refer to these regions as the Left and Right NAcc ROIs, for the rest of the manuscript.

In a second step, we directly addressed whether the activity in these ROIs, in addition to reflect the attainment of external reward, was also modulated by the different degree of precision. To this aim, we extracted beta values from the three ROIs and for each participant and condition separately. The resulting dataset was submitted to analyses of variance (ANOVAs), exploring significant effects with post-hoc comparisons and polynomial contrasts. In order to avoid double dipping (the same data were used to define the Reward Present/Reward Absent conditions during the ROIs definition) data of the three precision outcomes with monetary reward (Center, Near Center, Near Miss) were analysed separately from data of the two conditions with no reward (Near Win, Fail).

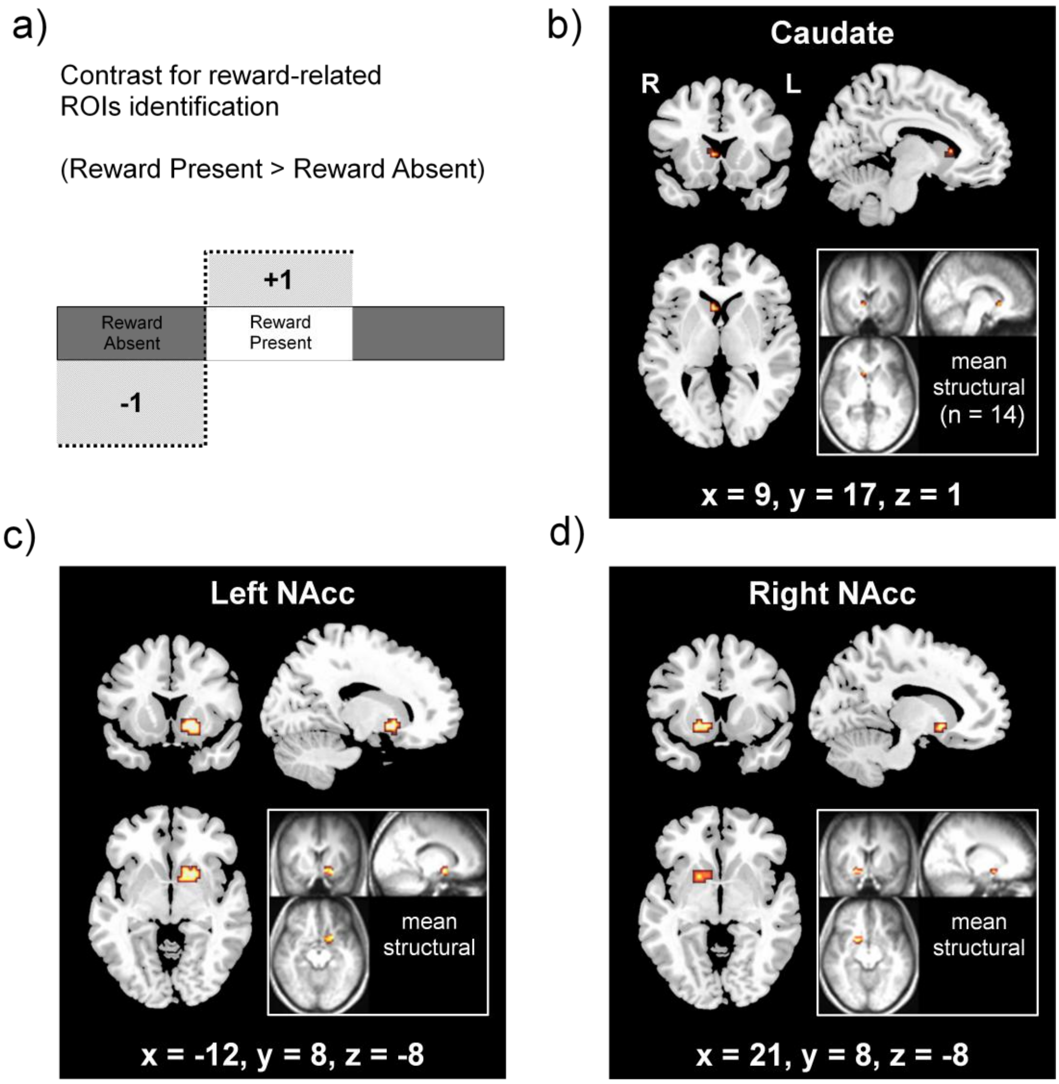

Figure 11. a) Contrast applied to the statistical map of each participant in order to localize regions more responsive during Reward Present than during Reward Absent outcomes. b), c) and d) Caudate and bilateral NAcc ROIs (see text for details). The delimited ROIs are superimposed on a standardized (MNI) atlas and on the averaged structural of fourteen participants (small white box on the bottom right corner).

