



UNIVERSITÉ
PARIS
DESCARTES

MEMBRE DE

U-S-PC
Université Sorbonne
Paris Cité

École Doctorale Cognition, Comportement et Conduites Humaines (ED3CH)
Laboratoire Psychologie de la Perception (LPP)

The Role of Attention in Conscious Access Mechanisms and their Influence on Visual Representation

Evidence from Psychophysics and fMRI

A Doctoral Thesis in Cognitive Neuroscience

Louis Thibault

2016

Advisors: Claire Sergent & Patrick Cavanagh

Keywords: consciousness, representation, attention, vision, fMRI, Multivariate Pattern Analysis, psychophysics, behavioral model

Abstract

A major finding in the scientific study of conscious perception has been the existence of two temporally-distinct phases of visual processing. The first, characterized by the feed-forward propagation of evoked activity in early visual cortex, is not typically associated with conscious perception. The second phase involves a reactivation of early sensory cortex by downstream regions and is often cited as a correlate — if not a proximal cause — of consciousness. This raises a few crucial questions: firstly, what causes this feedback process to emerge, and secondly, what distinguishes a stimulus representation that has undergone such feedback processing from one that has not?

At the time of writing, two competing theories have been proposed. The first theory, hitherto referred to as "early-and-local", posits that conscious access emerges from the very emergence of a feedback loop between high-level sensory cortex and its primary counterpart, and that this cortical resonance is driven entirely by upstream activations along the feed-forward chain. This implies that only those stimuli that exhibit high salience from the onset can become conscious, and by extension, that the stimulus' reportability is governed entirely by early evoked activity in primary sensory cortex. "Late-and-global" theory, by contrast, posits that conscious perception is the direct result of routing of information through a distributed cortico-cortical network called the Global Neuronal Workspace (hereafter GNW). By this account, visual information in various local cortical regions is given access to routing infrastructure by some selective process, namely attention.

In 2013, Sergent and colleagues tested a prediction derived from this second model: that an arbitrary sensory representation that has initially failed to become conscious can be hooked into the GNW by means of an attentional manipulation. To do this, a low threshold target Gabor patch was presented, followed by an extrinsic cue either at the location in which the Gabor had been presented, or on the opposite side of the screen. Subjects were better at discriminating the orientation of the Gabor in trials where the cue had been presented on the same side as the target, and also reported seeing the target more clearly, suggesting that the retrospective intervention of attention was enabling a weak signal to gain access to the global neuronal workspace.

In our first experiment, we test whether this so-called retro-perception effect indeed hinges on consciousness, rather than resulting from some process whereby features from degraded but still visible targets are sharpened. Using a mixture-modelling approach borrowed from working memory (WM) literature, we find that the precision with which subjects can recall the orientation of a target does not change as a function of retrospective attentional cueing, but the number of flat-out guesses (thought to correspond to cases where no conscious percept emerges) is reduced when the cue is indicative of the target's previous position. In a second experiment, we ensure that our model is sensitive to changes in perceptual precision under classical iconic-memory task conditions and find that precision but not guessing-rate is affected by post-cues. These results invite us to consider whether the effects classically reported in iconic memory (IM) literature may be partially driven by retroperception effects. In order to ascertain the presence of such effects, as well as validate our model in a well-understood setting, we conduct a second experiment using similar methods and stimuli. Whereas the first study placed a high strain on perception (via low-contrast stimuli) and a light load on working memory (via a small set-size), the second study applies an identical mixture-modelling technique to the converse conditions: multiple, high-contrast targets. We find evidence both for retroperception in such settings (increased guessing-rate as a function of SOA) and for multiple attentional strategies across subjects.

Having found evidence for an all-or-none triggering of conscious access by the attentional system, we proceed to ask whether said triggering process modifies the internal representation of the stimulus in early visual cortex. To do so, we turn to functional magnetic resonance imagery (fMRI) and multivariate pattern analysis (MVPA). We hypothesize that low-contrast Gabor patches at perceptual threshold-contrast are represented differently in V1 depending on whether they are classified as “seen” or “unseen” by an observer. We predict that this systematic variance in encoding scheme can be detected by a support vector classifier (SVC), in the form of improved performance in decoding stimulus orientation from occipital voxels. Such a result would suggest that the intervention of the attentional system has a net effect of boosting the signal-to-noise ratio in these cortical regions, resulting in a stronger and/or more stereotyped representation. While we fail to obtain above-chance classification, we propose several refinements to the protocol as well as suggestions for future work.

Taken together, our results point to a causal role for attention in the emergence of conscious perception, consistent with predictions made by Global Workspace Theory. The present work further supports the use of mixture-model analysis (MMA) as an investigative tool in the study of conscious perception and internal stimulus representations.

Acknowledgements

I would like to extend my sincere gratitude to the great many friends, family and colleagues who have lent their support during these past three years. There are too many to list.

In particular, my advisor Dr. Claire Sergent stands out as an exemplary scientist and mentor. I am immensely grateful for her patience, guidance and expertise, without which this adventure would not have been possible.

Likewise, I would like to thank my advisor Dr. Patrick Cavanagh for sharing his intuitions and knowledge during our brainstorming meetings. His experience was instrumental in avoiding methodological pitfalls on more than one occasion.

I would also like to thank my fellow Ph.D students, especially those in room H418B, who, in addition to being generous with their help and support, made for great company and will be sorely missed: Nika Adamian, Solène Le Bars, Marianne Duyck, Romain Vincent, Daphné Rimsky-Robert, Lucile Dupin, Sergiu Popescu, and Vincent Forma. I would be remiss if I didn't extend a similar acknowledgment to my dear friends and partners-in-crime outside of the lab: Vincent Fournier, Guillaume Collard and Casey Chance. Beyond making this possible, you've made it enjoyable.

My parents, of course, deserve special thanks for their constant encouragement and unwavering confidence in my abilities. I am grateful to my father in particular for sharing his wisdom and experience at the time of writing this manuscript and to my mother for her numerous care-packages.

Finally, I must express my most profound love and gratitude to Oriane Georgeac, for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without her.

À mon grand-père

Chapter 1: Introduction

It's Sunday afternoon in Paris, where the long-awaited return of vernal weather has brought with it the usual flock of tourists. Among them, a young American couple sits in a cafe terrace and recounts the day's adventures in great detail. The young woman, elated by the unfamiliar architecture, the aroma of bakeries and the sound of accordion music in the metro, offhandedly remarks that there is surely nowhere else in the world where one might experience all these things *at once*! This unfortunate choice of words sets into motion a long and heated debate about the nature of consciousness, for her companion is indeed an unequivocal pessimist, convinced that the majority of subjective experience is illusory, and that conscious perception is quite sparse.

He argues, "I can only be aware of what I am attending to in the exact moment! As my organ of perception is pointed from one object to the next, I experience each object individually, blissfully unaware of the gaps in my own awareness." She, on the other hand, could not disagree more.

"On some level, I am experiencing everything at once," she exclaims. "Oh sure, I can't remember every detail I see, but by focusing my attention on this thing or that, I can ensure that it is recalled, as well as glean some additional information in the process. Just because I can't remember what I saw a few moments ago doesn't mean I didn't experience it!"

Who, then, is right? Our female protagonist might contend that she has won the argument by virtue of the fact that she perceives a visual scene as unified, thus invalidating the assertion that she is only conscious of a subset thereof. This would be somewhat myopic, however, since she must also consider the possibility that an observer can attend to his entire field-of-view as a singular whole, and in so doing trade a detailed percept of a small object for an approximate percept of a large one. To this effect, our male protagonist might point to a rich

body of literature showing that spreading attention over a large area reduces one's ability to report fine-grained features (Gardelle et al. 2009). But this is hardly conclusive; he has indeed shown that attention is *sufficient* to induce awareness, but is it *necessary*?

From this exchange, the casual eavesdropper might conclude two things: firstly, that the debate is hardly settled, and more importantly, that the debate is not philosophical in nature but rather empirical. Again, bears repeating that our two protagonists concur, to a large extent, that the presence of certain perceptual features suffice to qualify a percept as *conscious*. Stimuli whose informational content can be intentionally communicated across interpersonal boundaries (e.g. the presence of an aroma or the melodic properties of acordeon music), are indisputably experienced with the full benefit of awareness. Likewise, informational quanta that can be integrated across sensory modalities (e.g. the sight of a musical instrument with it's sound), that can form the basis of rational inference (e.g. inferring the presence of a nearby bakery from a sweet aroma), and that can be recalled (e.g. the closing time of the Louvre) are most assuredly conscious percepts at some point or another. As far as this couple is concerned, the debate is not one of metaphysics or even epistemology. Their disagreement stems rather from two questions: which of the myriad concurrent cognitive processes in the wakeful brain are necessary and sufficient to produce this seemingly rich experience? And by what means is sensory information selected for promotion into the realm of awareness? So phrased, the scientific nature of these questions becomes obvious, and the search for answers is reduced to a conceptually simple (though often tedious) matter: the search for minimal contrasts between conscious perception and its absence. Only then can the limits of consciousness be described, and the debate settled.

Laypersons are often surprised that the topic of consciousness should so naturally engender testable, empirical questions. Indeed, one is hard-pressed to find a phenomenon as commonplace as consciousness that has historically been so opaque to empirical instruments. For this reason, a thorough review of the discipline's roots — both in philosophy of mind and the natural sciences — is essential if one is to understand exactly what can be gained from scientific study of consciousness.

Origins and History of the Scientific Study of Consciousness

Consciousness is a fascinating but elusive phenomenon; it is impossible to specify what it is, what it does, or why it evolved. Nothing worth reading has been written on it.

— Stuart Sutherland, 1989 *International Dictionary of Psychology*

The term *consciousness*, in the sense of subjective experience and awareness, entered mainstream use only in the early 20th century with the advent of depth psychology. Prior to this, the word was either uttered in tandem with more nebulous — if not mystical — terms such as *soul*, *mind* or *spirit*, or in the context of clinical descriptions of behavior such as *sleep* or *coma*. But what *is* it, ontologically speaking? What are its properties and behaviors? How is it produced? Does it have function?

While consciousness has entertained the curiosity of man since the dawn of its existence, its subjective nature long made it inscrutable by empirical means, and for good reason, for how does one observe the inner life of others through his own senses? At first glance, there is a strong intuition that consciousness is beyond the reach of empirical methodology. It is subjective, immaterial, and inextricably entangled with the very means by which we understand the world. However, as far back as the 5th century BCE, the Greek physician Hippocrates of Kos noticed that traumatic injury to the brain produced all manner of alterations to the mind, launching a lifelong study of various neurological phenomena. *On the Sacred Disease*, his seminal neurosurgical treatise on epilepsy, catalogs a long list of phenomena that are both immaterial and subjective, and whose origins he ascribes to the brain:

Men ought to know that from nothing else but the brain come joys, delights, laughter and sports, and sorrows, griefs, despondency, and lamentations. And by this, in an especial manner, we acquire wisdom and knowledge, and see and hear, and know what are foul and what are fair, what are bad and what are good, what are sweet, and what unsavory; some we discriminate by habit, and some we

perceive by their utility. [...] Wherefore, I say, that it is the brain which interprets the understanding.

Herein lies the foundation on which the scientific study of consciousness is built: the seat of experience is none other than the brain, and the very subjectivity that is characteristic of individual experience is attributed to some putative (meta)physical process in this organ. This observation, which is by no means Hippocrates' greatest contribution, is incalculable in value, as it asserts that experience — and by extension consciousness — interacts with the physical realm. It then follows that observation can reveal at least some of the more obscure properties of consciousness. Nevertheless, Hippocrates is not remembered as a great theorist of consciousness, for indeed the Hippocratic account of subjective experience provides little more than an anatomical locus; it is most notably lacking *mechanism*. We know where it resides, but how does it work?

Towards a Mechanistic Account of Phenomenology

In the mid 17th century, French philosopher René Descartes undertook the feat reducing man to simple biological and physical components, in effect attempting to discover the very mechanisms that are absent from the Hippocratic account of subjective experience. In his philosophical manuscripts, *l'Homme*, *La Description du Corps Humain* and *Les Passions de l'Âme*, Descartes proposes a system inspired by contemporary advances in the field of horlogery and automata in which an organism reacts to its external environment through various deterministic, mechanical processes (most notably a sort of pneumatic actuation). The nerves are indeed purported to be hollow tubes in which "a certain very fine air or wind" he calls *animal spirit* flow under pressure (Charles & Tannery, 1982 p. 331). Observing that nerves connect sensory organs to the brain, and the brain to various muscles, he further posits that the stock and flow of animal spirit is controlled by a series of valves in the ventricles of the brain, the latter of which serve as pressurized containers. When sensory organs are stimulated, Descartes explains, the flow of animal spirit in, say, the optic nerves actuate a subset of the ventricular valves, which in turn cause pressurized animal spirit to discharge down a nerve that leads to the pineal gland.

Ultimately projected upon the gland's surface, this animal spirit would give rise to a percept, the exact form of which depends on the combination of valves that were stimulated by the upstream eye. This percept, in turn, could actuate a series of valves guiding yet more animal spirit towards the motor apparatus of the body, giving rise to kinetic behavior. By appealing to such metaphors as clockwork and rigging, Descartes has no difficulty accounting for a wide range of low-level, automatic, reflexive behaviors. Moreover, this account cemented the idea that the brain's function related to the translation of sensation into behavior, and more specifically, that the nerves served as the medium through which information and motor impetus travel. Though modern science has substituted animal spirit for action potential, this model of neurological function is surprisingly accurate, at least in conceptual terms. In point of fact, the predominant neurobiological theory of the 19th century — reticular theory — retains an important element of the cartesian description of the body: the nervous system as a matrix of interconnected pipes. In the case of reticular theory, animal spirit was substituted for electrically-conductive substrate, reflecting contemporary discoveries in the field of electromagnetism. The final correction would come at the turn of the twentieth century where the discovery of neural synapses would contradict the model of the nervous system as literal pipework, subsequently leading to the discovery of chemical signalling across neurons.

For simple behaviors, the cartesian model is rather satisfying, and in effect constitutes a transduction system that converts physical energy of one kind into another. Thus, there is nothing inherently mysterious in how the radiant heat of an open flame might provoke a sudden withdrawal of one's limb. All that is required to understand this behavior is an understanding of how heat transfers to the surface of the skin, whereupon basic physical forces provoke the movement of animal spirit in the nerves in accordance with well-understood laws. Says Descartes in his closing of *l'Homme*:

I desire, I say, that you should consider that these functions follow in this [human] machine simply from the disposition of the organs as wholly naturally as the movements of a clock or other automaton follow from the disposition of its

counterweights and wheels. To explain these functions, then, it is not necessary to conceive of any vegetative or sensitive soul, or any other principle of movement or life, other than its blood and its spirits which are agitated by the heat of the fire that burns continuously in its heart, and which is of the same nature as those fires that occur in inanimate bodies

But beyond simple reflexes, Descartes has a problem. How can such a simple system account for man's rich inner life? That is to say, how can a physical system such as the brain act as a medium through which a full mental formulation is rendered? While Descartes' putative clockwork can certainly produce behaviors — even elaborate ones — there is no obvious means through which to select a seemingly arbitrary and specific response from the panoply of behaviors routinely elicited by a single stimulus. Consider, for instance, a case in which a mouse darts across some room (as is often the case in the offices of doctoral students). Witnessing this event can cause a wide range of reactions, from a sudden flinch to a verbal comment about leaving food at the workplace. What's more, the process through which one of these myriad behaviors is selected and executed appears to be largely non-deterministic, contrary to the fundamental means in which mechanical automata operate. Seeing the mouse may produce one reaction the first time, and an altogether different one thereafter!

Faced with the seemingly intractable problem of reconciling stochastic behaviors and the self-evident existence of a subjective inner life with a materialist account of human psychology, Descartes ultimately appeals to the idea that mind is ontologically distinct from the physical. In doing so, he formulates what would become the school of interactionist dualism, claiming that an ethereal mental substance (*res cogitans*) exists in a plane distinct from the physical world, and interacts with the latter through a specific anatomical locus in the human brain: the now (in)famous pineal gland. The problems with this account are, of course, obvious. A mechanism for mind-matter interaction is notably absent, as is an explanation for the process by which mental substance interacts only with the pineal gland while simultaneously evading detection by physical instruments of measure! A satisfactory mechanism remaining absent from the

description of consciousness, Descartes has merely substituted a broad locus of phenomenology (the brain) for a smaller one (the pineal gland).

In the end, Cartesian dualism falls short of providing a mechanistic explanation for consciousness, ultimately shifting the production of experience from the realm of the brain to the realm of mental substance. His work nevertheless provides some insight into the problem. Indeed, Descartes provides significant credence to the claim that the brain functions as an interpreter of sensory signals. To this point, his principal insight is that the various signals emitted by sensory organs (whether in the form of flowing animal spirit or action potentials) are devoid of meaning until the brain has performed its function. The mind, in the Cartesian schema, can be understood as a black box that somehow extracts or ascribes meaning to the information it receives from the senses. Again, for Cartesians, this notion of meaning is intimately intertwined with consciousness since it is the nature of *res cogitans* to be conscious, and the interaction of animal spirit with this mental substance constitutes the precise moment at which sensory signals gives rise to meaningful experience. But for any self-respecting empiricist of the 21st century, this state of affairs is profoundly dissatisfying, as it implies mind-matter interaction escapes the jurisdiction of the scientific method; after all, part of the Cartesian process is described as non-physical. The issue, then, revolves around finding an alternative explanation for the mind that does not appeal to any dualist notions.

Such an alternative appeared in the late 19th century when German physicist and all-around erudite Hermann von Helmholtz published his seminal treatise on physiological optics (*Handbuch der physiologischen Optik*). A massive opus by all accounts, the 1379 pages of the *Handbuch* serve as foundational literature for a breadth of topics studied by modern psychologists, from eye movements, to monocular and binocular vision, and depth perception. On this latter subject, Helmholtz observes that the accuracy with which an observer can estimate the distance of an object varies as a function of the observer's experience or knowledge of the object. Indeed, apparent size of an object conflates the object's true size with its distance from the focal point of the eye, producing an entire class of common interpretive failures. "A person

accustomed to a flat country,” writes Helmholtz, “may easily take a vineyard for a potato patch or pine trees on a distant high mountain for heather, and thus underestimate both the distance and the height of the mountains” (Southall, 1962 30:238). Moreover, he observes that children often make such judgement errors and recounts a relevant anecdote from his own childhood:

I can recall when I was a boy going past the garrison chapel in Potsdam, where some people were standing in the belfry. I mistook them for dolls and asked my mother to reach up and get them for me, which I thought she could do. The circumstances were impressed on my memory, because it was by this mistake that I learned to understand the law of foreshortening in perspective.