Anatomical Label	Cluster size (voxels)	Hemi	Peak MNI coordinates (*x,z,y*)	*t* Value
Caudate	5	Right	9, 17, 1	6,63
Nucleus Accumbens (Nacc)	20	Right	21, 8, -8	7.48
	44	Left	-12, 8, -8	10.3

Table 1. Overview of the three ROIs delimited with the contrast Reward Present>Reward Absent (see text for details).

5.3 Results

Beta values for the three conditions of reward were submitted to repeated-measure one-way ANOVA with a single factor (Precision) and three levels (Center, Near Center, Near Miss). The analysis was carried out separately for each ROI. The two conditions with no reward outcome (Near Win, Fail) were compared by means of paired *t*-tests, whose results revealed no significant difference in any of the three ROIs (all *p*s > 0.1).

The ANOVAs on the Reward Present conditions revealed a significant main effect of Precision in the Left NAcc, $F(2,26) = 4.16$, $p < 0.05$, $\eta_p^2 = 0.24$, and a reliable trend for significance in the Right NAcc, $F(2,26) = 3.36$, $p = 0.05$, $\eta_p^2 = 0.21$. No significant effect of Precision was found in the Caudate ROI, $F(2,26) = 0.80$, $p = 0.46$, $\eta_p^2 = 0.06$. In addition, beta values for the Left and Right NAcc showed a significant linear contrast (both *p*s < 0.05), with evidence of a graded increase in the BOLD amplitude as a function of precision. Such linear trend was validated by the significant difference between betas in the Near Miss and in the Near Center/Center conditions (Near Win versus Near Center, $p < 0.05$;

Near Win versus Center, *p* < 0.01; one-tailed *t*-test), as resulted from post-hoc comparisons in both the Right and Left NAcc.

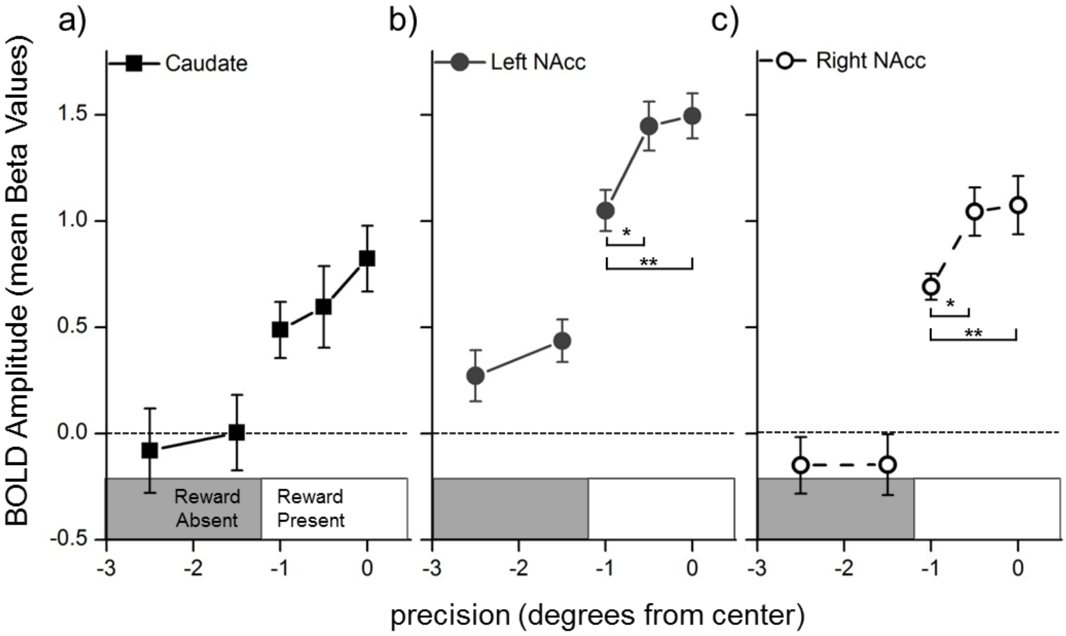

Figure 12. Averaged BOLD Amplitude for the Caudate (a), Left NAcc (b) and Right NAcc (c) ROI as a function of task precision. All the three regions show a consistent increase of activity for Reward Present outcomes (highlighted by the white rectangle in the lower right) than for Reward Absent outcomes (indicated by the gray rectangle in the lower left of each plot). Interestingly, the ROIs including the bilateral NAcc present significant differences in activity between the Reward Present outcome with the lower degree of precision and the other two, more precise outcomes. *, p < 0.05; **, p < 0.01. Error bars represent standard errors of the mean (corrected for within-subject design).