It so became apparent to Helmholtz that an observer must have some *a priori* notion of the object’s physical dimensions if he is to correctly interpret its distance. Helmholtz’s crucial insight, however, is not that judging distance requires learning but rather that the information available to the retina is rarely sufficient to reconstruct the exterior world. It therefore follows that certain properties of a visual scene (e.g. size, distance, color, etc) require that the brain make inferences on the basis of incomplete information and strong assumptions about the state of the perceiver’s environment.

To illustrate, consider the eponymous “crater illusion” in which an image of a crater is rotated 180° (Fig 1.1). The first time seeing this manipulation is usually surprising, as the crater appears to have suddenly been transformed into a mound!

Figure 1.1: Crater illusion under ecological conditions. Left: a satellite image of the Occator Crater on Ceres taken NASA space probe *Dawn*. Right, an identical image rotated 180°. (Public Domain, released by NASA/JPL-Caltech/UCLA/MPS/DLR/IDA)

The minimal example in (Fig 1.2) is helpful in illustrating why the illusion arises. In the associated publication, Indian neuroscientist Vilayanur Subramanian Ramachandran demonstrates that inverting a luminosity gradient within a circle is sufficient to produce the

appearance of convexity or concavity depending on the orientation of the gradient (Ramachandran, 1988). As is characteristic of perceptual illusions, the crater illusion is robust; even with posterior knowledge of the crater's true configuration, the observer cannot help but see a mound in the inverted image. Helmholtz accounts for the robustness of visual illusions by suggesting that they reflect statistical properties that are learned from the environment. Such "perceptual learning" results from the brain's ability to extract and retain statistical features of perceptual signals, eventually accruing enough evidence in favor of their invariance to regard them as inviolable constants. Such constants, in turn, are constitutive of the inferential assumptions in perception. Whether this is actually the case or whether the underlying assumptions that govern perceptual inferences are hereditary, the fact remains that the state of the world is decoded on the basis of *a priori* assumptions, and that these give rise to illusory percepts when violated.

Figure 1.2: Crater-illusion stimuli reproduced from Ramachandran (1988). The stimuli in panel (a) are spontaneously perceived as convex while those in panel (b) are usually seen as concave, as the perceptual system assumes that there is a single light source. It is surprising to note that the illusion depends on the luminosity gradient as it appears on the retina, independent of the viewer's sense of gravity. To quote the author, "the reader can verify this by viewing panel (a) while hanging upside-down from the ceiling".

In the case of the crater illusion, the brain infers the volumetric properties of an object on the basis of two regularities observed in nature: (1) there is exactly one major source of illumination and (2) this source of light is always located above the observer. Thus, when the shaded area of a three-dimensional object is situated along its upper half, it is safe to assume that the object is concave, and inversely, shade along the lower half of the object suggests a convex surface. Of course, such unconscious inferences are not limited to the visual domain. The amputation of limbs, for instance, violates a particularly strong assumption of the haptic system, namely that activity in certain nerves follows from the stimulation of its associated limb. When said limb is removed, patients experience all manner of sensations in arms and legs that are no longer present (due to spontaneous neuronal activity), and find themselves unable to suppress

these percepts despite conflicting input from the visual system (Bishnu & Grossberg, 2011; Ramachadran & Hirstein, 1998; Melzack 1989; Pons et al., 1991; Hill, 1999).

In this manner, Helmholtz builds on the initial Cartesian project of providing a materialist account of cognition. The principle of inductive inference indeed provides a precise mechanism by which the brain can give rise to perception and mind, and further suggests that this same principle might also underpin higher-level, conscious reasoning (Helmholtz 1867/1910). This having been said, it does not trivially follow from Helmholtz' observations that the brain can and does perform inferential computations. Even if the principle of unconscious inference was eagerly revisited with the support of novel computational tools and mathematical models from the onset of the cognitive revolution (Barlow, 1974; Gregory 1997), the *Handbuch* does not resolve the question of whether this principle can fully account for subjective experience. Consequently, the possibility of studying consciousness by empirical means continued to meet strong resistance, both from dualists who continued to claim that it constitutes a metaphysical phenomenon, and from behaviorists who reject the notion that the mind can be studied through empirical means (Watson 1913).

The bulk of dualist objections were overcome in 1949 with the publication of Gilbert Ryle's *The Concept of Mind*. Through a linguistic analysis of Descartes' *Meditations on First Philosophy* (1641), Ryle argues that the mind-body problem is in effect a mere pseudo-problem, originating from what he calls *category mistakes*. In brief, the argument begins with the observation that dualism explicitly posits a polar opposition between mind and matter, as the properties of *res cogitans* are described in terms of logical negations relative to their physical counterparts. By virtue of this fact, Ryle concludes that mental and physical events belong to the same category, since properties of an object can oppose one another only insofar as they belong to the same *kind* of thing (Ryle, 1949). The opposing properties *Hot* and *cold*, for instance, belong to the category of *temperature*; the properties *large* and *brief*, by contrast, cannot be said to be opposites since one belongs to the category of the *magnitude* and the other to the category of *temporal*. While his analysis was heralded as having "put the final nail in the coffin of

Cartesian dualism” (Tanney, 2007), the remaining dualists and subsequent behaviorists continued in their objections to the scientific study of consciousness, this time on the basis that science — an objective, third-person discipline by all accounts — cannot study a “subjective” phenomenon such as consciousness, since it amounts to first-person experience.

More specifically, the invalidity-by-subjectivity argument takes the following form: because science is an *objective* method, it cannot account for *subjective* phenomena. This objection, according to philosopher John Searle, rests on a failure to distinguish between the epistemic sense of the objective-subjective dichotomy and its ontological sense. Consciousness is most assuredly *subjective* in the sense that its mode of existence requires a *subject*, that is, the *experience* of conscious content requires an *experiencer*. This constitutes *ontological subjectivity* and is completely irrelevant to the question of whether or not consciousness is amenable to scientific inquiry. What would be problematic for the scientific study of consciousness, would be if the *study* of conscious content were to be found subjective. Indeed, an empiricist would like to measure consciousness in a manner independent of his opinions, aesthetic tastes and personal disposition. This requirement is called *epistemic objectivity*, and it is in no way contradicted by its ontological counterpart (Searle 1998). What is needed, then, is an epistemically objective method for measuring consciousness, such that its variance across experimental conditions can be measured. This problem turns out to be non-trivial because of the so-called *other minds problem*, which essentially states that one cannot introspect into another individual’s mind. As a consequence, the existence of a mental state in a third party appears fundamentally unknowable.

The Science of Consciousness

Circumventing the apparent obstruction caused by the *other minds problem* relies upon an analysis of the levels of abstractions with which empirical phenomena can be understood. When trying to understand or predict the behavior of a system, an observer can reason on the basis of more or less abstract systematization. More concrete levels of reasoning *in principle*

produce more accurate predictions, since they rely on fewer assumptions, but incur a trade-off as these are costly to compute by virtue of the fact that they often include irrelevant details. Dennett proposes three levels of abstraction that constitute rational strategies for understanding a complex system (Dennett 1987, pp. 43-68).

The most basic of such strategies, or “stances”, is the physical stance. Consistent with its low degree of abstraction, the physical stance requires the fewest assumptions: merely those that are inherent to our understanding of physical forces such as mass, energy and momentum. Predicting the path of an electrical current through a conductive medium constitutes a physical stance. Some phenomena, however, benefit from a more abstract approach. The reasoning strategy that minimally abstracts the physical stance is what Dennett calls the design stance. The design stance differs from its physical counterpart in that it generates predictions and explanations on the basis of the system’s purpose and function, resting on the added assumption that said system is functioning as intended (i.e. there are no malfunctions). This confers the advantage of enabling the thinker to predict behavior with little or no knowledge of the underlying physics of its parts, thus simplifying certain problems while producing true conclusions in most cases (Dennett, 1995, p.229). For instance, understanding a conductive medium as an integrated circuit allows one to discuss the tasks that this circuit can accomplish, making it a more useful description of a television remote-control than a description that appeals to infrared waves, circuitry and photovoltaic effects. But just as it is unappealing to describe a remote control in terms of electrons and photons, so too is it unappealing to describe mental states in terms of function. Indeed, the differentiation of physical and mental phenomena from a design stance does is bound to include much irrelevant information. Describing attention in terms of a neurological system that performs signal amplification and band-pass filtering overlooks the emergent, and therefore interesting, properties of the system. For instance, the fact that attention permits behaviors that are directed towards objects in the environment. Again, such properties can be derived *in principle* from a design-level description, but adopting a design stance in describing mental process forces one to consider a great deal of ultimately-irrelevant information. With such intellectual cruff, the risk of confusion and false-inference increases, so

one is naturally inclined to abstract away notions of “design” when dealing with brain processes. Instead, it is more appealing to consider only those features that minimally distinguish mental systems from their component parts. The question therefore becomes: what is the characteristic property of mental events?

The characteristic property of mental events — that which distinguishes them from their physical counterparts — is generally accepted by philosophers of mind to be *intentionality*: the property of referring to something as an object (Brentano 1973, p.97). According to Searle, “the primary evolutionary role of the mind is to relate us in certain ways to the environment, and especially to other people”. Mental states, with their ontological subjectivity, relate their experiencer (the subject) to one or more elements of the outside world (the object) thereby forming relationships.¹ An intentional stance, by consequence, is simply a method of reasoning that ascribes intentionality to the object of study: a cognitive strategy Searle argues is necessary for predicting and providing *post-hoc* accounts of all human behavior (Searle 1999, p. 85). The intentional stance, consistent with its increased level of abstraction relative to the design stance, rests on further assumptions. It assumes firstly that an agent forms beliefs on the basis of its perceptual facilities, its memory of past experience and its epistemic needs (that is, the level of abstraction and other constraints it seeks in its explanation), and secondly that said beliefs lead the agent to form desires, which in turn motivate concrete actions (Dennett 1971, 1978, 1987).

Otherwise stated, reasoning from an intentional stance enables a thinker to predict a system’s behavior by treating it as an *intentional being* whose actions are rationally-determined by goals (i.e. desires that ought be fulfilled), which in turn are based on beliefs about the world. In this manner, understanding the brain in terms of beliefs and rational actions based thereupon is to neuroscience as understanding a remote-control in terms of form and function is to electronics; while it is conceivable that scientists might describe the mechanism by which populations of neurons represent abstract objects and their relationship with the host-organism, it is more

¹ This echoes the description of mental events by Gilbert Ryle (1949) as belonging to the category of *relations* rather than the category of *substance*. Mental events fundamentally relate a subject to its object, a property that is reflected in language.

natural (and arguably safer with respect to the risk of false inference) to take intentionality as a given and describe the behavior of interest directly.

Determining the appropriate stance for a given problem is very much a matter of pragmatic tradeoff analysis. For instance, it quickly becomes impossible to predict the behavior of an integrated circuit that has been overclocked or overvolted, as it's behavior will seem erratic from a functional point of view. In such cases, one must drop to a physical description — one that can account for induced currents from electromagnetic emissions and the melting points of various components — in order to effectively predict the circuit's output. Additionally, an excessively abstract stance can fail to provide a useful conceptual framework without necessarily producing erroneous conclusions. Accounting for the behavior of a remote control from an intentional stance, suggesting for instance that it prefers to watch *South Park* rather than *Game of Thrones*, fails to adequately predict the changing of a channel as compared to relating a change in channel to a button that was pressed. *In extremis*, such premature abstraction increases the risk of coming to absurd conclusions, such as claiming that the remote is not working because it doesn't want to watch daytime television.

When applied to the study of consciousness, an intentional stance again allows the investigator to view self-report of a conscious experience as an intentional act, that is, an act that rests upon beliefs and satisfies desires. In a typical experimental setup, statements from a participant such as “I saw a Gabor patch oriented in a clockwise fashion” can be explained in terms of the locutor having consciously experienced the corresponding percept and desiring to perform well enough to be compensated. Note how the validity of this intentional reasoning rests upon the implicit assumption that the subject's and investigator's goals are aligned. The burden is upon the experimenter to design a study that does not incentivize deceit or subversion of the experimental measure. This indirect approach to studying consciousness by interpreting self-report from an intentional stance is what Dennett calls the *heterophenomenological principle*. It is, in other words, a method that requires consciousness researchers to take self-report seriously. More generally, “self-report” can include a wide variety of behaviors ranging

from verbal utterances to button-presses, but the important point is that these behaviors are interpreted as propositions the subject means to assert. Treated as such, these responses constitute raw data that are emitted as a function of some phenomenon of interest, within the usual confines of noisy signals. Having established a means through which to measure an individual's beliefs regarding a conscious experience, we turn to the problem of validating these beliefs. How can we link these beliefs to physical events?

Detecting correlation and inferring causality is the bread and butter of science. Thus, relating a noisy "belief signal" from a self-reported measure to its corresponding source does not present any novel problems in and of itself. Said signals need only be compared against external events (such as the appearance of a backwards-mask on a display) and against internal neurological phenomena (such as the amplitude of an event-related electroencephalographic potential) and paired with a well controlled manipulation in order to pinpoint the conditions that are both necessary and sufficient for a belief signal (about a subject's own consciousness) to become detectable by experimenters. Importantly, researchers must take care to dissociate contents of a stated belief from its credibility, as the heterophenomenological principle demands that subjects be given free reign to describe how their experiences *seem* to them and that the experimenters remain neutral with respect to the question of whether or not these beliefs about experience are accurate. So long as this prerequisite is respected, the domain of heterophenomenology contains all the necessary component pieces for scientific inference. In point of fact, the heterophenomenological dataset includes the raw intentional behavior of the subjects, concurrent physiological events inside the brain and body, and the concurrent physical events in the surrounding environment.

Having thus reduced the empirical study of consciousness to the study of individual beliefs about conscious experience, the scientist is required to explain two things: (1) the neural mechanisms through which beliefs are formed, and (2) the process through which beliefs give rise to the behavior that is measured — that's it (Dennett, 1993)! The implied equivalence between belief of a conscious experience and an actual conscious experience may seem

problematic at first glance, but quickly reveals itself to be minor. Consider for instance the objection that a subject might have a conscious experience but be missing the belief. In such cases, the subject will report that he experienced nothing, in which case we might ask whether we can speak of a conscious experience in any meaningful sense. Moreover, if the event is just as inscrutable to the subject as to the investigator, there is no way in which to collect usable, empirical data, so the scientist cannot be expected to provide any explanation. In the inverse case, when the subject believes he has experienced something he has not, a similar argument can be made. Indeed, claiming to have experienced specific features of a non-experience is self-contradicting, though this absurd claim is still amenable to heterophenomenological inquiry. The investigator should *in principle* be able to account for the brain functions that give rise to this false impression; the false belief still demands an explanation, while the non-experience does not. Likewise, if the claim is instead that the contents of the experience is somehow ineffable, then the heterophenomenologist must explain that belief too (Dennett 2005).

One of the more controversial features of the heterophenomenologist's methodology is its strong appeal to functionalism. Heterophenomenology indeed asks which brain processes are both necessary and sufficient to produce a report of conscious experience, and thus claims no jurisdiction over questions of *qualia*. Critics of heterophenomenology point out that a functional explanation of consciousness — one that describes consciousness in terms of what it allows a system to do — is not satisfactory because functional descriptions cannot account for individual experience (Chalmers 1995). Why should the accumulation of mental processes we call "consciousness" feel like *something*, rather than feel like *nothing*? Relating qualia to brain function constitutes what philosopher David Chalmers coined the *hard problem* of consciousness, to be contrasted with the *easy problem* of providing a functional, mechanistic explanation for the phenomenon. Arguments in favor of such a distinction often rest on philosophical thought experiments, most famously Frank Jackson's *Knowledge Argument*, which rests upon a long line of arguments to the effect that some knowledge is non-physical in origin (Jackson, 1982). The thought-experiment is stated thusly (Jackson, 1986):

Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black and white room via a black and white television monitor. She specializes in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red', 'blue', and so on. She discovers, for example, just which wavelength combinations from the sky stimulate the retina, and exactly how this produces via the central nervous system the contraction of the vocal cords and expulsion of air from the lungs that results in the uttering of the sentence "The sky is blue". [...] What will happen when Mary is released from her black and white room or is given a color television monitor? Will she learn anything or not?

Jackson's intended conclusion is that, yes, Mary will learn something from the first-hand experience of color, and that there therefore exists non-physical knowledge that escapes functional explanation. A similar interpretation by Paul Churchland concludes on similar ground, holding that (1) Mary has complete knowledge of brain states and their properties (2) Mary is ignorant with respect to certain perceptions, having not experienced them before, and that therefore (3) perceptions are distinct from brain properties (Churchland, 1985). That said, these conclusions result from some ambiguity in the premise, as evidenced by the disagreement between Jackson and Churchland, and this ambiguity is difficult to resolve. For Dennett and others, the claims are downright contradictory. By virtue of the fact that Mary has perfect knowledge of the physical world, she should be able to differentiate any two given qualia in the brain, which in turn implies that she already knows how a given qualia will affect her belief about the world (Dennett 1991, 2006). Likewise, Mary's perfect knowledge of perception and physics gives her the ability to relate a certain known state of perception to other known states, so even if she has not experienced a given state, she is still able to understand it in relation to others (Maloney, 1985).

The impatient reader might object that the hard-versus-easy problem of consciousness is neither here nor there. To his point, the existence and consequences of extra-physical qualia is beyond the scope of heterophenomenology (and by extension, the scope of the present work), and there seems to be no proof positive that a hard problem exists at all. Couldn't this just be another pseudo-problem? While there is no proof positive that the hard-versus-easy problem is a false dichotomy either, the argument bears repeating if only to clarify some terminology. *Consciousness*, to the heterophenomenologist, is a physical phenomenon that has informational content; one is conscious *of something*. This feature of heterophenomenological consciousness has also earned it the name of *access consciousness*, since it is understood that the functional machinery of consciousness “has access” percept or representation. Furthermore the term *access consciousness* places a distinction between sets of perceptual stimuli in which a subset is perceived consciously whereas another (sometimes physically identical) subset is perceived without giving rise to awareness. In keeping with the heterophenomenological tradition, we assert that the hard-versus-easy distinction is largely irrelevant with respect to a functional or mechanistic description of consciousness. Our primary interest is indeed to describe the neurobiological and algorithmic functions underpinning access consciousness, so we deliberately ignore questions relating to qualia, and assume that a functional description of consciousness will produce scientifically-satisfactory knowledge.

Neural Correlates of Consciousness

With the arrival of the Cognitive Revolution in the 1950's, psychologists began studying and characterizing those internal brain processes that had been so diligently ignored by their behaviorist predecessors. Of these mental processes, attention and working memory seemed ubiquitous, showing signs of involvement in virtually every task researchers could throw at a human subject, from dichotic listening, to change-detection, reading, mental calculation, etc. Such experimental results quickly gave rise to a class of cognitive models whose function is to mediate contention over shared cognitive resources in the processing of low-level perceptual signals. The earliest of such models was proposed by Donald Broadbent as an explanation for the inherent difficulty of dual-task paradigms. He proposes a two-stage system in which the stream of sensory information that enters early sensory cortex is placed into a perceptual buffer as it awaits access to a limited-capacity resource. The buffering process could be performed in parallel, but consumption of buffered data from the downstream computational resource was supposed to be strictly serial (Broadbent 1957). Twenty years later, Baddeley and Hitch clarified the nature of the sensory buffering system by suggesting that a *phonological loop* (in effect, repeated vocal or subvocal rehearsal of auditory information) was in play, and by positing the existence of a supervisory entity whose role was to exercise volitional cognitive control over behavior: the *central executive* (Baddeley & Hitch 1975). Roughly ten years later, Norman and Shallice proposed a distinction between supervised processes whose mode of action imposes serial access, and their parallel counterparts (Norman & Shallice 1986), ultimately laying the framework for Bernard Baars' *global workspace* theory (GW), which trades the notion of a serial supervising entity for one that instead routes information around the brain (Baars 1988). Baars' major innovation is that the unsupervised, parallel processes are neither idle as in the models of Broadbent's, Bradley and Hitch's, nor are they simply prerequisite inputs into Norman and Shallice's *supervisory attentional system*. Rather, each process performs its specialized function either locally and asynchronously (*i.e.* in a constrained cortical region, independently of other

similar processes) when it does not have access to the workspace's routing infrastructure, or globally and synchronously when it manages to take hold of said workspace. Further, and perhaps most importantly, Baars explicitly suggests that this broadcasting corresponds to the functional underpinning of conscious perception, a notion that was famously summarized by Dennett as "consciousness is fame in the brain" (Dennett 2005).

If consciousness is fame, then surely there must be groupies: signatures of fame's arrival. In the consciousness literature, these signatures are often called *neural correlates of consciousness* (NCCs), and these exist in a variety of recording modalities. One of the first experimental paradigms to produce reliable NCCs was backwards-masking. This manipulation is unusually appealing because it allows investigators to keep all properties of the masked stimulus constant, thus achieving a minute contrast between trials in which subjects report having seen the stimulus and those in which they do not. Early studies using this paradigm found a simple, but robust finding: while BOLD activity in the primary visual cortex and its neighboring regions is roughly constant across conditions, certain high-level regions sharply increase their global activity across trials in which subjects report seeing the target (Grill-Spector, Kashmir, Hendler & Malach 2000). Importantly, many of these regions exhibit some non-zero level of BOLD activity even during "unseen" trials, but their activity systematically increases in amplitude when the subject is aware of the target, and brain regions that are more frontal (that is, further away from sensory cortex) exhibit progressively stronger responses (Dehaene, Naccache, Cohen, Le Bihan, Mangin, Poline & Rivière 2001). This finding has been replicated in other paradigms and modalities including single-unit recordings in awake monkeys (Leopold & Logothetis 1996; Logothetis, Leopold & Sheinberg 1996), and in humans with the added finding that the time-course of the burst of activity is synchronized across regions (Marois, Yi & Chun 2004; Kouider, Dehaene, Jobert & Le Bihan 2007; Stephan, Thaut, Wunderlich, Schicks, Tian, Tellmann, Schmitz et al., 2002; McIntosh, Rajah & Lobaugh 1999; Petersen, van Mier, Fiez & Raichle 1998). In sum, there is robust evidence for a domino-effect of sorts; a stimulus in early sensory cortex would seem to propagate in a feed-forward pattern towards the rostral part of the brain, accumulating in strength as it progresses, while exhibiting strong

synchrony along the lateral axis. This pattern has been taken to suggest that the implicated anatomical structures are exchanging information, or mutually-reinforcing the signal (Dehaene, Naccache, Cohen, Le Bihan, Mangin, Poline & Rivière 2001; Rodriguez, George, Lachaux, Martinere, Renault & Varela 1999; Varela, Lachaux, Rodriguez & Martinerie 2001).

Beyond the feed-forward wave, conscious perception displays a number of distinct temporal features that are most apparent when measured with encephalographic techniques. As a case and point, consider a study by Sergent and colleagues in which EEG recordings were taken while subjects were engaged in an attentional blink task (Sergent, Baillet & Dehaene, 2005). Subjects were to attend to a rapid serial presentation of character chains and, consistent with the attentional blink task, asked to spot two targets (T1: a chain consisting of Xs and Os, T2: a valid word). At the end of each trial, subjects were asked to identify the characters in the first target as a control and, crucially, to rate the visibility of the second target along a continuous scale. The primary behavioral result is clear: participants had an overwhelming tendency to report a clear perception of T2, or no perception at all. Furthermore, this all-or-none event is mirrored in the time-course of event-related potentials. Indeed, the ERPs for T1 and T2 systematically shared a robust P1/N1 complex roughly 170 ms after the onset of the epoched target, and displayed a propagation consistent with the feed-forward cascade described above. Within the set of T2 epochs, however, the authors report a divergence in the time-course between targets that had been reported as highly visible and those that were reported as invisible. In the ERPs of visible T2s only, a large central positivity emerges in the central electrodes and works its way towards the back of the head over the course of 250ms, ending in a sustained re-activation of visual cortex. One particularly interesting aspect of this P3b component is interaction with scalp-voltages elicited by T1. Specifically, the ERPs from T1 also exhibited P3b components in trials where subjects correctly identified its ordered sequence of Xs and Os, and moreover, the authors report a consistent trade-off between the presence of a P3b in T1 and the presence of the same component in T2. Where T1 exhibits a P3b, T2 will generally not and vice-versa. One exception to this rule exists, however, where instead of being extinguished, the conscious perception of T2 seems to be delayed (Marti, Sigman & Dehaene, 2012). Taken together, these results show a

clear-pattern; in addition to behaving in an all-or-none fashion, conscious access is a strictly serial operation. The feedback wave is now a ubiquitous feature of perceptual neuroimaging and encephalography (Williams, Baker, Op de Beck, Shim, Dang, Triantafyllou & Kanwisher 2008; Roelfsema, Lamme & Spekreijse 1998; Roelfsema, Khayat & Spekreijse 2003; Supèr, Spekreijse & Lamme 2001a; Supèr, Spekreijse & Lamme 2001b; Haynes, Driver & Rees 2005; Williams, Visser, Cunnington & Mattingley 2008), as well as one of the most robust correlates of conscious access (Lamy, Salti & Bar-Haim 2009; Del Cul, Baillet & Dehaene 2007; Donchin & Coles 1988; Bekinschtein, Dehaene, Rothaut, Tadel, Cohen & Naccache 2009; Picton 1992; Melloni, Molina, Pena, Torres, Singer & Rodriguez 2007; Dehaene & Changeux, 2011; Marti, Thibault & Dehaene, 2014).

Though the initial feed-forward cascade and subsequent feedback wave differ in their spatial dispersion patterns and time-course, the behavior of both NCCs suggests that consciousness may depend on a signal amplification or stabilization function. In point of fact, the main distinguishing feature of the feed-forward cascade is none other than progressive increase in signal amplitude as activity propagates towards the rostral pole, and the secondary, latent feedback wave manifests itself as a re-activation of upstream cortex after a brief period of quiescence. A closer look at the full electroencephalographic time-course provides a great deal of insight, and strong evidence to this effect. In 2007, Del Cul and colleagues performed an experiment in which they flashed a digit for a single frame (roughly 16 ms) followed by a backwards mask. Scalp-level voltages were recorded via EEG while the subjects performed a simple forced-choice task in which they had to compare the target number to the numeral 5, thus providing a reliable metric of conscious perception. Again, Del Cul and colleagues compared the ERPs evoked by “seen” targets with those evoked by “unseen” targets and found the usual markers of conscious access: a P1/N1 complex which, early visual cortex, remained invariant across conditions and a P3 complex that appeared concomitantly with correct numeric magnitude judgements (see Fig 1.3). In higher visual areas such as inferior temporal gyrus, however, the P1/N1 exhibited an amplitude that was directly proportional to SOA and whose overall magnitude decreased with distance from visual cortex. This monotonic decrease in P1/N1

amplitude is in sharp contrast to the behavior of the P3b, which increased in direct proportion to SOA with a sudden, nonlinear jump between SOAs for which the subject consistently scored above chance and those where he could not (Del Cul, Baillet & Dehaene, 2007).

Figure 1.3. Nonlinear dynamics of consciousness in EEG. In early visual areas, the P1/N1 amplitude is constant, but varies proportionally to SOA in higher-level visual areas. The P3 amplitude exhibits a nonlinear jump in amplitude between mask SOAs that are too short for the subject to consciously perceive the target and those that are not. Note that the P3 amplitude at threshold is likely an artifact of a mixture of seen and unseen trials.

These results show that the manifestation of a conscious percept corresponds to a sudden change in brain state, supporting the notion that the former is an all-or-none event. What's more, this sudden change in jump from one brain state to another is the partially the result of signal amplification, but to what end? And what drives this increase in P3b amplitude?

The Global Workspace model posits that consciousness is the product of a global sharing of information in the brain, and therefore predicts that transitions into conscious states will be accompanied by an increase in cortical connectivity. Given this strong prediction, one might ask whether there is a known mechanism by which brain connectivity is increased that also increases the amplitude of a signal, and one might turn to oscillatory synchrony as a viable candidate. The putative involvement of neural synchrony in conscious perception is, incidentally, a fairly old idea in the field of consciousness. In the late 1990's, Crick & Koch speculated that the 40-Hz band might be a signature of consciousness (Crick & Koch 1998, 2003), though it has since been shown that unconscious stimuli can still elicit high-frequency activity throughout the entire gamma band (Gaillard, Dehaene, Adam, Clemenceau, Hasboun, Baulac, Cohen & Naccache 2009; Fisch, Privman, Ramot, Harel, Nir, Kipervasser, Andelman et al., 2009; Aru, Axmacher, Do Lam, Fell, Elger, Singer & Melloni 2012). This having been said, gamma-band activity is still significant with respect to conscious perception. Short gamma-band bursts, however, are associated with conscious perception (Gaillard, Dehaene, Adam, Clemenceau, Hasboun, Baulac, Cohen & Naccache, 2009; Fisch, Privman, Ramot, Harel, Nir, Kipervasser, Andelman, et al., 2009; Rodriguez George, Lachaux, Martinerie, Renault & Varela, 1999; Gross, Schmitz,

Schnitzler, Kessler, Shapiro, Hommel & Schnitzler, 2004), and have been shown to facilitate the transmission of information between distal brain regions (Fries, 2005; Womelsdorf, Schoffelen, Oostenveld, Singer, Desimone, Engel & Fries 2007; Varela, Lachaux, Rodriguez & Martinerie, 2001). To boot, several studies have applied Granger causality analysis to encephalographic data and concluded that consciousness induces sudden, bi-directional changes in causality during conscious perception, suggesting a non-linear jump in connectivity between disparate regions of the brain (Varela, Lachaux, Rodriguez & Martinerie, 2001; Gaillard, Dehaene, Adam, Clemenceau, Hasboun, Baulac, Cohen & Naccache, 2009).

In sum, there is ample evidence to suggest that consciousness co-occurs with a sudden increase in signal amplitude in frontal regions, neural synchrony and cortical connectivity, all of which lend credence to the Global Workspace model. Still, the whirling of this conscious machinery raises a few non-trivial questions: what triggers this cascade of events, and how do these changes in brain state affect the brain's internal representation of the stimulus? To this first question, the attentional system is a likely candidate due to its selective signal-enhancement properties, although its involvement in consciousness is controversial. Qualifying the interplay between attention and consciousness, to be sure, remains an arduous task for modern cognitive science, whose tools and methods are at best indirect and limited in scope. This is especially problematic when trying to disentangle two processes that are as intimately related as attention and conscious perception

Attention and Consciousness

To say that attention and consciousness are difficult to disentangle is an understatement. The two processes often present synchronously such that paying attention to, say, a visual percept produces a concurrent conscious experience with no perceptible effort beyond the orientation of one's gaze. This tight coupling has led some philosophers and scientists to postulate that attention and consciousness are strictly equivalent (Merikle & Joordens, 1997; O'Regan & Noë, 2001; Posner, 1994), whereas others have argued for their separation into two distinct processes (Baars, 2005; Block, 2005; Stanislas Dehaene et al., 2006; Iwasaki, 1993;

Kentridge, Heywood, & Weiskrantz, 2004; Koch, 2004; Lamme, 2003; Naccache, Blandin, & Dehaene, 2002; Woodman & Luck, 2003).

Consciousness Requires Attention

Findings from inattention blindness studies suggest that attention plays a crucial role in the emergence of conscious percepts, to the point that its absence might fully explain why stimuli sometimes fail to be consciously perceived (Mack & Rock, 1998, Most, 2010). These findings have been extended to modalities outside of vision, where the contrast between attended and unattended stimuli yields similarly striking results. In the auditory modality, for example, allowing a subject to daydream is apparently sufficient to eliminate both verbal reports of conscious perception and evoked potentials associated with conscious access (Bekinschtein et al., 2009; Marti, Thibault & Dehaene, 2014). In clinical settings, the inability of hemineglect patients to report features of their visual environment can be accounted for in terms of an attentional deficit stemming from lesions along fronto-parietal cortex, particularly in the white-matter tracts (Kooistra & Heilman, 1989; Walker, Findlay, Young, & Welch, 1991; Ward, Goodrich, & Driver, 1994).

Attention Without Consciousness

Somewhere between the extrema of attentive consciousness and inattentive non-consciousness, experimental subjects sometimes attend to an area without consciously perceiving any of the stimuli presented. In lateral masking paradigms, subjects do not report seeing peripherally presented gratings, although the disappearance of these stimuli are still capable of eliciting an afterimage (Heemskerk, DiNardo, & Kosterken, 1996). This, along with further evidence from motion-induced blindness, suggests that the mechanism by which the gratings remain unseen is not a simple function of stimulus intensity (Mitroff & Scholl, 2005).

Subjects also show temporal effects of attention in the form of semantic priming for backwards-masked words, but only in trials where they attended the temporal interval in which

the prime and target words were presented (Naccache et al., 2002). An inverse effect, that is to say the attracting of attention by non-conscious stimuli, has also been demonstrated in a particularly amusing study. Jiang and colleagues demonstrated that images of male and female nudes, rendered “invisible” by continuous flash suppression, still attracted attention. In heterosexual subjects, the effect appeared only for nudes of the opposite sex, suggesting that a high level of processing occurred unconsciously (Jiang, Costello, Fang, Huang, & He, 2006).

Consciousness Without Attention

Dual-task paradigms have been used to argue that consciousness without attention is possible, since the primary task monopolizes top-down attentional resources, whereas subjects are nevertheless able to pick out certain salient stimuli from the background noise (Koch & Tsuchiya, 2007). Similarly scene gist has been used to argue that phenomenal consciousness is somehow special in that it is not dependent on attentional processes. Indeed, subjects can accurately report the gist of a photograph flashed unexpectedly for a mere 30 ms — an interval too short for the engagement of top-down, endogenous attention (Mack & Rock, 1998). This reasoning assumes that top-down attention cannot act retroactively upon traces of evoked cortical activity. Under such an assumption, consciousness must be a fundamentally ballistic process wherein the emergence of a conscious percept is a function of the brain's current state and features of the stimulus itself. Importantly, this excludes the possibility of arbitrary or “at will” deployment of attentional resources, and instead suggests that attention is predetermined by these parameters as well. More recently, however, these claims of attention without consciousness have been somewhat overturned (Cohen, Cavanagh, Chun, & Nakayama, 2012).

The Putative Role of Attention in Theories of Consciousness

As previously mentioned, the presentation of a visual stimulus evokes fast propagation of cortical activation from V1 towards the rostral end of the brain, in a general progression towards higher-level processing centers. More exactly, it can be said that several caudo-rostral sweeps take place in parallel, albeit with different propagation rates due to anatomical variations.

Though different, these propagation rates are fast by cortical standards. By 120 ms post-stimulus, nearly all cortical areas (including primary motor cortex) show evoked responses to visual stimuli (Lamme & Roelfsema, 2000). Perhaps the most striking aspect of this initial cascade is the fact that cells with receptive fields tuned to specific features already exhibit their full responses (Tovée, 1994). In other words, the full gamut of elementary features can be decoded out of approximately the first 100 ms of evoked activity, including complex features like face selectivity in inferior temporal cortex (Oram & Perrett, 1992). In light of these facts, a neural theory of consciousness must account for the mechanism by which the apparently complete set of perceptual information encoded in the initial feed-forward sweep reaches consciousness.

We distinguish between two broad categories of theories of consciousness: (1) those which claim that attention provides gatekeeping functionality for the phenomenological machinery, and (2) those which propose secondary, non-causal involvement of attention in conscious perception. We begin by presenting some details about the second group.

Early and Local Models

Theories positing a weak, non-causative role of attention in the emergence of visual awareness are perhaps best described as being early in time-course and local in topography. Specifically, they contend that certain bits of information are irrevocably biased towards (or against) conscious processing prior to attentional selection, and that these a priori biases are the principal determining factor in arbitrating the contents of awareness (Crick & Koch, 2003; Lamme, 2003).

By far the most influential of the early-and-local theories is Victor Lamme's extension to Francis Crick and Christoph Koch's Coalition of Neurons model, the latter of which posits that conscious perception can be equated with the existence of cortical feedback towards primary sensory areas (Crick & Koch, 2003). When discussing Lamme's Local Recurrence Theory, it is crucial to distinguish between the consciousness dimension, which is mediated by the presence

(or absence) of recurrent activations in upstream cortex, from the dimension of reportability, which is purportedly mediated by attention.