5.4 Discussion

5.4.1 Main effect of reward

We observed that during the execution of a visual game activity in the bilateral NAcc and caudate nucleus is higher for trials leading to monetary reward than for non-rewarded trials. This result is in line with the robust activation of mesolimbic projection regions (including the caudate nucleus and NAcc) observed when humans are engaged on interactive gameplays (Koepp, Gunn, Lawrence, Cunningham, Dagher et al., 1998). Furthermore, this pattern of activation is consistent with the involvement of the NAcc in processes related to reward anticipation and detection, as reported by previous imaging studies. Knutson, Adams, Fong and Hommer (2001), for example, demonstrated that the NAcc plays a selective role in the anticipation of reward versus punishment and that activity within this nucleus correlates with the feeling of “happiness” elicited by reward-predictive cues. This evidence may suggest a direct involvement of the NAcc in mediating hedonic aspects of reward, such as the sense of achievement and satisfaction that derives from goal attainment. In line with this hypothesis, several studies have identified an opioid substrate (Peciña, 2008; Smith & Berridge; 2007), the so-called “hedonic hot spot”, in the medial shell of the NAcc, that may represent a likely candidate for internal signals of “liking” and reward-related gratification (Smith & Berridge; 2007). In following this hedonic account, the higher activity in the NAcc that we observed for Reward Present outcomes may be interpreted as an immediate hedonic reaction to the attainment of valuable goals. However, a detailed analysis of the NAcc functionalities may suggest a further interpretation.

According to a recent reconceptualization of the NAcc's anatomy (Kelley, 2003; Salamone, Correa, Mingote and Weber, 2005), this nucleus is composed by two sub-regions, the core and the shell, which seems to have distinct functionalities. On the one hand, the NAcc core, surrounding the anterior commissure and projecting mainly to the basal ganglia, seems to be involved in learning and voluntary actions execution in context of adaptive instrumental behaviour. On the other hand, the NAcc shell, extending ventrally around the core and projecting to subcortical nuclei (including VTA), may instead have a critical role in appetitive and motivational behaviour.

This anatomical and functional distinction suggests that diffuse activity within this region, as the one we observed, may correspond to processes of instrumental learning, such as the developing of associations between actions and their outcomes, and to processes related to appetitive and motivational ("wanting") behavior when rewarding events occur. Crucially, both possibilities are plausible in our task, since participants were constantly trying to "learn" the game by inferring the causality of their responses, and, at the same time, successful outcomes represented actual cues for a final reward, acting as potential motivational attractors.

To summarize, on the basis of the known functionalities of the NAcc, the "main effect" of reward that we observed could have been driven by several distinct aspects of task execution, from instrumental learning to motivational or hedonic reactions to the final reward.

5.4.2 The impact of task precision on NAcc

The novelty of our result is that activity in NAcc was modulated by the precision of task execution. This observation is substantiated by the fact that, in addition to the main effect of reward, activity patterns in the bilateral NAcc showed reliable differences as a function of the distance of the bullet from the center of the target. Interestingly, this result could help to shed light on the actual role of NAcc in circumstances as the one proposed by our paradigm.

According to the main roles of the NAcc described above, the motivational "wanting" account would favor the prediction of a pure binary response in this region, by which activity is modulated exclusively by the presence or absence of reward. By contrast, the modulation of the NAcc activity as a function of the precision or "goodness" with which reward was attained, seems to suggest the involvement of different processes, that can be potentially explained by both the instrumental learning and the "liking" account.

As introduced above, the instrumental learning account proposed that activity in the NAcc underlies critical aspects of instrumental behavior, from aiding the ability of an organism to overcome obstacles, to foster learning and performance improvements during reward-seeking behavior (Salamone et al., 2005). Under this perspective, our results may indicate that, even when external reward is obtained, an internal mechanism constantly evaluates the "goodness" of performance and reinforces behaviors proportionally to their precision. In other terms, the NAcc may have a role in mediating endogenous signals that maximize the likelihood of “precise” behaviors in the future.

A different interpretation, on the other hand, can be drawn if one refers to the more hedonic impact of “being precise”. In following this rationale, it is plausible to assume that during the game participants were experiencing a different sense of accomplishment for the precision of each outcome, and thus, the observed activity in NAcc could resemble endogenous signals of “liking”.