Summarily, this model proposes a simple threshold system whereby sub-threshold stimuli are represented in the initial feed-forward cascade of cortical activity, but ultimately fail to evoke sufficiently strong downstream responses to trigger feedback into early visual areas. By contrast, supraliminal stimuli manage to establish a standing wave of cortical activity through the feed-forward and recurrent feedback connections discussed in the previous section. This metastable configuration sustains the cortical activations in early visual cortex for certain stimuli, resulting in a persistent but wispy form of retinotopic memory, which is said to account for the phenomenal experience of consciousness. This memory, aside from being highly volatile, is only partially reportable, thus accounting for the vagueness of scene gist reports. Moreover, as this retinotopic memory is populated with new information, old data rapidly decays. The logical conclusion of this theory is that individuals are conscious of almost everything, but quickly forget the fine details of a scene. Attention, again, does not intervene until after the emergence of conscious perception, where it fulfils its traditional role as a filter for extended processing, thus acting as a gatekeeper for working memory, the contents of which enjoys the full benefits of reportability. Change blindness, under such a model, can be explained in terms of replacing information in a scene at a rate comparable to the natural rate of decay. Without the intervention of attention, or without a fast, localized change in a small area to trigger the orientation of exogenous attention, these changes will be consciously imperceptible. In this framework, classical cuing counteracts the change blindness mechanism by dramatically increasing the chance that a given feature is encoded into working memory (Lamme, 2003).

Lamme and colleagues propose a mechanism through which neural activity can bias selection of certain stimuli both for encoding into conscious perception and for selection by working memory. By their very nature, some stimuli evoke stronger activations than others (high vs. low-contrast stimuli, moving vs. stationary objects, foveal vs peripheral object, stimuli with optimal spatial frequencies, etc...). These differences in encoding yield a highly heterogeneous

pattern of activation at all levels of the feed-forward hierarchy (Lamme, 2003). These differences in activation topography can radically change the course of processing for a subsequent stimulus by either facilitating the activation for a stimulus (as in priming) or by interfering with its activity (Dehaene et al., 1998; Egeth & Yantis, 1997).

In summary, early-and-local theories predict a non-causal role for attention with regards to conscious processing, insofar as it is neither necessary nor sufficient for a conscious percept to enter awareness *per se*.

The late and global approach to consciousness is more or less synonymous with the Global Neuronal Workspace (GNW) model of consciousness. According to this model, localized activity in sensory cortex is capable of activating other excitatory neurons with long-distance cortico-cortical projections. The activation of these so-called global workspace neurons is thought to enable the integration of a stimulus across distant cortical processing centers, thus enabling elaborate tasks involving, for example, planning or verbal report (Dehaene et al., 2003). In the GNW model, feedback is necessary, but not sufficient to produce a conscious percept. Instead, excitatory re-entry into primary sensory areas from workspace neurons maintains the representation across time, such that it can be accessed by other processing modules.

Because the engagement of GNW neurons is a discrete phase transition, conscious perception exhibits itself as a nonlinear function of stimulus salience (Dehaene et al., 2003; Sergent et al., 2004; Sergent et al., 2005). A gradual increase in the contrast of a Gabor patch, for example, will suddenly yield a discrete moment in which the stimulus accesses the global workspace. It is important to note, however, that under this model, attention influences the degree of salience attributable to our hypothetical Gabor patch. The GNW, to be sure, predicts that attention is a *sine qua non* condition for the emergence of conscious perception, and as such, there exist three separate conditions of (non)consciousness: (1) Subliminal stimuli are not consciously perceived because they lack raw signal strength. Low-contrast stimuli are a typical example. (2) Preconscious stimuli are those with a sufficiently strong signal-to-noise ratio to potentially give rise to a conscious percept, but lack the dedicated attentional resources to realize this feat. Inattention blindness and change blindness stimuli correspond to this category of non-conscious perception. (3) Finally, conscious stimuli are those which have both the signal strength and the attentional resources required to access the global neuronal workspace (Stanislas Dehaene et al., 2006).

The early-and-local family draws support from observations of early predictors of consciousness, including many of the aforementioned neural correlates of consciousness (Pins & Ffytche, 2003; Palva, Linkenkaer-Hansen, Naatanen & Palva, 2005; Fahrenfort, Scholte & Lamme, 2007; Railo & Koivisto, 2009; Koivisto, Lahteenmaki, Sorensen, Vangkilde, Overgaaard & Revonsuo, 2008), but do not necessarily appear in all instances of similar experimental paradigms (van Aalderen-Smeets, Oostenveld & Schwarzbach, 2006; Lamy, Salti & Bar-Haim, 2009; Sergent et al 2005). Thus, the question remains. Is consciousness responsible for mediating conscious access to perceived stimuli, and how does the transition between conscious and unconscious perception in any way alter the percept?

In the subsequent chapters, we investigate the role of attention in mediating access to stimuli at threshold and attempt to provide insight into the distinguishing representational features of conscious and nonconscious percepts. We will attempt to partially settle the debate between our fictional couple, showing both *what* consciousness is made of and *how* it is made from raw materials.

Chapter 2: Seeing better versus seeing more often

Abstract

Cueing attention after the disappearance of visual stimuli biases which items will be remembered best. This observation has historically been attributed to the influence of attention on memory as opposed to subjective visual experience. We recently challenged this view by showing that cueing attention after the stimulus can improve the perception of a single Gabor patch at threshold levels of contrast. Here, we test whether this retro-perception actually increases the frequency of consciously perceiving the stimulus, or simply allows for a more precise recall of its features. We used retro-cues in an orientation-matching task and performed mixture-model analysis to independently estimate the proportion of guesses and the precision of non-guess responses. We find that the improvements in performance conferred by retrospective attention are overwhelmingly determined by a reduction in the proportion of guesses, providing strong evidence that attracting attention to the target's location after its disappearance increases the likelihood of perceiving it consciously.

Introduction

What is the role of attention in conscious perception? This question is central in current discussions of the neural mechanisms of conscious perception (Dehaene, et al., 2006; Koch & Tsuchiya, 2006; Lamme 2003). Some authors propose that conscious perception is entirely determined during the build-up of representations within sensory areas, and that, although attention can modulate this process, it is not part of the core mechanisms of awareness (Lamme, 2003; Tsuchiya & Koch, 2005). In contrast, other authors propose that conscious perception arises when and only when sensory representations are broadcast, shared and maintained within a wider network of cortical regions, including supramodal areas (Dehaene, et al., 2006; Koch & Tsuchiya, 2006; Lamme 2003; Tsuchiya & Koch, 2005; Baars, 2005). In this latter view, attention would act as a gatekeeper that mediates this broadcasting event. This second

proposition leads to a strong prediction: if a sensory representation initially fails to become conscious, it should still be possible to promote this representation into awareness by orienting attention towards its residual sensory trace (Tanaka & Sagi, 1998), even after the stimulus itself has disappeared.

Cueing attention after a visual display has classically been used in “iconic memory” experiments, where each display contains several high-contrast items, for example an array of letters (Sligte, Scholte & Lamme, 2008; Sperling 1960). These experiments show that, although participants are limited in the number of letter identities they can report from a briefly presented array (no more than 3 or 4), cueing attention to one specific row within one second after the display can still improve how well these cued letters are recalled. A classical interpretation of this effect is that post-cued attention can bias which items are transferred to working memory. In such protocols it is difficult to assess whether conscious perception itself is affected by post-cueing. Specifically, when the number of items presented exceeds working memory capacity, one can argue that what is reported is less than what has been consciously perceived. In other words, in these type of experiments there may be a dissociation between the content of conscious perception and the content of conscious access (i.e. the representations that are present in working memory and can be reported).

In order to test our prediction that perception itself can be influenced by retrospective attention, we developed a protocol where we ask participants to report a single Gabor patch at threshold contrast. In this case, the stimulus does not exceed working memory capacities and report should faithfully reflect conscious perception. In a series of experiments we tested the influence of retrospective attention (or “retro-cueing”) on the perception of this single Gabor patch (Astle, Summerfield, Griffin & Nobre, 2012; Sergent et al., 2013). We showed that attracting exogenous attention to the stimulus location after its disappearance improved objective orientation and detection sensitivity (d') as well as subjective visibility (Sergent et al., 2013), suggesting that retrospective attention can indeed improve conscious perception.

While this effect could be taken as evidence that retro-cueing elicits a discrete transition to conscious access, and thus from no conscious perception to conscious perception, another possibility is that performance improves because retro-cueing affects the fidelity with which an already conscious content is maintained. Studies on working memory suggest that retrospective attention can prevent rapid forgetting of fine-grained information in displays with multiple high-contrast items (Griffin & Nobre, 2003), so the same process could be at work in this retro-perception phenomenon. A two-alternative forced-choice (2AFC) task does not allow these options to be disentangled; changes in the number of seen trials or changes in the quality of a conscious representation produce similar changes in performance for forced choice.

The present study directly tests these two options through the use of a finer grained measure of perceptual content: a continuous, stimulus-matching task or “reproduction task” (Fig 2.1). Instead of choosing between two options (correct orientation versus orthogonal orientation), subjects were instructed to continuously adjust the orientation of a probe in order to match the orientation of the preceding target. Previous literature shows that the response distributions in such reproduction tasks can often be accurately described by a mixture between a Gaussian distribution around the target’s true orientation with a certain standard deviation, and a uniform distribution, due to trials where subjects guessed, i.e. responded in the absence of information about the target’s actual orientation (Asplund, Fougner, Zughni, Martin & Marois, 2014; Zhang & Luck, 2008). A mixture model analysis of these distributions allows for separate estimates of the proportion of “guess” trials in which the target was not consciously accessed, and the precision of the consciously accessed representations. The two accounts of retro-perception make opposite predictions regarding these measures (Fig 2.2). In the first account (Fig 2.2A and 2.2B), retro-cued attention may prevent the typical loss of precision of the target with time for targets that are already in awareness. Consequently valid retro-cues (same side as the target) should increase the precision of reported target orientation relative to invalid cues, without affecting the percentage of guesses (Fig 2.2A and 2.2B). Alternatively, in the second account (Fig 2.2C and 2.2D) retro-cued attention may act on targets that have not reached awareness and bring their initially unconscious sensory trace into awareness. In this case, the frequency of guesses should

decrease with valid retro-cues compared to invalid ones while the precision of responses should be unaffected (Fig 2.2C and 2.2D).

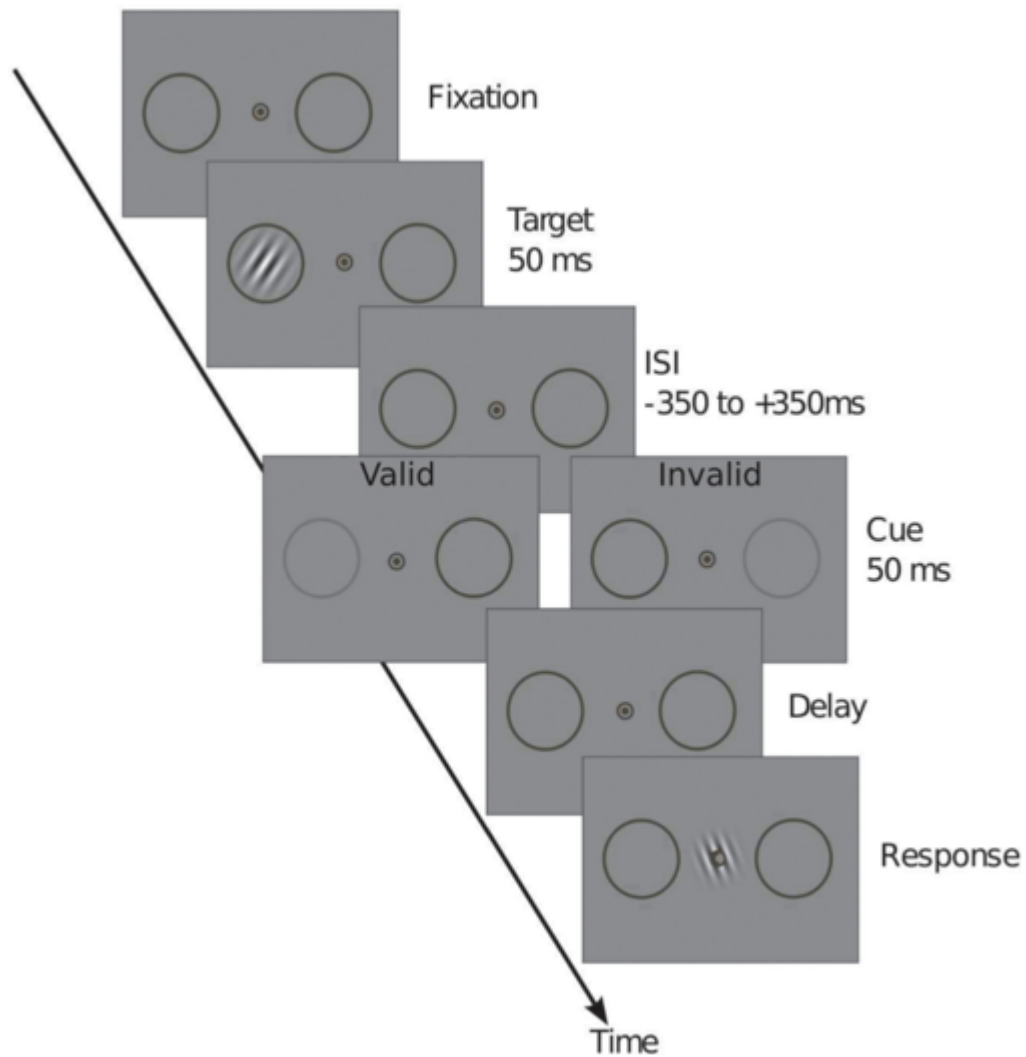


Figure 1.1. Experimental design. A target appeared in either one of the circular placeholders and was preceded or followed by a pre or retro cue in the form of a brief dimming of one of the placeholders. Subjects reported the orientation of the target using the central Gabor patch. On the response screen, a report cue (thickening of one side of the fixation circle) indicated where the target had been presented so that there was no location uncertainty at the time of the response. Note: stimuli are not to scale on this representation.

Materials and Methods

Participants

The number of participants was fixed to twenty prior to the experiment, based on our previous observation and replications of the retro-perception effect (Sergent et al., 2013). Twenty participants between the ages of 18 and 32 took part in the study, each exhibiting normal or corrected-to-normal vision. Of these, three were excluded because they failed to converge on a stable contrast threshold during the initial staircase, or because they failed to perform above chance in all conditions. The 17 remaining subjects (9 women, 8 men) had an average age of 23.7 years \pm 2.1. All participants gave informed consent in writing prior to participation, and the Université Paris Descartes Review Board, CERES, approved the protocols for the study in accordance with French regulations and the Declaration of Helsinki. Participants received a compensation of 10€ per hour for their time.

Apparatus and Stimuli

Stimuli were generated and responses recorded using the Psychophysics Toolbox for Matlab (Brainard, 1997). Stimuli were presented on a CRT monitor (Sony Trinitron GDM-F520). Refresh rate was 60 Hz and screen resolution was 1280 by 1024 pixels. Participants were seated 60 cm away from the monitor, in a dimly lit room. Eye fixation was monitored and recorded using an Eyelink 1000 (SR Research Ltd., Osgoode, Ontario, Canada). We verified that subjects maintained fixation during the majority of trials. This was determined by counting the number of trials during which the mean fixation exceeded 1 degree of eccentricity from the central fixation point (the border of the placeholders were at 3 degree on each side). On average, subjects exceeded this threshold on 0.3% of the trials.

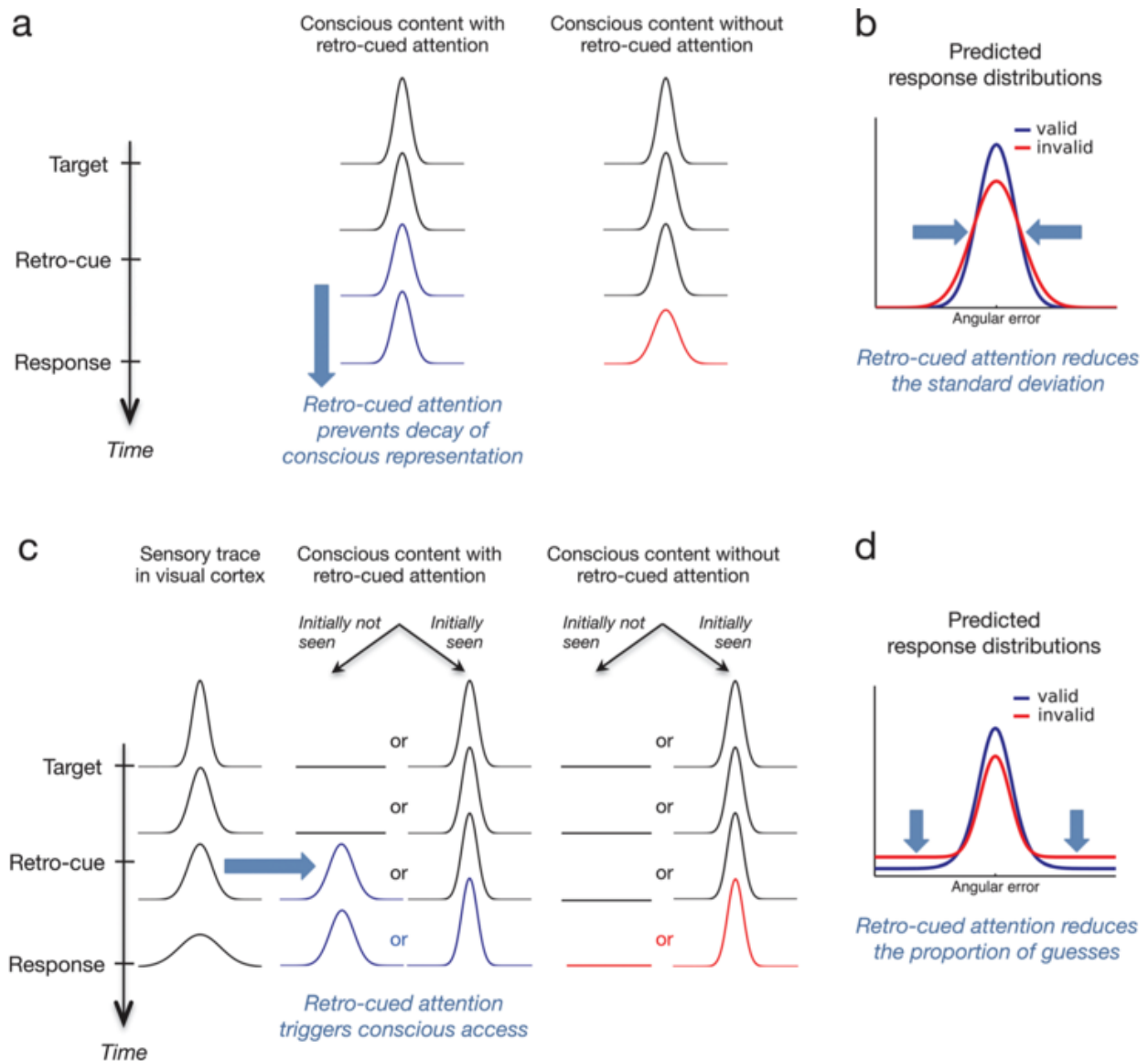


Figure 2.2. Hypotheses and predictions. According to a first hypothesis (a) retro-cued attention prevents the decay of an existing conscious percept. In this proposition, even when the target is conscious, in the absence of retro-cued attention the precision of this conscious representation decays slightly with time (right column). When retro-cued attention is focused on the target's location (left column), it would prevent this slight decay of the conscious representation. This hypothesis thus predicts that the precision of the response on target's orientation should be increased for valid retro-cues (blue curve) compared to no cues or invalid retro-cues (red curve) (b). The alternative hypothesis (c) is that retro-cued attention triggers conscious perception on trials where the target would otherwise have been missed. In this proposition, the target is not always consciously accessed, and thus not always consciously seen following its presentation (right column), but it always leaves a sensory trace in the visual cortex (left column). On trials where the target initially failed to reach conscious access, retro-cued attention at the target's location could still promote the remaining sensory trace in visual cortex at this location to be consciously accessed (middle column). This hypothesis predicts that valid retro-cues (blue curve) should decrease the number of guesses compared to no cues or invalid retro-cues (red curve) (d). It also predicts a decrease in the precision of the accessed information: indeed valid retro-cues trigger conscious access to a degraded sensory trace on trials that otherwise would have counted as guess. Thus, less precise representations get included in the standard deviation estimate.

Stimuli (Fig 2.1) were presented on a gray background (12 cd/m²), and participants were told to fixate a small black circle at the center of the screen (.6° in diameter). Two larger black circles (2.4° in diameter) were always present bilaterally, with their centers positioned 4° to the left and right of central fixation, and served as placeholders for the two possible target positions. They also provided a means for attentional cueing, as introducing a brief decrease in one of the placeholders' contrast produced an attention-grabbing flash.

Targets were Gabor patches subtending 2° in diameter (2 cycles per degree with a randomized phase; Gaussian envelop with 1° full width at half maximum) and were presented in one of twelve orientations spanning 7.5° to 172.5° in increments of 15°. The contrast of the target was determined for each individual using a staircase procedure that converged on a hit-rate of 80% (proportion of trials with an absolute angular error smaller than 45°).

Each trial began with the onset of a dot at the center of the fixation circle. Following a random delay between 500 ms and 900 ms, a target was presented for 50 ms within either of the placeholders. A brief reduction in the contrast of one of the placeholders, turning from black to dark gray (6 cd/m²) for 50 ms, drew attention to the side of the target (valid cue) or the opposite side (invalid cue). This attentional cueing could take place before (SOA -100 ms) or after the target (SOA 100 ms or 400 ms). Each experimental block of 156 trials contained 12 trials where no cue was presented. A response screen appeared following another 500 to 900 ms delay, comprising a response cue in the form of a thickening on one side of the central fixation circle that indicated where the target had appeared, so that there was no uncertainty about the target's location at the time of the response. The response screen also included a response Gabor patch presented at fixation and subtending 2 degrees of visual angle that subjects used to reproduce the target's remembered orientation. Its parameters were the same as the target except that its contrast was 100%, its sinusoidal phase was fixed at .5 radians, and its initial orientation was random.

The participant's task was to reproduce the remembered orientation of the target by freely and continuously varying the orientation of the response patch using the mouse. A small black dot above the response patch indicated the mouse position on the screen. Subjects were not limited in their response times. They indicated their final choice with a left click. Subjects were instructed to always provide a response, and guess in the event that they had not seen the target. Feedback was provided at the end of each block in the form of percentage of hits (a response deviating more than 45° relative to the target orientation was considered as a “miss”).

Procedure

The experiment consisted of two or three staircase blocks of 80 trials and 8 experimental blocks of 156 trials each. The “staircase” blocks consisted of a psychometric staircase function (weighted up down procedure) that converged on a hit-rate of 80% (proportion of trials with an absolute angular error smaller than 45°) (Kaernbach, 1991; Appelle, 1972). Staircase blocks were identical to their experimental counterparts with two exceptions: (1) no cues were presented and (2) target contrast initially began at 100% and was decremented/incremented as a function of the correctness of the previous response (an absolute angular error smaller than 45° was considered as a correct response). In exceptional cases, a third staircase was performed to help stabilize performance (2 subjects).

In the standard experimental block, all targets were presented at the contrast for which the staircase function predicted 80% hit-rate in the absence of cueing (contrast was on average $3.38\% \pm .72\%$). At the end of each block, the participants received feedback in the form of their overall hit rate (percent of trials with angular error inferior to 45°). For four participants, the target's contrast had to be readjusted between experimental blocks (once for two subjects, twice for one subject and three times for one subject) because the overall hit rate had become too high (> 90%) or too low (< 70%). Responses were collected over eight blocks of 156 trials resulting in a total tally of 192 trials for each Validity x SOA condition and 96 “no-cue” trials per subject.

Analysis

Overview of analysis steps

The main question that we wish to answer is whether retrocuing affects (i) the probability that subjects have conscious access to information about a past stimulus, (ii) the precision of this information, or (iii) both. To answer this question, we fit a range of plausible models to our data and examine the parameter estimates of the model that best accounts for the data. Before we fit the models, we remove bias from our data caused by the oblique effect (Appelle, 1972). All steps are described in detail below.

Bias correction

For each trial, the orientation of the subject's response was subtracted from the target's true orientation, yielding the angular error. Biases in angular error varied across target orientations due to oblique effects (Appelle, 1972). After verifying that the magnitude of this oblique effect was unaffected by our experimental conditions (see S1 Fig), we normalized our data in the following way. This bias estimation and correction was performed independently for each participant. For each participant, we took the median angular error as a function of the 12 possible target angles across all experimental conditions as a first estimate of the biases profile. Since oblique effect biases were symmetrical around the vertical and horizontal meridians for each subject, we further averaged the absolute bias across symmetrical angles and replaced the initial estimates with this average, correctly signed. This bias estimate for each target angle was subtracted from the corresponding error distribution, thereby yielding error distributions centered around zero for all target angles.

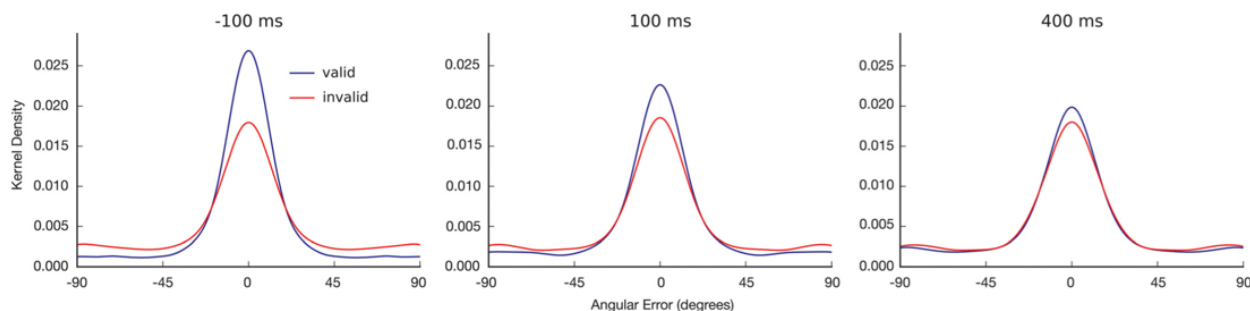


Figure 2.3. Observed response distributions. Distributions of the angular response errors around the target's true orientation (kernel density estimation) for valid and invalid cues at the three different SOAs: pre-cues (-100 ms, left

panel) or retro-cues (100 ms and 400 ms, middle and right panels). A reduction in the uniform component (P_{guess}) is apparent across SOAs as a difference between the valid and invalid distributions in the extrema of the curve.

Basic mixture model

We hypothesize that the distribution of a subject's orientation judgment errors (Fig 2.3) reflects two kinds of trials: trials in which orientation information was consciously available and trials in which such information was not available (Zhang & Luck, 2008). In the first type of trial, errors are expected to follow a Von Mises distribution (the circular equivalent of a normal distribution) that is centered on the target's true orientation. The width of this distribution reflects the average precision with which the orientation was remembered: a narrower distribution means that the orientation was on average remembered with higher precision. In the second type of trial, responses are expected to be pure guesses, thus producing a uniform error distribution. The predicted error distribution, i.e. the probability of producing an angular error x , is thus of the form:

$$P(x \vee P_{\text{guess}}, \kappa) = P_{\text{guess}} \frac{1}{2\pi} + (1 - P_{\text{guess}}) \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(x)} \quad (1)$$

where

- The first term specifies the (uniform) guessing component and the second term specifies the (Von Mises) non-guessing component;
- x is the angular error (in radians);
- P_{guess} is the proportion of guess trials (a free parameter);
- κ is the concentration parameter of the Von Mises distribution; this free parameter can be interpreted as the precision of the memory (higher κ produces a narrower error distribution);
- $I_0(\cdot)$ is the modified Bessel function of the first kind of order 0 (the formula is provided as supplementary information)

Factorial design of 4 mixture models for model selection. The basic mixture model specified above assumes that memory precision is a fixed quantity throughout the experiment. However, several studies have found that, in classical working memory experiments where each

display contains several high-contrast items, working memory errors are often better accounted for by models in which working memory precision varies across items and trials (van den Berg, Awh & Ma, 2014; van den Berg, Shin, Chou, George & Ma, 2012; Fougine, Suchow & Alvarez, 2012). Moreover, such variable-precision models do not necessarily need a guess rate to successfully explain memory errors. To examine whether our data are best accounted for by an equal-precision (EP) or variable-precision (VP) model and whether or not a guessing component is required, we implemented a factorial model design with 2 factors (“variability in precision” and “guessing”) with 2 levels each (“absent” and “present”). This 2x2 design thus gives rise to the following 4 models:

1. Equal precision without guessing
2. Equal precision with guessing (i.e. the basic mixture model described above)
3. Variable precision without guessing
4. Variable precision with guessing

If we find that the data are best accounted for by a model without a guessing component, we should conclude that retro-cueing can only affect precision (or variability in precision) and not the probability with which a subject has conscious access to information about a past stimulus. If, on the other hand, we find that the data are best accounted for by a model with a guessing component, then we can analyze the parameter estimates to examine the effect of retro-cues on recall precision and the guess rate. Following previous work (van den Berg, Awh & Ma, 2014; van den Berg, Shin, Chou, George & Ma, 2012), we model variability in precision across trials by using a gamma distribution. Defining precision as the concentration parameter of the Von Mises distribution, κ , the predicted distribution of orientation errors in the VP-with-guessing model is thus specified as

$$P(x \vee P_{guess}, \kappa) = P_{guess} \frac{1}{2\pi} + (1 - P_{guess}) \int_0^\infty \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(x)} \gamma(\kappa \vee \underline{\kappa}, \tau) d\kappa \quad (2)$$

Where $\gamma(\kappa \vee \kappa, \tau)$ is the Gamma distribution with a mean κ and shape parameter κ . The predicted error distributions of the EP and VP models without guessing are identical to the models specified in Eqs (1) and (2), respectively, but with P_{guess} fixed to 0.

Model fitting

We divided the data of each subject into 7 subsets (2 cueing conditions, valid or invalid, times 3 SOAs plus a no-cue condition). We fit all models separately to each of the 7 subsets. Fitting was done using Matlab's `fminsearch` function to find the maximum likelihood parameter values.

Statistics

When reporting the ANOVAs F , we report corrected degrees of freedom using Greenhouse-Geisser.

Results

Model-free analyses

Both pre and retro-cueing improved performance in reporting the target's orientation, as reflected in the response distributions (Fig 2.3) and in the average absolute angular error around the target's true orientation (Fig 2.4A). We analyzed how the average absolute angular error varied as a function of our experimental conditions (Fig 2.4A) using a repeated-measures analysis of variance on Validity x SOA (2x3). Participants were more accurate in reproducing the target's orientation (decrease in absolute angular error) on trials where the cue attracted their attention to the side where the target appeared (valid cue) compared to trials where the cue was on the opposite side (invalid cue), $F(1, 16) = 79.96, p < .001, d = 1.57$. This was true for cues presented before the target, $t(16) = -.79, p < .001, d = 2.28$, as expected from the classical literature on attention [22–24], and also for cues presented after the target disappeared: SOA 100 ms, $t(16) = -6.10, p < .001, d = 1.21$, and SOA 400 ms, $t(16) = -3.60, p < .005, d = .63$. The effect of SOA was significant, $F(2, 32) = 18.36, p < .001$, as was the interaction between validity and

SOA, $F(2, 32) = 37.56$, $p < .001$. These results replicated the effects observed in previous retro-perception experiments with this new measure of angular error (Sergent et al., 2013).

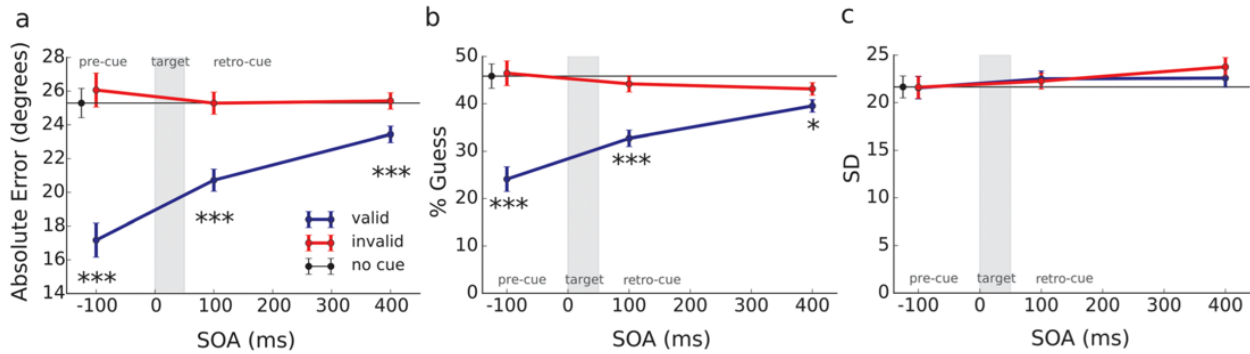


Figure 2.4. Angular error and parameter estimates. Effect of cue validity and SOA on mean absolute response error (a), on percentage of guesses (parameter P_{guess} of the model) (b), and on standard deviation (SD) (c). Error bars represent standard error of the mean effect size. For “no cue” trials, error bars represent the standard error of the mean.

We found no significant difference in average error on trials that were invalidly cued versus those where the cue was absent. This comparison with the “no-cue” baseline condition revealed that the above effect of validity on task accuracy was due to a benefit of valid cueing rather than a cost of invalid cueing.

Model selection

To obtain insight into the statistical nature of the error distributions, we fit 4 different mixture models (see Methods) to the data of each subject. For each of these 68 model fits (17 subjects times 4 models), we computed the Bayesian Information Criterion (BIC), which is a measure of how well a model accounts for the data, taking into account the number of free parameters (Schwarz, 1978). For 16 out of 17 subjects, the EP(equal precision)-with-guessing model was the preferred model, outperforming the runner-up model with an average BIC difference of 19.5 ± 4.2 (mean \pm sem). For the remaining subject, the VP(variable precision)-without-guessing model was the preferred model. However, the BIC difference with the EP-with-guessing was only 0.6, which is negligible. Hence, a guessing component is important to account for the data, but variability in precision across trials is not required. This observation contrasts with recent studies of visual working memory using displays with several high-contrast items

where estimation error data are better accounted for by variable-precision models (van den Berg, Awh & Ma, 2014; van den Berg, Shin, Chou, George & Ma, 2012). However, those studies all used set sizes larger than 1. If variability in memory precision is caused by an inability in dividing mnemonic resources exactly equally across multiple items, then it is not surprising that we do not find evidence for such variability in our study.

Precision versus guessing for the selected model

Having identified the best model (the standard mixture model, i.e. equal precision with guessing), we investigated whether the improvement observed with pre and retro-cued attention was due to a decrease in either the proportion of guesses or the standard deviation (Fig 2.3). In the absence of cueing the proportion of guesses was around 45% (Fig 2.4B). Both pre and retro-cueing to the target's side reduced the number of trials in which subjects guessed their responses, as evidenced by a reduction in P_{guess} for valid trials relative to invalid trials, $F(1, 16) = 86.70$, $p < .001$, $d = 1.45$. This was once again true for trials in which the cue preceded the target, $t(16) = 8.49$, $p < .001$, $d = 2.03$, and when the cue followed the target by 100 ms, $t(16) = 6.64$, $p < .001$, $d = 1.23$, or even by 400 ms, $t(16) = 2.67$, $p = .017$, $d = .42$. We found a significant main effect of SOA, $F(1.76, 28.23) = 9.27$, $p = .001$, and a significant interaction between validity and SOA, $F(1.51, 24.09) = 27.94$, $p < .001$, mirroring the pattern of results observed for angular error.

In contrast, cue validity did not significantly affect the precision of report for seen trials, as reflected by the standard deviation (SD) parameter (Fig 4C), $F(1.00, 16.00) = .44$, $p = .250$. There was a modest increase of SD with SOA, $F(1.85, 29.49) = 3.90$, $p = .034$. No interaction was found between validity and SOA, $F(1.71, 27.28) = .489$, $p = .589$.

Discussion

Our aim here was to test the prediction that retrospective attention can trigger conscious perception. We asked participants to report a single target Gabor patch, shortly after it has been presented (less than a second), with no uncertainty on where it has been presented (thanks to a response cue). In this setting, if participants fail to report this target, it is reasonable to assume

that they also failed to perceive it consciously. Conversely, improved performance should reflect improvement in conscious perception.

The present results confirm our previous observation of a retroactive effect of attention on conscious perception, a phenomenon we call “retro-perception” (Sergent et al., 2013): although the target was a single Gabor patch at threshold, attracting exogenous attention on its location 100 ms or 400 ms after its presentation substantially improved participants’ ability to reproduce its orientation (Fig 2.4A). In our previous studies, results on subjective visibility suggested that retro-attention triggered discrete transitions in conscious access, and thus in conscious perception (Fig 2.4 in Sergent et al., 2013). In the present study we formally tested this proposition using a continuous orientation matching task and a mixture model analysis.