Whether the modulation of NAcc activity derives from a positive feeling associated with task precision or from a more mechanistic process of reinforcement, cannot be inferred from the current state of our research, however, what seems to be an interesting evidence is that the NAcc may be crucial in driving internal signals related to the execution of a task.

A similar endogenous mechanism engaged by task-related episodes has been already hypothesized by previous research (Pascucci & Turatto, 2013; Huang & Watanabe, 2012; Seitz & Watanabe, 2003; Satterthwaite et al., 2012; Daniel & Pollmann, 2012). In one of these studies, for example, Satterthwaite and colleagues (Satterthwaite et al., 2012), observed an increased activity in the ventral striatum for correct versus incorrect responses while humans were engaged on a visual memory task without external reward. Interestingly, across the large sample of participants that they tested, they found that the correct execution of a difficult task activated the striatum more that the correct execution of an easy one. This results lead the authors to suggest that, in the absence of external valuable feedback, humans are provided with an endogenous system that evaluates task performance and release reinforcement signals according to the correctness of task execution.

Our results represent a further step toward the definition of endogenous reinforcement signals, indicating that not only correct responses are reinforced internally but also a more fine-grained aspect of performance, that is, the precision of execution. In addition, it is worth to notice, that contrary to previous work on this topic (Satterthwaite et al., 2012), in the present study external (monetary) reward was effectively paired with successful outcomes, and thus, our paradigm represents the first attempt to isolate and address endogenous modulations of reinforcement signal.

5.5 Conclusion

In the present study, we provided preliminary evidence that neural activity within the NAcc can be driven by the precision of a task execution. The paradigm we have adopted, allowed us to contrast the pure impact of an obtained reward from a more qualitative aspect of task performance. Interestingly, we found that in addition to be overall more responsive during reward outcomes, regions of the ventral striatum and NAcc showed a reliable sensitivity to the goodness of the event that lead to reward. This result supports the existence of an internal mechanism of reinforcement that drives plasticity in a weighted fashion, promoting the strengthening of actions or episodes where the desired goal was achieved with higher precision.

5.6 Open issues

A further ROI that we found active during Reward Present outcomes included a group of voxels within the head of the right caudate nucleus. This region showed significant activity modulation as a function of reward but not of precision (although the graphical description of the results may suggest a linear trend). Activation of this particular portion of the caudate has been associated with feedback processing (Seger & Cincotta, 2005) and with the automatic guidance of saccades toward stimuli that indicate immediate reward outcomes (Hikosaka, Kim, Yasuda & Yamamoto, 2014). It is therefore plausible that the pattern of activity that we detected in the head of the caudate could have been induced by the presentation of positive performance feedbacks (the ending position of the bullet) or by a differential drive in saccadic movements towards the winning or losing outcomes. Because of these particular properties of the caudate nucleus, and because of the partially different pattern of results with respect to the NAcc, we have focused our discussion on the bilateral and more robust activation observed in the NAcc.

However, further data collection and analysis will help to clarify the role of this region as well as to understand the lack of any reliable activation for other reward-related structures, such as VTA, in our paradigm.

Chapter 6

Summary

In the last decade, there has been growing interest in the role of task-related processes in shaping humans perceptual and cognitive abilities. One of the main findings is that stimuli conveying no useful information for a task, can be learned or memorized when they are extensively paired with task performance. In the present work I further investigated this issue, by addressing the impact of task-related episodes on short- and long-term visual plasticity.

Early research on this topic has demonstrated that the repeated pairing of irrelevant stimuli with task-relevant events can lead to long-term changes in the sensory representation of the irrelevant items. In a first study, we asked whether a similar form of target-elicited plasticity could be observed on the frame of a single trial and, to address this question, we developed a new paradigm. Our approach relied on the hypothesis that recognizing targets would strengthen the transitory processing of irrelevant stimuli (Seitz & Watanabe, 2003). From this assumption, we examined whether the TAE, a visual illusion caused by the prolonged exposure (adaptation) to tilted images, could be magnified by the occurrence (and correct identification) of visual targets during adaptation. The results of these experiments, presented in *Chapter 2,* confirmed our prediction, providing converging evidence of the effect of targets recognition on visual plasticity. Furthermore, we reported first evidence of an immediate phenomenological change in perception triggered by the occurrence of relevant events. In line with previous research, the effect that we reported was restricted to stimuli (the adapters) presented at the same time of targets and the overall illusion was less pronounced when targets appeared after the adapter was removed.