Although mixture model analyses have mainly been used in working memory experiments with several high contrast items (Luck & Vogel, 2013; Ma, Husain & Bays, 2014), Asplund and colleagues recently used this method to confirm that the attentional blink, which is known to impair perception, is not due to a degradation of the sensory representation of the “blinked” stimulus but to a discrete blocking of conscious access to that information (increase in the number of guesses) (Asplund, et al., 2014). This was an elegant way to corroborate, using an objective measure, observations that were initially made using a subjective visibility measure (Sergent, Baillet & Dehaene, 2005; Sergent & Dehaene, 2004). Here we adopted the same strategy to probe the effect of retrospective attention on precision and guessing. A comparison of four plausible models confirmed that the standard mixture model (equal precision across trials plus guessing) was the one that accounted best for the response distributions obtained in the present study. The parameters estimated from this model showed a very clear-cut pattern whereby the benefits of pre or retro-cued attention were accounted for by a reduction in the number of “guesses” (Fig 2.4). By contrast, the precision of representation was not affected by whether the cue was valid or invalid (Fig 2.4C).

These results rule out the hypothesis that the retro-perception effect stems from a memory rather than a perceptual effect; if retro-perception prevented a rapid decay of seen representations in memory, we should have observed an improvement of precision in valid retro-cue trials compared to invalid or no-cue trials (see hypotheses and predictions in Fig 2.2). We find no evidence for such improvement. So, in contrast with the widely-held assumption that events occurring after the disappearance of a stimulus can only affect post-perceptual processes such as decision or working memory (Kinchla, Chen & Evert, 1995; Prinzmetal, Long & Leonhardt, 2008; Vogel, Woodman & Luck, 2005; Yeshurun, Montagna & Carrasco, 2008), the present results show that retro-cued attention can also directly affect whether the stimulus is seen or not.

Our results provide strong support for models of consciousness according to which conscious perception arises when and only when representations held in sensory cortex are broadcast and maintained within a “global workspace” that includes higher-level cortical areas (Dehaene et al., 2006; Sergent & Naccache, 2012; Dehaene, Sergent & Changeux, 2003). In such models conscious access and conscious perception are tightly linked and these models suggest that attentional selection acts as a gatekeeper for such broadcasting mechanisms. The present study validates a very strong and counterintuitive prediction of these models: even when a stimulus initially fails to be perceived consciously, inducing a reactivation of the associated sensory trace by attention can promote it to awareness. In other words, attention can cause conscious perception after the stimulus has disappeared. The present observations also support the notion of “preconscious representations” that we developed earlier (Dehaene et al., 2006): in the present experiment, when a target initially fails to become conscious, it is preconscious in the sense that its conscious fate is still uncertain, since retro-cueing can still promote it to conscious access and conscious perception.

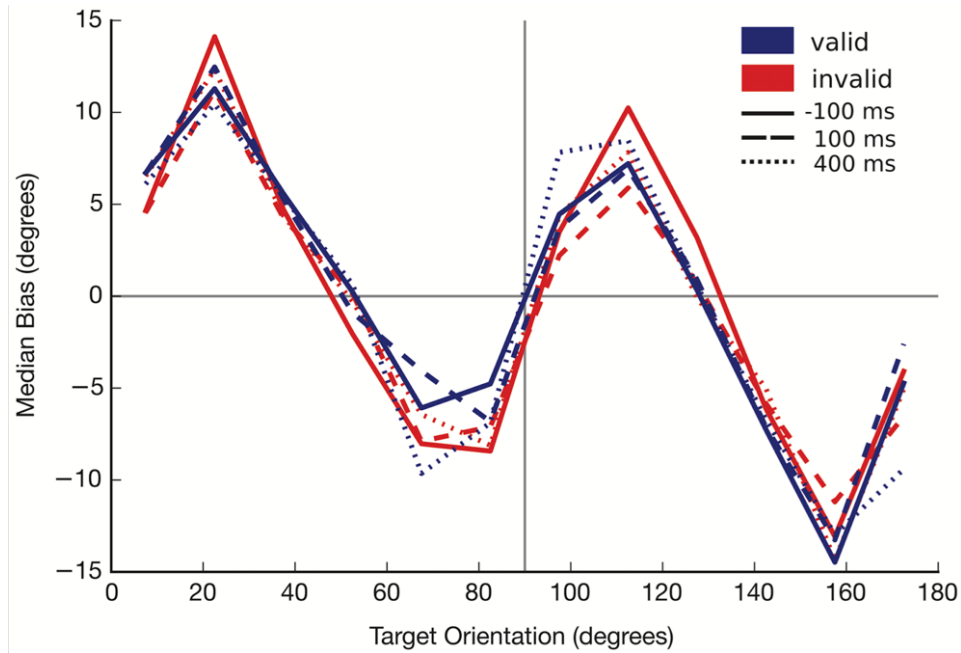
Here we chose an experimental setting in which, by construction, behavioral report should faithfully reflect conscious perception. In iconic memory experiments, by contrast, the link between report, conscious access and conscious perception is less straightforward and still very much debated (Block 2007; Cohen & Dennett, 2011). In these experiments the number of

items, and more generally the complexity of the display, exceeds working memory capacity, thus opening the possibility of a dissociation between what is perceived and what can be reported. The beneficial effect of post- cueing in iconic memory experiments has been taken as an indication of such a dissociation: since what is reported can be flexibly influenced by a post-cue, this might indicate that the initial percept is richer than the subset that is extracted for report. This interpretation of the iconic memory phenomenon has become a core argument in favor of the existence a form of conscious perception, called phenomenal consciousness, which may greatly exceed the scope of details available to conscious access (Block, 2007; Lamme, 2006). However, this interpretation relies on the assumption that what comes after the stimulus cannot induce conscious perception of elements that were not initially perceived consciously. The current results show that this assumption is not supported: here a discrete transition to conscious perception was induced by orienting attention after the stimulus had disappeared. As such, one cannot exclude the possibility that retro-perception mechanisms are also at play in iconic memory experiments, and hence account or partly account for the beneficial effect of post-cueing, as suggested by alternative interpretations of the iconic memory phenomenon (Kouider et al., 2010).

In summary, the present results show that conscious access displays a discrete component, and that attentional cueing can gate this discrete transition to conscious access and conscious perception, even after the stimulus is gone. This provides strong evidence for the hypothesis that attention plays a crucial role in conscious perception.

Supporting Information

Angular Bias



S1 Fig. Bias profile and oblique effect. The graph represents the median of signed angular error relative to the target's true orientation for each target orientation and each experimental condition averaged across participants.

While the absolute angular error gives us an estimate of the dispersion of the errors, the median of signed errors indicates the center of the error distribution. When this center is 0, it means that there is no bias in the perception of the target's orientation. Here we see the classical "oblique effect" bias as a deviation from 0 for target orientations close to the horizontal or vertical. This oblique effect profile did not vary significantly across experimental conditions.

Formulae

1. Bessel equation of order v :

$$I_v(z) = \sum_{k=0}^{\infty} \frac{\left(\frac{z^2}{4}\right)^k}{k! \Gamma(v+k+1)}$$

where Γ is the gamma function.

2. The conversion of the precision parameter κ to standard deviation (SD) follow the equation:

$$SD = \sqrt{\frac{1}{\kappa}}$$

Chapter 3: Retroperception in iconic-memory conditions

Introduction

The previous experiment (hereafter referred to as *Study 1*) suggests that an attentional manipulation can exert influence upon the encoded representation of a visual stimulus, granting access to the machinery of consciousness. To wit, behavioral responses (in effect, the reproduction of a previously-displayed stimulus) are best described as an aggregate of two trial types: trials in which the subject has access to some nonzero amount of information about the target's orientation — assumed to reflect conscious access to the stimulus — and trials in which he does not. As such, the main finding of Study 1 is that so-called “valid” cues, which direct attention towards the region of a display in which a low-contrast Gabor patch was previously displayed, reduces the number of trials falling into the first category as compared to an “invalid” cue. In terms of the above model, this dissociation appears as a reduction in the uniform component, which corresponds to a subject's guessing rate (which in turn is assumed to reflect the number of trials where the subject did not see the target consciously). These results validate our primary hypothesis, which states that attention triggers conscious perception on trials where the target would otherwise have been missed. In this proposition, the target is not always consciously accessed, and thus not always consciously seen following its presentation. It does, however, leave a sensory trace in the sensory cortex, so we propose that the retro-cue promotes these traces to be consciously accessed.

The results of Study 1 raise questions about whether retrospection might partially generalize to a classical iconic memory paradigm. Could a portion of the post-cue's effect on recall precision be driven by perceptual (as opposed to mnemonic) processes? The answer to such a question is to be found in the model comparison procedure from the previous study. If retroperception is at play, then we would expect either that data from subjects are best described by an EP+g model (indicating that these subjects are unaware of a subset of the array), or that the

BIC value in the comparison between EP+g and VP models is quite small, which in turn would indicate that both models account for a significant portion of the total variance. If, however, retroperception is not involved in the Sperling paradigm and its derivatives, then we would expect the data to be best described by a model that assumes the contents of iconic memory is always consciously perceived, and therefore doesn't account for informationless guesses.

In typical iconic memory experiments, a display containing several high-contrast items such as a matrix of capital letters (Sligte, Scholte & Lamme, 2008; Sperling 1960), is briefly presented to a participant. Absent any experimental manipulation, observers are able to accurately report the identity or features of a subset (usually three or four) of these targets. However, cueing attention to one specific subset of targets within one second of SOA will dramatically improve the accuracy with which these stimuli are reported, and because the targets in question are highly-visible, this improvement is attributed to the limited capacity of working memory. The classical interpretation of this effect is that all targets are initially perceived — consciously — and the cue serves only to increase the odds that a given object is transferred from the ineffable, volatile sensory store of iconic memory into the reportable, durable store of working memory. Under this classical interpretation, there is no *a priori* change in conscious state, so we predict that EP+g mixture-model applied to an iconic-memory paradigm employing stimuli similar to Study 1 would show no effect of SOA on the guessing rate. Instead, due to the high memory load of a multi-target array, we predict a significant shift in the recall precision due to the progressive deterioration of perceptual representations in iconic memory.

To reiterate, Study 2 consists of an exploratory study into the involvement of retroperception in classical iconic memory tasks. We expect the involvement of this phenomenon to become apparent in the model-selection stage of analysis.

Materials and Methods

Participants

The number of participants was fixed to twenty prior to the experiment, based on our previous observation and replications of the retro-perception effect (Sergent et al., 2013; Thibault et al., 2016). Twenty participants between the ages of 18 and 32 took part in the study, each exhibiting normal or corrected-to-normal vision. Due to hardware failure, data from two participants were lost, reducing the total subject pool to 18 individuals. The remaining subjects (13 women, 7 men) had an average age of 24.59 years \pm 3.7. All participants gave informed consent in writing prior to participation, and the Université Paris Descartes Review Board, CERES, approved the protocols for the study in accordance with French regulations and the Declaration of Helsinki. Participants received a compensation of 10€ per hour for their time.

Apparatus and Stimuli

Stimuli were generated and responses recorded using the Psychophysics Toolbox for Matlab (Brainard, 1997). Stimuli were presented on a CRT monitor (Sony Trinitron GDM-F520). Refresh rate was 60 Hz and screen resolution was 1280 by 1024 pixels. Participants were seated 60 cm away from the monitor, in a dimly lit room. Eye fixation was monitored and recorded using an Eyelink 1000 (SR Research Ltd., Osgoode, Ontario, Canada). We verified that subjects maintained fixation during the majority of trials. This was determined by counting the number of trials during which the mean fixation exceeded 1 degree of eccentricity from the central fixation point (the border of the placeholders were at 3 degree on each side). On average, subjects exceeded this threshold on less than 1% of the trials.

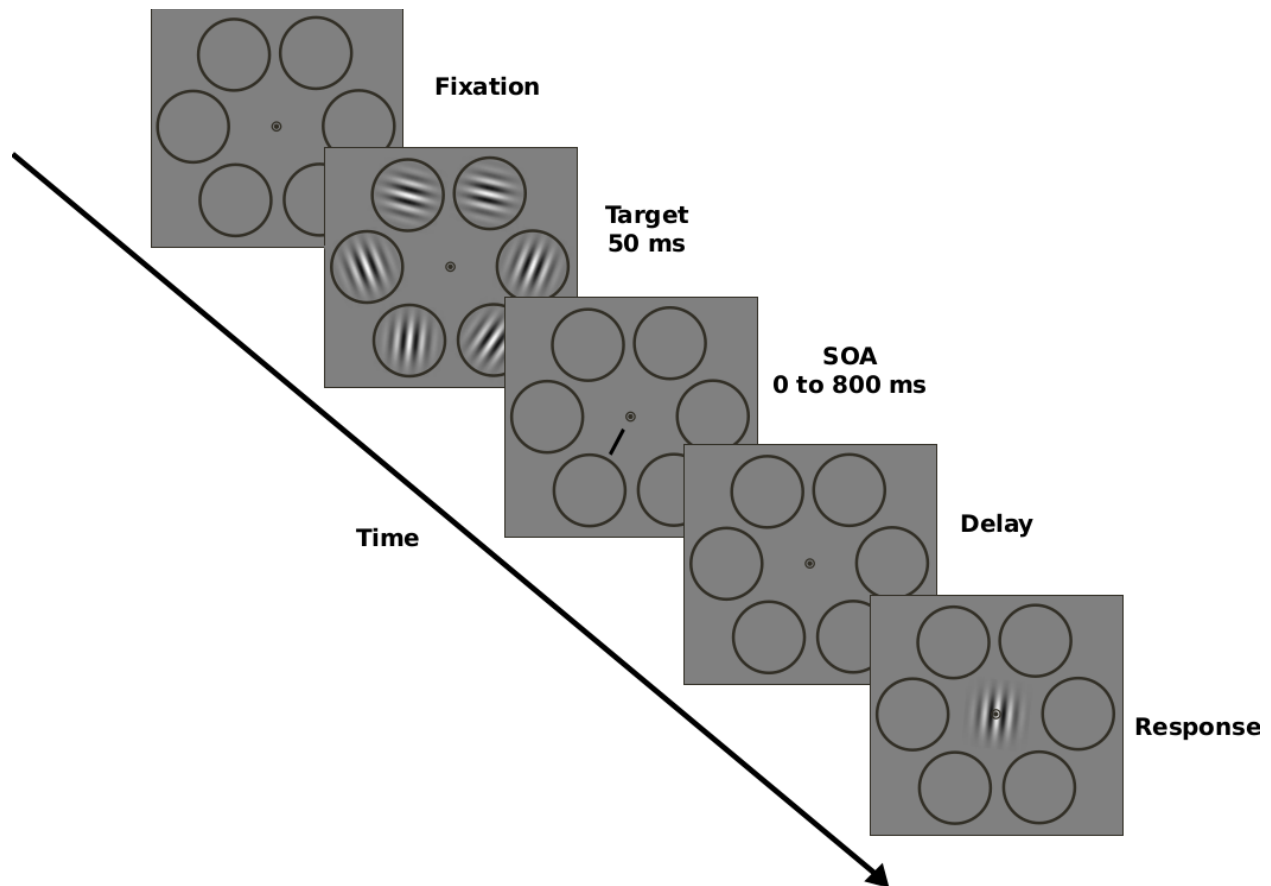


Figure 3.1. Experimental design. A target appeared in each of the circular placeholders and one such target was followed by a concurrent or post-cue in the form of a black bar that briefly pointed to the target of interest. Subjects reported the orientation of the target using the central Gabor patch presented at the end of the trial. *N.B:* stimuli are not drawn to scale on this representation.

Stimuli (Fig 3.1) were presented on a gray background (12 cd/m²), and participants were told to fixate a small black circle at the center of the screen ($.6^\circ$ in diameter). Six larger black circles (2.4° in diameter) were always present at the angles of an invisible equilateral hexagon centered on the fixation point. The eccentricity of these black circles relative to their centers was 4° from central fixation. As in Study 1, these circles served as placeholders that indicated the area of the display in which target Gabor patches would be presented. Contrary to the previous study, however, they did not provide a means for attentional cuing. Cueing was instead achieved by means of a black bar 1° in length that appeared between the placeholder and the fixation point, superimposed upon the invisible line that connects the two, and appearing for 50 ms.

Targets were again Gabor patches subtending 2° in diameter (2 cycles per degree with a randomized phase; Gaussian envelop with 1° full width at half maximum) and were presented in one of twelve orientations spanning 7.5° to 172.5° in increments of 15° . The contrast of the target was fixed at 100% for all targets.

Each trial began with the onset of a dot at the center of the fixation circle. Following a random delay between 500 ms and 900 ms, six Gabor patches were presented for 50 ms, one in each placeholder. Exactly one of these targets was selected at random to be cued at one of four possible SOAs (0, 100, 400 or 800 ms) and all trials contained a cue. A response screen appeared following a 500 to 900 ms delay relative to the offset of targets, containing a response Gabor patch presented at fixation and subtending 2 degrees of visual angle that subjects used to reproduce the cued target's remembered orientation. Its parameters were the same as the target except that its sinusoidal phase was fixed at .5 radians, and its initial orientation was selected at random from a uniform distribution ranging from zero to 180° .

The participant's task was to reproduce the remembered orientation of the target by freely and continuously varying the orientation of the response patch using the mouse. A small black dot above the response patch indicated the mouse position on the screen. Subjects were not limited in their response times. They indicated their final choice with a left click. Subjects were instructed to always provide a response, and guess in the event that they had not seen the target. Feedback was provided at the end of each block in the form of percentage of hits (a response deviating more than 45° relative to the target orientation was considered as a “miss”).

Analysis

Overview of analysis steps

The present study extends the results of Study 1 with two principal objectives. Firstly, we seek to determine whether the mixture-modelling strategy employed in Study 1 is capable of detecting changes in recall precision under classical iconic-memory conditions. Under such

conditions, a load on the detection system (i.e. low-contrast) is traded in favor of a load on the working-memory system (i.e. multiple targets), meaning that all targets should be visible *in principle* while an individual's ability to recall a particular target is a function of set-size. In order to compare the effects of retro-cueing in iconic memory (Study 2) with the effects of retro-cueing in a retroperception protocol (Study 1) we employ identical methods with respect to bias correction, model fitting, and model comparison unless otherwise noted.

Results

Model-free analyses

As the SOA between target and cue increased, the mean absolute response error also increased, $F(3, 51) = 79.67, p < .001$. This was expected: it mirrors the well known decline in performance with the delay of post-cueing in iconic memory experiments. Modeling of these response errors was further performed in order to dissect the origin of this decline.

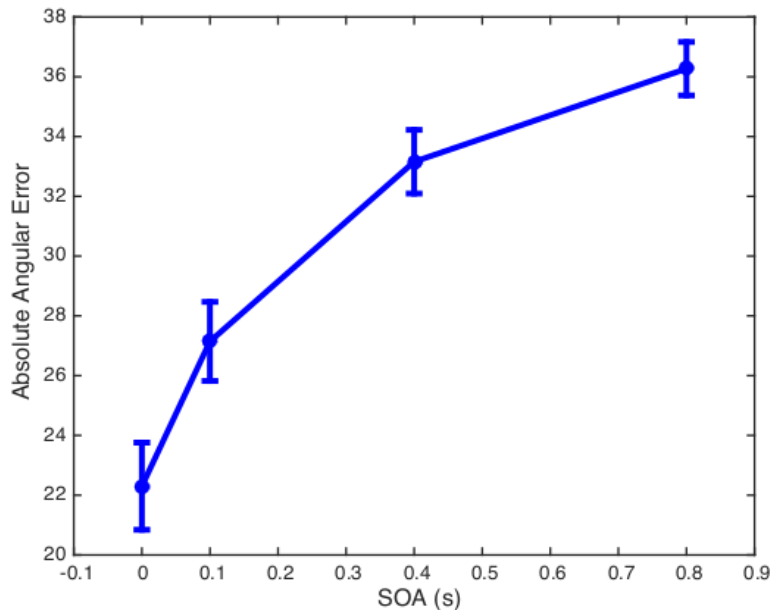


Figure 3.2. Absolute angular error as a function of SOA. Increasing the SOA between target and cue apparition reduces the precision in the orientation matching response, as evidenced by an increase in the average absolute value of angular response-error.

Model selection

As with Study 1, we fit 4 different mixture models to the data of each subject. For each of these 72 model fits (18 subjects times 4 models), we again computed the Bayesian Information Criterion (BIC) taking into account the number of free parameters (Schwarz, 1978). When considering individual fits, no single model provides a consistently better fit for all subjects in all conditions (see Table 3.1). We observe that the equal-precision-without-guessing (EP) and variable-precision-with-guessing (VP+g) models perform relatively poorly, that is, they rarely (if ever) outperform the other models.

Preferred Model

<i>SOA</i>	<i>EP</i>	<i>EP+g</i>	<i>VP</i>	<i>VP+g</i>
<i>0 ms</i>	1	11	6	0
<i>100 ms</i>	0	12	6	0
<i>400 ms</i>	1	14	3	0
<i>800 ms</i>	0	13	4	0

Table 3.1. Preferred model for each subject. For each subject, each of the four models was fitted to data from each of the four possible conditions (i.e. SOA). Each fit is accompanied by a Bayesian Information Criterion (BIC) values for which a lower magnitude indicates a better model fit (including a correction for the number of parameters in a given model). The preferred model is the model for which the BIC is lowest. Table values represent the number of subjects preferring a given model for each SOA, revealing a strong preference for the EP+g model (as in Study 1), followed by a preference for the VP model. In two conditions, a subject preferred the EP model, though the difference in BIC relative to VP was marginal.

The remaining EP+g and VP models provide mixed results, but it is highly unrealistic to assume that the process by which a subject generates his response varies across experimental conditions. In the next step of the analysis we thus make the assumption that the same underlying process is at work across all SOAs. We therefore collapse our data across conditions and compare the model fits across the aggregate whole for each subject. Table 3.2 shows the relative BIC values for each model in each subject across conditions of SOA. We observe that for 12 out of 18 subjects, the EP+g model emerges as a winner, as evidenced by the relative BIC of zero. Further, in cases where the VP model defeats its EP+g counterpart, the BIC difference is of modest size (mean= 9.79 ± 6.98), suggesting that the two models might be approximately equal in their predictive power.

Relative BIC Scores

<i>Subject</i>	<i>EP</i>	<i>EP+g</i>	<i>VP</i>	<i>VP+g</i>
<i>1</i>	10.13	0.12	0	6.90
<i>2</i>	359.21	0	35.24	7.01
<i>3</i>	161.80	0	26.71	7.05
<i>4</i>	104.51	5.65	0	6.64
<i>5</i>	88.64	0	5.76	6.79
<i>6</i>	296.91	0	26.91	6.70
<i>7</i>	326.86	0	42.37	7.05
<i>8</i>	262.10	0	45.60	7.05
<i>9</i>	160.67	0	9.28	5.97
<i>10</i>	171.95	16.72	0	6.99
<i>11</i>	116.41	0	2.29	4.88
<i>12</i>	398.94	0	69.27	7.05
<i>13</i>	251.80	9.87	0	2.96
<i>14</i>	400.98	0	52.80	7.05
<i>15</i>	94.73	18.75	0	7.05
<i>16</i>	227.41	0	14.77	6.07
<i>17</i>	201.67	7.63	0	3.84
<i>18</i>	307.64	0	51.97	7.05

Table 3.2. Preferred model for each subject. For each subject, the model with a relative BIC of zero is the winning model. The above table shows that for the majority of subjects, the EP+g model is the winning model and that in cases in which the VP model is the winner, the difference in BIC with EP+g tends to be marginal.

Similarly to Study 1, data from a majority of subjects was best explained by a model that assumes a fixed perceptual precision across trials, but allows for subjects to flat-out guess the orientation of unseen targets. Compared to Study 1, the results of the model selection analysis are less clear. Whereas the previous study revealed one subject with a marginal preference for the non-predicted model (VP in Study 1 with a BIC of .06), the present study reveals six "minority" subjects with a moderate preference for the VP model, compared to the EP+g model.

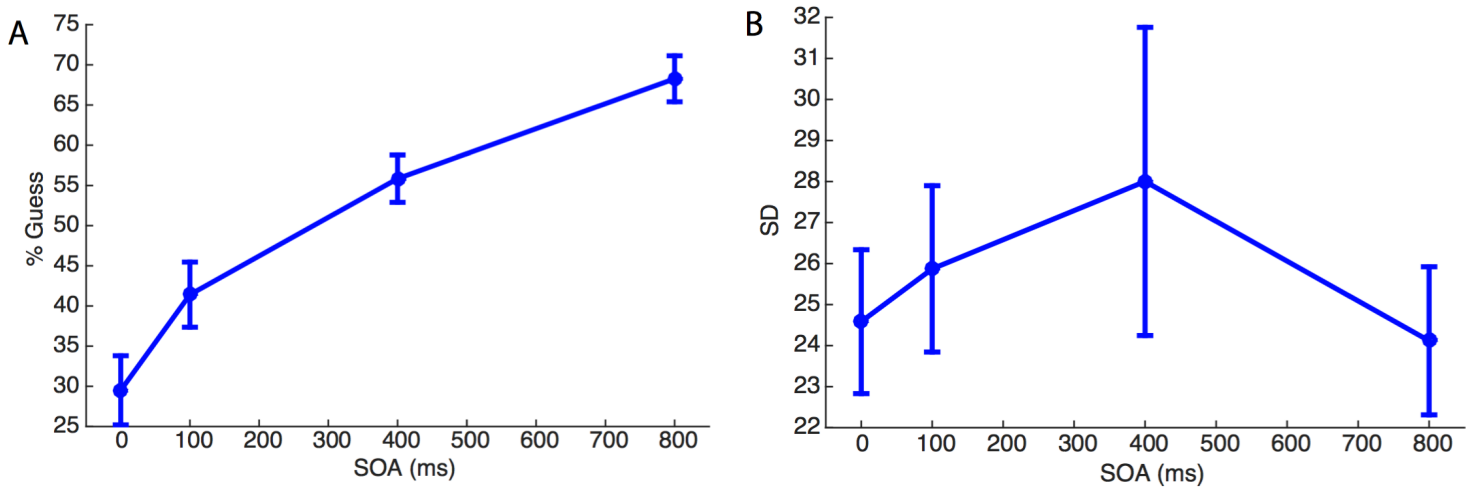


Figure 3.3. EP+g model parameters as a function of SOA. Percentage of guessed trials as a function of SOA (A) and mean precision of response as a function of SOA (B).

It is possible that there are some subjects for which a phenomenon like retroperception is taking place (the subjects for which EP+G is better) and some other subjects for which the classical iconic memory interpretation is more probable (VP subjects). These two type of subjects might have developed different strategies. Imagine, for instance, that a subject applies the full extent of his attentional resources on a subset of targets S . By virtue of the maximal contrast of the Gabor patches, all of S would be immediately encoded in working memory, and be available for recall with minimal degradation when the subject is prompted for a response. As such, we might expect that when said subject were prompted to report the orientation of a target not included in S , he would guess, mimicking the results of Study 1. If, however, another subject were to “spread” his attention over the entirety of the scene, the strategy more closely resembles that which is tacitly assumed in iconic-memory experiments. Indeed, the large number of high-

visibility targets implies contention over working memory resources, and thus, the attentional cues would bias which targets were granted access. An important corollary of this putative system is that the cue SOA would thus control the amount of signal degradation that occurs before a target is placed into WM, and thus we would predict a systematic decrease in recall precision as a function of SOA.

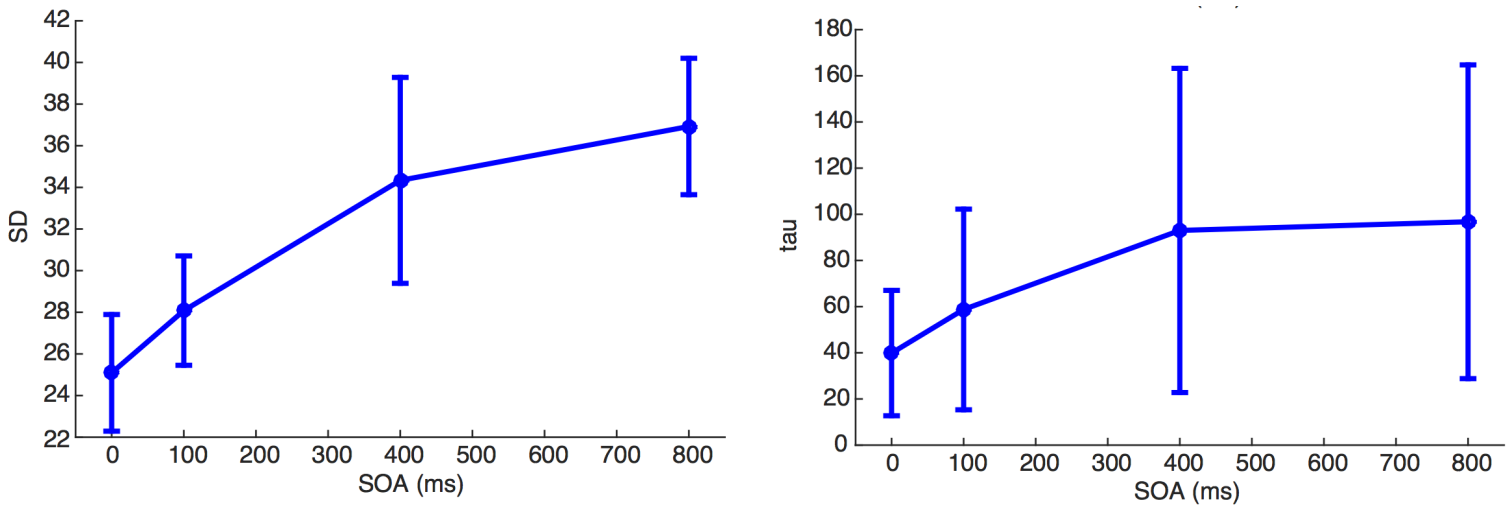


Figure 3.4. VP model parameters as a function of SOA. Mean precision of response as a function of SOA (A) and variability of precision between trials (B).

In addition to the standard deviation, which represents recall precision, the VP model also presents a tau parameter, which represents the variance in recall precision across trials (and thus, in effect, across items of the display). With regards to tau, it is possible that the variance of recall precision across items increases with time, i.e. the more we wait, the more the subject will be biased towards certain items and not other. All in all, the VP model accounts for more variance across subjects, and thus we select the VP model for further analyses.

Precision versus guessing for competing models

Having identified the best model for most of our subjects (the standard mixture model, i.e. equal precision with guessing), we investigated whether target-cue SOA influenced the proportion of guesses and the precision of orientation-matching. We again report that an increase

in SOA produced an increase in the guessing rate, $F(3, 51) = 50.43$, $p < .001$, see Fig 3.3. Similarly to Study 1, variance in the SOA did not translate into a systematic shift in recall precision, as evidenced by the absence of a significant effect on the standard deviation of angular error, $F(3, 51) = .77$, $p = .46$.

When we restrict the same analysis to only those subjects for whom the EP+guess model was preferred, we obtain similar results for guessing rate, $F(3, 15) = 9.97$, $p < .05$, and recall precision, $F(3, 15) = .67$, $p = .51$.

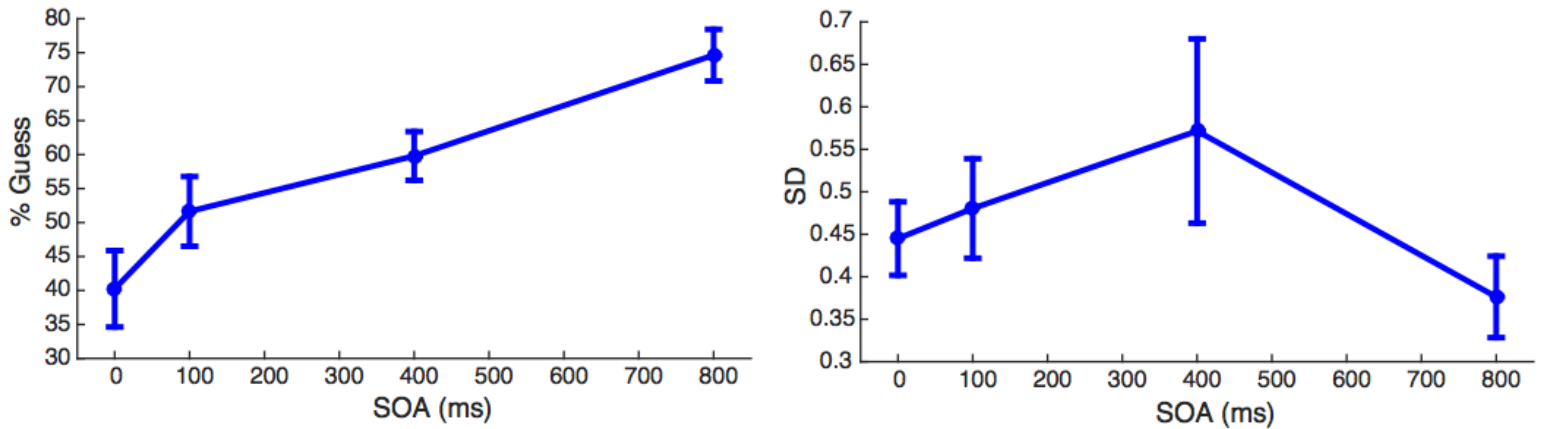


Figure 3.4. EP+g model parameters as a function of SOA for subjects preferring said model. Percentage of guessed trials as a function of SOA (A) and mean precision of response as a function of SOA (B) for subjects whose behavior is best described by the EP+g model, as determined by BIC.

As previously mentioned, the data suggests that a subset of subjects may be employing a radically different strategy for completing the task. These data are supported by the assertion that increasing SOA also increases the standard deviation (that is, *decreases* the precision) of the angular error reported by the subjects during an orientation matching task, $F(3, 51) = 7.77$, $p = .001$. Readers should note that the reported effects were computed using data from all subjects, including those whose behavior was best explained by the EP+g model, since the subset of subjects exhibiting VP-like behavior is too small to produce reliable statistics. This lack of statistical power is perhaps responsible for the null effect we find in tau, both over the full cohort and within the subset of VP-like subjects.

Discussion

The effects of post-cueing on objective report in iconic-memory settings have historically been interpreted as a biasing of some selective process that transfers stimulus representations from iconic to working memory. Implicit in this interpretation is the strong assumption that all stimuli in the display are — at least initially — consciously perceived by the observer (Sperling, 1960). Study 1 exposes an alternative mechanism by which post-cueing can improve objective report: a sort of perceptual resuscitation in which the subject suddenly becomes aware of a stimulus he did not consciously access at first sight. The existence of this “retroperception” phenomenon invites us to revisit long-held assumptions about the mechanism of attentional post-cueing in classical iconic-memory paradigms. While results from Study 1 point to the existence of retroperception as a phenomenon, its design precludes us from drawing any conclusions about the involvement of retroperception in traditional iconic-memory tasks. Indeed, Study 1 achieves its objectives by ruling out effects of working-memory capacity and placing a heavy strain on the perceptual system, in effect inverting the characteristic manipulations of Sperling-esque paradigms. In Study 2, we trade the high-perceptual load (i.e. low contrast) for a perceptually easy task and a mnemonically hard task (i.e. large set-size) for a more challenging analog. In doing so, we explore the extent to which data from iconic-memory tasks are explained by the EP+g model, which would suggest conscious-access-driven effects, versus the VP model, which would suggest behavioral effects stemming from working-memory capacity. We report evidence in favor of both models, which raises the question of whether attentional strategies differ between subjects. Though this finding contradicts some recent findings (van den Berg, Awh & Ma, 2014; van den Berg, Shin, Chou, George & Ma, 2012), it is important to note that the behavior of a non-negligible portion of subjects is better explained by a model that assumes both that perceptual precision might vary across trial and that subjects always consciously perceive the target (that is, their responses are always informed by some information, i.e. there are no guesses). While the exact source of this difference requires further investigation, the above studies manipulate color rather than polar orientation, a visual feature which is known to be

integrated pre-attentively (Theeuwes, 1992). In the present experiment, we hypothesize that our variable results may stem from a difference in attentional strategy between subjects.

Attention is known to be heterogeneously allocated in space under a wide variety of conditions, both voluntarily (Giordano, McElree, & Carrasco, 2009; Kinchla, 1980; Mangun & Hillyard, 1990; Sperling & Melchner, 1978) and not (Cheal et al., 1991; Giordano, McElree, & Carrasco, 2009; Jonides, 1981; Nakayama & Mackeben, 1989). As such, one might imagine several distinct strategies for attending to the target display. This could explain why there was no single model that could account for the behavior of all subjects. In the present experiment, the EP+guess model was the best model in a majority of subjects. What was the attentional strategy of these subjects? One possibility is that these subjects used a strategy which attempts to maximize response precision by sacrificing all information about certain targets in favor of more detailed information about others. By focusing one's attention on a subset of targets such that all stimuli are encoded into working memory, a subject could produce behavior resembling the retroperception reported in Study 1. We might predict that said subject gleaned no orientation information from unattended targets, and therefore responds with a guess when said target is cued for a response (Mack & Rock, 1998; Simons & Rensink, 2005; Simons & Chabris, 1999; Simons, 2000; Scholl et al., 2003; Newby & Rock 1998; Most, et al., 2005; Most et al., 2001).