The study reported in *Chapter 3* investigated two factors that may be determinant when irrelevant stimuli are paired with task-relevant events, namely, the difficulty of the task and the number of targets. In a series of experiments we adopted the same paradigm as in our previous study (Pascucci & Turatto, 2013), but we varied the number of RSVP targets and we manipulated the difficulty of the RSVP. We found that the number of targets recognition was the crucial factor modulating the TAE, which was insensitive to task difficulty. This pattern of results was in contrast with other recent evidence showing that difficult tasks evoke higher TIPL (Huang & Watanabe, 2012), a discrepancy that, however, could reveal something significant about the mechanism behind. We proposed that the expectation of “being correct”, as the expectation of reward in classical models of reinforcement learning, could play a crucial role. When such expectation is low, as in difficult blocks of trials, the internal reinforcing signal triggered by targets recognition would be stronger, as in Huang and Watanabe (2012). By contrast, when task difficulty is unpredictable, as in our interleaved procedure, modulation of endogenous signals due to expectation is not possible. On the other hand, the fact that the number of targets recognized, independently of their difficulty, modulated TAE, suggested that, when there is no expectation about the probability of a correct response, the internal reinforcement mechanism may act in a rather stereotyped fashion, causing perceptual boosts whenever a target is recognized.

Throughout the studies presented above, we interpreted and discussed our results mainly relying on the hypothesis of endogenous mechanisms of reward. To briefly recall this account, several studies on TIPL postulated that the perceptual effect induced by targets recognition may be the consequence of internal signals of reward for task accomplishment. Although the complex interplay between task difficulty and expectation seems to support the involvement of reinforcement processes, the reward account remained mainly speculative.

In the study presented in *Chapter 4*, we directly address the pivotal role of reward-related mechanisms in TIPL by manipulating the incentive value of visual targets recognition. We reasoned that, if reward processing is a key component of TIPL, then perceptual changes associated with the recognition of targets may show differential patterns as a function of the targets incentive value. This work included two important manipulations that we applied to a canonical paradigm of TIPL. In the first experiment, we assigned high or low incentive value (corresponding to the amount of a final monetary reward) to the recognition of targets with different colours. In the second experiment, we indicated the incentive value of each trial with a visual cue around the targets location and we presented targets only on half of the trials. This allowed us to disentangle potential effects induced by motivational cues predicting reward (as in classical conditioning) from the role of target-triggered endogenous signals. Interestingly, we found that high incentive-value targets induced greater TIPL, whereas learning was prevented for high-incentive trials on which targets were absent. This result was in agreement with what reported by previous work on TIPL, confirming that targets need to be successfully identified in order to evoke TIPL whereas stimuli that do not require a response (e.g., visual cues indicating monetary reward) fail to induce learning. Furthermore, our study revealed a new important aspect of this phenomenon, that is, the presence of a final reward for task accomplishment can differentially modulate TIPL.

With the aim to gain a deeper understanding of how endogenous signals of reward are modulated by task execution, in the exploratory study presented in *Chapter 5*, we moved to a different approach and measure. This study posed an interesting question: When the same external reward is obtained, is there any endogenous signal reflecting the “goodness” of the action that lead to reward?

To address this question we measured haemodynamic changes with functional magnetic imaging as human subjects were engaged on a particular visual game. The task resembled a simple videogame where the goal was to launch a “bullet” toward the centre of a target element in order to gain a fixed amount of monetary reward. Crucially, we manipulated the degree of precision with which successes (reward) or failures (no reward) were collected. The idea behind this study was first to delimitate brain regions for their predominant responsiveness to reward (present/absent), and then to inspect locally the activity in these regions as a function of different degrees of precision. Our preliminary results indicated a crucial role for a specific subcortical structure, namely the nucleus accumbens, in the processing of reward outcomes during our task, a result that is consistent with the well-known sensitivity of this nucleus to reward-related information. However, the novelty of our results is that activity in this regions seems to be further modulated by the goodness (precision) of the event that lead to reward, supporting the existence of an internal mechanisms of reinforcement that may drive plasticity in a weighted fashion, promoting the strengthening of actions or episodes where the desired goal was achieved with higher precision.

The picture that emerged from this collection of studies provides supporting evidence of perceptual changes at the time of behaviourally relevant events. Our results contribute to the literature on this topic by showing that targets recognition can also evoke immediate phenomenological changes in the perception of irrelevant stimuli (*Chapter 2-3*).

Our theoretical approach focused on the hypothesis of internal reinforcement signals triggered by task-related events. In following this view, we reported evidence of reward-related mechanisms, engaged by the execution of a task (*Chapter 3-4-5*), whose role could be to regulate target-triggered signals on the basis of reward expectation and reward magnitude. We speculate that a potential candidate for mediating these endogenous signals could be the nucleus accumbens (*Chapter 5*) where activity seems to be modulated by intrinsic aspects of task performance.

An interesting question, that arises from this work and that deserves further investigation, is whether this endogenous activity results from hedonic aspects linked to goal attainment, or represents the expression of a more straightforward mechanism of reinforcement that constantly fosters instrumental learning without the mediation of positive or motivational aspects of task accomplishment.

References

Ahissar, M., & Hochstein, S. (1993). Attentional control of early perceptual learning. *Proceeding of the National Academy of Science USA, 90*, 5718-22.

Ainslie, G. (2013). Grasping the impalpable: the role of endogenous reward in choices, including process addictions. *Inquiry, 56*, 446-469.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences, 108*, 10367-10371.

Apicella, P. (2007). Leading tonically active neurons of the striatum from reward detection to context recognition. *Trends in Neurosciences*, *30*, 299-306.

Arsenault, J. T., Nelissen, K., Jarraya, B., & Vanduffel, W. (2013). Dopaminergic reward signals selectively decrease fMRI activity in primate visual cortex. *Neuron*, *77*, 1174-1186.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403-450.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology: Section A, 49*, 5-28.

Bao, S., Chan, V. T., & Merzenich, M. M. (2001). Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature*, *412*, 79-83.

Cador, M., Robbins, T. W., & Everitt, B. J. (1989). Involvement of the amygdala in stimulus-reward associations: interaction with the ventral striatum. *Neuroscience*, *30*, 77-86.

Carter, R. M., MacInnes, J. J., Huettel, S. A., & Adcock, R. A. (2009). Activation in the VTA and nucleus accumbens increases in anticipation of both gains and losses. *Frontiers in Behavioral Neuroscience*, *3*, 21.

Choi, H., Seitz, A. R., & Watanabe, T. (2009). When attention interrupts learning: inhibitory effects of attention on TIPL. *Vision Research, 49*, 2586-2590.

Corsini, R. J. (1999). The dictionary of psychology. *Philadelphia: Brunner/Mazel.*

Daniel, R., & Pollmann, S. (2012). Striatal activations signal prediction errors on confidence in the absence of external feedback. *Neuroimage, 59*(4), 3457-3467.

D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science, 319*(5867), 1264-1267.

Dayan, P., & Balleine, B. W. (2002). Reward, Motivation and Reinforcement Learning. *Neuron, 36*, 285-298.

Delgado, M. R. (2007). Reward Related Responses in the Human Striatum. *Annals of the New York Academy of Sciences, 1104*(1), 70-88.

Della Libera, C., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary rewards. *Psychological Science, 17*(3), 222-227.

Della Libera, C., & Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychological Science, 20*(6), 778-784.

Demos, K. E., Heatherton, T. F., & Kelley, W. M. (2012). Individual differences in nucleus accumbens activity to food and sexual images predict weight gain and sexual behavior. *The Journal of Neuroscience, 32*(16), 5549-5552.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193-222.

Dodge, R. (1917). The laws of relative fatigue. *Psychological Review*, *24*(2), 89.

Dragoi, V., Sharma, J., & Sur, M. (2000). Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron*, *28*(1), 287-298.

Frankó, E., Seitz, A. R., & Vogels, R. (2010). Dissociable Neural Effects of Long-term Stimulus–Reward Pairing in Macaque Visual Cortex. *Journal of Cognitive Neuroscience*, *22*(7), 1425-1439.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping*, *2*(4), 189-210.

Gibson, J., & Radner, J. M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines. I. Quantitative studies. *Journal of Experimental Psychology*, *20*, 453–467.

Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron*, *31*(5), 681-697.

Haber, S. N., & Knutson, B. (2009). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*, *35*(1), 4-26.

Haruno, M., & Kawato, M. (2006). Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *Journal of Neurophysiology*, *95*(2), 948-959.

Herzog, M. H., & Fahle, M. (1998). Modeling perceptual learning: difficulties and how they can be overcome. *Biological Cybernetics, 78,* 107-117.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *The Journal of Neuroscience*, *30*(33), 11096-11103.

Hikosaka, O., Kim, H. F., Yasuda, M., & Yamamoto, S. (2014). Basal Ganglia Circuits for Reward Value-Guided Behavior. *Annual Review of Neuroscience*, *37*(1).

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679.

Huang, T. R., & Watanabe, T. (2012). Task attention facilitates learning of task-irrelevant stimuli. *PloS one*, *7*(4), e35946.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, *43*(4), 346-354.

Karni, A., & Bertini, G. (1997). Learning perceptual skills: behavioral probes into adult cortical plasticity. *Current Opinion in Neurobiology*, *7*(4), 530-535.

Kiss, M., Driver, J., & Eimer, M. (2009). Reward priority of visual target singletons modulates event-related potential signatures of attentional selection. *Psychological Science*, *20*, 245–251.

Koepp MJ, Gunn RN, Lawrence AD, Cunningham VJ, Dagher A, et al. (1998). Evidence for striatal dopamine release during a video game. *Nature*, 393: 266–268.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive science*s, *9*(2), 75-82.

Leclercq, V., & Seitz, A. R. (2012a). The impact of orienting attention in fast task-irrelevant perceptual learning. *Attention, Perception, & Psychophysics*, *74*(4), 648-660.

Leclercq, V., & Seitz, A. R. (2012b). Fast task-irrelevant perceptual learning is disrupted by sudden onset of central task elements. *Vision Research*, *61*, 70-76.

Leclercq, V., & Seitz, A. R. (2012c). Enhancement from targets and suppression from cues in fast task-irrelevant perceptual learning. *Acta Psychologica*, *141*(1), 31-38.

Leclercq, V., & Seitz, A. R. (2012d). Fast-TIPL occurs for salient images without a memorization requirement in men but not in women. *PloS one*, *7*(4), e36228.

Maunsell, J. H. (2004). Neuronal representations of cognitive state: reward or attention? *Trends in Cognitive Science, 8,* 261-265.

Moscatelli, A., Mezzetti, M., & Laquaniti, F. (2012). Modeling psychophysical data at the population level: The generalized linear mixed model. *Journal of Vision*, *12*, 1–17.

Nishina, S., Seitz, A. R., Kawato, M., & Watanabe, T. (2007). Effect of spatial distance to the task stimulus on task-irrelevant perceptual learning of static Gabors. *Journal of Vision, 7,* 1-10.

Niv, Y., Duff, M. O., & Dayan, P. (2005). Dopamine, uncertainty and TD learning. *Behavioral and Brain Functions, 1*(6), 1-9.

Olivers, C. N., & Meeter, M. (2008). A boost and bounce theory of temporal attention. *Psychological Review*, *115*(4), 836.

Pascucci, D., & Turatto, M. (2013). Immediate effect of internal reward on visual adaptation. *Psychological Science*, *24*(7), 1317-1322.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological Bulletin*, *116*(2), 220.

Peciña, S. (2008). Opioid reward 'liking'and 'wanting'in the nucleus accumbens. *Physiology & Behavior*, *94*(5), 675-680.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437-442.

Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R. J., & Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron*, *59*(4), 561-567.

Pilly, P. K., Grossberg, S., & Seitz, A. R. (2010). Low-level sensory plasticity during task-irrelevant perceptual learning: Evidence from conventional and double training procedures. *Vision Research, 50,* 424–432.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13,* 25–42.

Raymond, J. E., & O'Brien, J. L. (2009). Selective Visual Attention and Motivation The Consequences of Value Learning in an Attentional Blink Task. *Psychological Science, 20*(8), 981-988.

Reader, T. A. (1978). The effects of dopamine, noradrenaline and serotonin in the visual cortex of the cat. *Experientia, 34*(12), 1586-1588.

Roelfsema, P. R., van Ooyen, A., & Watanabe, T. (2010). Perceptual learning rules based on reinforcers and attention. *Trends in Cognitive Sciences, 14*(2), 64-71.

Rose, J., Schmidt, R., Grabemann, M., & Güntürkün, O. (2009). Theory meets pigeons: the influence of reward-magnitude on discrimination-learning. *Behavioral Brain Research, 198,* 125-129.

Rutherford, H. J. V., O'Brien, J. L., & Raymond, J. E. (2010). Value associations of irrelevant stimuli modify rapid visual orienting. *Psychological Bulletin & Review***,** *17,* 536-542**.**

Salamone, J. D., Correa, M., Mingote, S. M., & Weber, S. M. (2005). Beyond the reward hypothesis: alternative functions of nucleus accumbens dopamine. *Current Opinion in Pharmacology*, *5*(1), 34-41.

Sasaki, Y., Náñez, J. E., & Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nature Reviews Neuroscience, 11*, 53–60.

Sasaki, Y., Náñez, J. E., & Watanabe, T. (2012). Recent progress in perceptual learning research. *Wiley Interdisciplinary Reviews: Cognitive Science, 3*(3), 293-299.

Satterthwaite, T. D., Ruparel, K., Loughead, J., Elliott, M. A., Gerraty, R. T., et al. (2012). Being right is its own reward: Load and performance related ventral striatum activation to correct responses during a working memory task in youth. *Neuroimage, 61*(3), 723-729.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology, 80*, 1-27.

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews Neuroscience*, *1*(3), 199-207.

Schultz, W. (2002). Getting Formal with Dopamine and Reward. *Neuron*, *36*, 241–263.

Schultz, W. (2007). Reward. *Scholarpedia*, *2*(3), 1652.

Seger, C. A., & Cincotta, C. M. (2005). The roles of the caudate nucleus in human classification learning. *The Journal of Neuroscience*, *25*(11), 2941-2951.

Seitz, A. R., Nanez, J. E., Holloway, S. R., Koyama, S., & Watanabe, T. (2005). Seeing what is not there shows the costs of perceptual learning. *Proceedings of the National Academy of Sciences, USA*, *102*, 9080–9085.

Seitz, A. R., & Dinse, H. R. (2007). A common framework for perceptual learning. *Current Opinion in Neurobiology*, *17*, 148-153.

Seitz, A. R., & Watanabe, T. (2003). Psychophysics: Is subliminal learning really passive?. *Nature*, *422*(6927), 36-36.

Seitz, A. R., & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Science, 9*, 329-334.

Seitz, A. R., & Watanabe, T. (2008). Is task-irrelevant learning really task-irrelevant?. *PloS one*, *3*(11), e3792.

Seitz, A. R., & Watanabe, T. (2009). The phenomenon of task-irrelevant perceptual learning. *Vision Research, 49*, 2604-2610.

Seitz, A. R., Kim, D., & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron, 61*, 700-707.

Seitz, A. R., Nanez, J. E., Holloway, S. R., Koyama, S., & Watanabe, T. (2005). Seeing what is not there shows the costs of perceptual learning. *Proceedings of the National Academy of Sciences. USA*, *102*(25), 9080-9085.

Seitz, A. R., Lefebvre, C., Watanabe, T., & Jolicoeur, P. (2005). Requirement for high-level processing in subliminal learning. *Current Biology, 15*, 2073.

Seitz, A. R., & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Sciences*, *9*(7), 329-334.

Shuler, M. G., & Bear, M. F. (2006). Reward timing in the primary visual cortex. *Science, 311,* 1606–1609.

Smith, K. S., & Berridge, K. C. (2007). Opioid limbic circuit for reward: interaction between hedonic hotspots of nucleus accumbens and ventral pallidum. *The Journal of Neuroscience,* *27*(7), 1594-1605.

Spivey, M. J., & Sprin, M. J. (2000). Selective visual attention modulates the direct tilt aftereffect. *Perception & Psychophysics, 62,* 1525–1533.

Swallow, K. M., & Jiang, Y. V. (2010). The attentional boost effect: Transient increases in attention to one task enhance performance in a second task. *Cognition, 115*(1), 118-132.

Swallow, K. M., & Jiang, Y. V. (2011). The role of timing in the attentional boost effect. *Attention, Perception, & Psychophysics, 73*(2), 389-404.

Swallow, K. M., & Jiang, Y. V. (2012). Goal-relevant events need not be rare to boost memory for concurrent images. *Attention, Perception, & Psychophysics, 74*(1), 70-82.

Swallow, K. M., & Jiang, Y. V. (2013). Attentional load and attentional boost: a review of data and theory. *Frontiers in psychology, 4,* 274.

Tsushima, Y., Seitz, A. R., & Watanabe, T. (2008). Task-irrelevant learning occurs only when the irrelevant feature is weak. *Current Biology*, *18*(12), R516-R517.

Watanabe, T., Náñez, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature, 413,* 844-848.

Watanabe, T., Nanez, J. E., Koyama, S., Mukai, I., Liederman, J., & Sasaki, Y. (2002). Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. *Nature Neuroscience, 5*, 1003–1009.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function. I. Fitting, sampling and goodness-of-fit. *Perception & Psychophysics*, *63*, 1293–1313.

Wise, R. A., & Rompré, P. P. (1989). Brain dopamine and reward. *Annual Review of Psychology, 40*(1), 191-225.

Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance*, *19,* 676-681.

Yotsumoto, Y., & Watanabe, T. (2008). Defining a link between perceptual learning and attention. *PLoS Biology*, *6*(8), e221e-223.