Another possibility is that what was counted as “guesses” in the EP+g model were actually “misbinding errors”: the subjects reported the orientation of one of the items next to the cued target (Zokaei et al., 2014).

While the above strategy maximizes the odds of having *all* orientation information about the selected target, another approach to the behavioral task might be to ensure that *some* information is available for all targets. Understanding that each target exhibits an equal probability of being cued for response, a subject might attend to the display as a whole, trading the precision with which they encode each target for a guaranteed non-zero quantity of information about all targets (Luck & Vogel, 1997; Linke et al., 2011; Sims et al., 2012; Ma et

al., 2014). In such cases, it stands to reason that all targets are consciously perceived, and that the resolution of the encoded percept degrades as a function of time, as observed in the VP model fits. Individuals executing this strategy never guess, as they always have some indication of the targets orientation, and their behavior is thus accounted for by the VP model, whose SD parameter is positively correlated with SOA.

As previously mentioned, we had no strong predictions with respect to the VP model's tau parameter, it is worth noting that the number of subjects whose preferred model was VP was possibly too small to capture a small effect of SOA on tau. We indulge, however, in some speculation about what a future study might reveal with regards to such effects, since the presence significant effects on this variable would comfort us in the notion that a VP process is at play. Recall that tau represents the variance in perceptual precision across trials. In the simplest case, the source of this variance is fully random, for instance stemming from a systematic attentional bias towards a region of the display (Giordano, McElree, & Carrasco, 2009; Kinchla, 1980; Mangun & Hillyard, 1990; Sperling & Melchner, 1978; Cheal et al., 1991; Giordano, McElree, & Carrasco, 2009; Jonides, 1981; Nakayama & Mackeben, 1989), coupled with the pseudorandom selection of a target for post-cueing. In such a case, we would expect no effect SOA on tau, as with the present study, but we should see a systematic decrease for targets located in the preferred attentional fields. If, on the contrary, we were to observe a relationship (positive or otherwise) between tau and SOA, whence might it come? One possibility is that it might reflect some difference in encoding-fidelity of targets into iconic memory. Consider, for example, a case in which attention were unevenly (albeit stochastically) distributed across the display. In such a case, certain targets would be better encoded than others, simply because of the conjunction of the target's position and the current attentional topology. If we also suppose that the rate of perceptual decay is a partial function of encoding fidelity — that is, high-fidelity representations decay at a slower rate than their low-fidelity counterparts — then we would observe an uneven rate of decay when comparing high versus low-fidelity targets (formally, we would see this as an interaction between encoding fidelity and SOA). Though hypothetical, these effects would paint a more detailed picture of the attentional, mnemonic and conscious

dynamics at play during what has hitherto been described as an “iconic memory task”. We recommend pursuing such follow-up studies with the goal of determining whether or not this name is well-suited to the phenomenon being probed.

Finally we would like to make a remark on the sensitivity of mixture-model analysis to fluctuations in the fidelity of perceptual encoding — we did not find any significant effect of SOA on the *SD* parameter in EP+g model. This could reflect the fact that most of the post-cueing effect is carried by the guess rate, in the absence of any substantial decay of the representations with time, but this could also be due to a lack of sensitivity of the model to variations in SD. This could be tested by applying the same methodology and models to a situation where it is certain that only the precision of information is varied with the experimental manipulation. For example, reproducing the orientation of well contrasted Gabors at fixation, presented for various durations.

Chapter 4: Consciousness and representational change, an fMRI study

Introduction

The previous two chapters outline a suppositional mechanism through which attention promotes a sensory representation to awareness. These results lend credence to models of consciousness that describe a process in which stimulus representations held in sensory cortex are broadcast throughout a “global workspace” composed of a network of high-level cortical areas (Dehaene et al., 2006; Sergent & Naccache, 2012; Dehaene, Sergent & Changeux, 2003). These models draw a distinction between pre-conscious stimuli, which are salient and well-encoded enough to elicit a conscious percept, but which have not been granted access to the workspace by an attentional gatekeeper, and those which enjoy the full benefits of conscious perception (Dehaene et al., 2006). In theory, nothing distinguishes the representations of pre-conscious and conscious stimuli with regards to their composition, and the selection process, as evidenced by myriad studies in which consciousness is manipulated without modifying the stimulus of interest (Raymond et al., 1992; Broadbent and Broadbent, 1987; Sergent & Dehaene,

2004; Marti et al., 2012; Kim & Blake, 2005; Leopold & Logothetis, 1999; Strezer et al., 2009; Breitmeyer & Ögmen, 2006; Dehaene et al., 2001; Del Cul et al., 2007; Albrecht et al., 2010). Of course, it is reasonable to assume that the intervention of attention somehow alters the preconscious percept during the process of rendering it conscious. One might therefore wish to investigate (1) whether such a change effectively takes place and (2) the nature of the change.

Analyzing the internal representations of sensory stimuli is largely a question of inferring a mental state on the basis of neuroimaging data, and such an analysis involves a unique set of challenges. At the time of writing, the bulk of neuroimaging studies in consciousness research have relied on contrastive analyses, which works by comparing conditions in which a subject is systematically conscious of a stimulus with conditions in which he/she is not (Baars, 1994; Dehaene et al., 2001; Kjaer et al., 2001). Because the principle of careful experimental control dictates that stimuli across conditions should be kept as simple as possible, there is an inherent tendency for neurological (e.g. BOLD) responses to covary along with the features of said stimuli. A problem therefore arises when trials are pooled within conditions; the data being analysed represent a uniform amalgamation of non-independent sources of variance, which cannot be decomposed and analyzed individually (Haynes & Rees, 2006; Kamitani & Tong, 2006). To make matters worse, popular statistical methods such as statistical parametric mapping (SPM) assume linear relationships between independent and dependent variables (Haynes & Rees, 2006; Sandberg et al., 2014), whereas a great many perceptual effects — particularly those whose locus is in primary sensory cortex — are best predicted by nonlinear (e.g. sigmoidal) models (Knierim & van Essen, 1992; Kapadia et al., 2000; Gilbert et al., 2000). Luckily, a new class of analyses, collectively referred to as *multivariate pattern analysis* (MVPA) address precisely these issues, gracefully accounting for intrinsic covariance between electrodes or voxels, and exhibiting the ability to model arbitrary non-linear processes, with the trade-off of potential overfitting (Sandberg et al., 2014). Judicious use of MVPA demonstrably improves sensitivity to statistical effects in neuroimaging data (Haynes & Rees, 2006, Norman et al., 2006), for instance improving the degree to which voxels from V1, V2 and V3 could predict the orientation of a square-wave grating during a binocular rivalry task relative to the mean

BOLD activity across all voxels in these regions (Haynes & Rees, 2005). To boot, multivariate tests are relatively insensitive to differences between conditions that appear only during certain trials, and show an increased sensitivity to systematic differences between experimental conditions (Sandberg et al., 2014). This is because contrary to univariate analyses, MVPA estimates a model on a trial-by-trial bases, and as such is able to disentangle systematic sources of (co)variance from stochastic ones. Haynes (2009), further proposes that MVPA, by virtue of its ability to estimate arbitrary nonlinear functions from a sample of the data they produce, are theoretically capable of fully-decoding any stimulus from neuroimaging data given sufficient temporal and spatial resolution in the recording. This in turn implies that an increase in MVPA performance is synonymous with an increase in information in the data from which it decodes, and that as such, MVPA analyses are a valid indicator of the quality of internal representations (Kamitani & Tong, 2006; Sandberg et al., 2014). This is a particularly appealing property for the problem at hand, since we endeavor to track changes in the informational content and representational fidelity of brain activity evoked by conscious and nonconscious percepts.

Using Gabor-patch stimuli similar to those employed in studies 1 and 2, we examine the degree to which conscious perception alters the performance of a support vector classifier (SVC) in the context of MVPA. We first predict that our classifier's performance will follow a sigmoid law with respect to increasing contrast, mirroring subject performance on an objective performance task. From there, we expect that if awareness does indeed alter the contents of a perceptual representation, classifier accuracy on those trials in which the subject reports having experienced the sight of a target to increase relative to those in which he indicates that no such percept occurred. Such an effect would provide strong evidence that consciousness, by way of the signal-amplification and stabilisation effects of the attentional system, modifies the perceptual representation of a stimulus.

Methods

Participants

Participants for the present study were recruited in two phases. First, 26 individuals with normal or corrected-to-normal vision were recruited. After giving their informed consent, each subject from the pool performed a pre-test (slightly modified version of the experimental task) outside of the scanner, where we tested the psychometric curve of each participant for 6 preselected contrast values (2%, 3%, 3.5%, 4%, 4.5% and 12%). The final cohort was selected based on this training. Participants were retained if the full rising part of their psychometric curve (from chance to perfect performance, going through threshold) was present within our predefined bounds. In total, 19 participants (8 male, 11 female; ages 22 to 36, mean 27.5 SD 3.19) passed selection and were included in the study. The experiments were conducted under the ethics committee agreement INSERM C10-47.

Stimuli

The stimuli for both the pre-test and the main experimental session, created using custom software built on top of the PyGame SDL wrapper for Python, were identical. Stimuli were presented on a gray background (12cd/m^2) while participants fixated a small black circle at the center of the screen ($.6^\circ$ in diameter). Two large black ellipsoids (3° in width and 6° in height) were always present bilaterally, with their centers positioned 5° to the left and right of the central fixation circle, and served as placeholders for the two possible target positions.

Targets were ellipsoid Gabor patches that occupied the space within the placeholders, and were characterized by three constants: the width of their gaussian envelope (1.5° in width and 2° in height) and the frequency and phase of the sinusoidal component (2.5° per cycle, i.e. 0.4 cycles per degree with a randomized phase; Gaussian envelop with 1° full width at half maximum).

Pre-Test Session

Apparatus

The selection pre-test took place outside of the scanner, in a dimly-lit room equipped with a chin-rest, which was fixed 60 cm away from the CRT Monitor (Sony Trinitron GDM-F520). The refresh-rate on said monitor was set to 60 Hz and the screen resolution was of 1280 by 1024 pixels. Participants provided their responses via the i and j keys on a French AZERTY keyboard, mimicking the digit-mappings on the FORP controllers in the fMRI.

Procedure

The pre-test consisted of 8 blocks of 48 trials, identical in form and content to the main experimental task, albeit with shortened trials. Each trial began with the a variable-length warning interval spanning .9 to 2.7 s, whose last frame presented the target within either of the two placeholders (duration: 16.667 ms). This was followed by a delay of 1.8 s. Both target orientation (45° and 135° from the vertical meridian) and target contrast were selected at random and counterbalanced within blocks, yielding 12 trials per contrast and 24 trials per orientation, for each block. During the ensuing response interval of 3.6 seconds, the first response screen appeared at most for 1.8 seconds, asking for orientation discrimination, before giving way to the second response screen, asking for subjective visibility, again within 1.8 sec.

The first response collected was the so-called objective measure, in which two small black line segments oriented at 45° and 135° appeared $.4^\circ$ above and below the central fixation circle. Participants were instructed to press the key corresponding to the segment that shared the same orientation as the target, and the position of each segment was randomized across trials such that participants could not predict whether correct response would appear above or below the fixation circle.

The second response covered the participant's subjective impression of target visibility. An integer ranging from zero to 7 was selected at random and presented within the central fixation circle, whereupon subjects were to rate the visibility of the target by respectively incrementing or decrementing the value with the i and j keys, reserving the value zero for cases

in which the target was not perceived at all and the value 7 for trials in which they had very clearly seen both the target and its orientation. At the end of the response interval, the value displayed was automatically submitted as a response without further user intervention.

Scan Sessions

Apparatus & Acquisition Parameters

Anatomical and BOLD data were acquired using a MAGNETOM Trio 3T system (Siemens, Erlangen, Germany). Functional images were acquired using a T2*-weighted gradient-echo EPI sequence (FA = 62°, FOV = 68 x 68 x 42 voxels, voxel size = 3 x 3 x 3 mm, TR/TE = 900/27 ms).

Visual stimuli were projected onto a screen situated at the entrance of the MRI tunnel, 120 cm away from the periscopic mirror above the participant's eyes, at a resolution of 1280 by 1024 pixels.

Procedure

Experimental Blocks

Experimental blocks were identical to those in the selection pre-test, save for their altered time-course and contrast values.

The present study followed a slow event-related design in which trial duration ranged from 9 to 10.8 seconds (10 to 12 TRs). An initial warning period lasted one to three TRs (.9 to 2.7 s), as before, with the target being presented during the last frame of this time interval. Target offset was immediately followed by a 5-TR-long (4.5 s) delay before the 4-TR-long (3.6 s) response interval appeared, polling participants for both objective judgements of orientation and subjective visibility responses.

Four possible contrast levels were selected from the original 6 present in the pre-test individually for each subject based on his/her performance on the task, save for one subject for whom a 6% contrast level was preferred. This, along with the two possible target orientations and the two possible target positions, produced a 2x2x4 design with 24 trials per condition in

total. For three participants, these contrasts were adjusted between blocks in order to better sample from the psychometric function.

Additional Localizer Blocks

Following the eight experimental sessions, participants underwent a localizer scan, during which they were presented with 64 blocks of 11.9 seconds of full-contrast targets flashing at a rate of 4.3 Hz, counterbalanced across each combination of target orientation and target position (16 blocks per position-by-orientation condition). Participants were instructed to monitor the central fixation circle for a brief change in luminance, and to press a button when such an event was detected. After 32 blocks, subjects were given 30 seconds of rest before resuming the task.

The same behavioral task was applied during a retinotopic mapping session. Participants fixated centrally while viewing a wedge, consisting of a flashing black-and-white checkerboard pattern (3 Hz), first rotating clockwise for 9 cycles and then anticlockwise for another 9 cycles (at a rotation speed of 24 s/cycle).

Analysis

Pre-Processing

Data were preprocessed individually for each subject. Preprocessing consisted of realignment, co-registration and low-pass filtering at 128 Hz, using SPM12 (<http://www.fil.ion.ucl.ac.uk/spm>, Wellcome Trust Centre for Neuroimaging, London, UK). A separate pre-processing pipeline was constructed for second-level univariate analyses, which was identical to the above procedure, save for the addition of a normalization stage during which anatomical and functional volumes were aligned to a standard MNI template.

Univariate Analysis

We constructed independent models for each independent scan-session type: experimental, localizer and retinotopic mapping. For experimental sessions, we first examined the relationship between target contrast and BOLD response by modelling a factorial design consisting of target position (left versus right), target orientation (clockwise versus anticlockwise) and target contrast (four levels). In order to cope with the heterogeneous target contrasts between (and in some cases, within) subjects, we discretized contrasts into four bins,

creatively labelled as contrast-level 1 (2% for all subjects) 2, 3 and 4 (12% for all subjects). During certain scanning sessions, some of the mid-tier contrasts had to be adjusted in flight because the subject was either at ceiling or floor performance across two or more contrast levels. In these cases, contrast levels 2 and 3 varied within subjects, so these were binned with their nearest neighbors in order to perform the following univariate analyses. To illustrate, consider subject 5 (see Fig 4.1), where five distinct contrasts were presented. In this particular instance, the 24 trials in which a target was presented at 4.5% contrast were binned with their 96 4%-contrast counterparts. A similar factorial design was modelled for localizer blocks containing only factors for target position and target orientation since contrast was fixed at 100%.

Using our localizer block for its intended purpose, we selected, on a subject-by-subject basis, those voxels that maximally responded to stimulation by our Gabor patches. To do so, we first computed a voxel mask in each subject's native voxel-space (i.e. using non-normalized scan volumes) using target position contrasts from the localizer block. We started from a target position contrast for each subject (left versus right and right versus left), family-wise error corrected to a threshold of $p < .05$, and identified the voxel corresponding to peak activation. From there, we retained the corresponding parent cluster of this peak and combined clusters for both hemispheres in order to construct a third single region-of-interest (ROI) mask. This selected voxels across several visual regions.

Using the combined mask, we were able to isolate corresponding voxels from non-normalized volumes of estimated beta-values generated by our model fit (i.e. beta-maps) and analyse their evolution across various conditions. We began by verifying that our GLM effectively captures the relationship between contrast and overall bold activity by comparing the mean beta value in our ROI for each of the four possible contrast levels independently for left and right-hemisphere ROIs and target position. Applying a repeated-measures ANOVA in order to quantify the relative contributions of target position, contrast and ROI laterality provides some coarse-grained evidence that the BOLD signal in occipital cortex (early visual cortex and

surrounds) varies with with respect to the strength of the visual input, and paves the way for a more thorough analysis of the information contained within.

Multivariate Analysis

For MVPA analyses, fMRI volumes were manipulated using the NIPY v0.4.0 supported by NIBabel v2.0.2 (Brett et al., 2015). The support-vector classifier implementation used for decoding target orientation was that of Scikit-learn v0.17.1 (Pedregosa et al., 2011). The above libraries were executed under Python 3.5.1.

Decoding was performed based on beta values located in the ROI described in the univariate analysis methods, which again were computed in each individual subject's native voxel-space. For each subject, localizer data was split into a “target left” and “target right” group and a linear support-vector classifier (SVC) was trained for each group using per-trial beta maps from the right-hemisphere ROI and the left-hemisphere ROI, respectively. The training phase was cross-validated using an iterative leave-one-out procedure in which two of the 24 trials for each subject (one for each target orientation) were omitted from the training set. These two trials were then decoded and their scores noted. The above process was repeated until each trial had been left out exactly once, whereupon the mean classification score across all iterations of the cross-validation procedure were averaged, yielding a general assessment of decoding accuracy. From there, the two classifiers were trained without excluding any trials from their group, and the resulting kernels were used to decode the low-contrast trials from the main experimental blocks. A similar procedure was also employed using the high-contrast experimental targets ($c=4$, i.e. 12% contrast) as a training set, prior to generalizing to the three remaining levels.

Results

Behavioral

All subjects underwent 96 trials for 2% and 12% contrast whereas the middle-two contrasts for three subjects were adjusted between runs due to ceiling effects (two subjects) or floor effects (one subject) in performance on the 2AFC task. For one subject, the adjustment was

made after two runs, yielding 72 trials at 3% contrast, but for all other adjustments were made after the first run, yielding 84 trials at the optimized contrast.

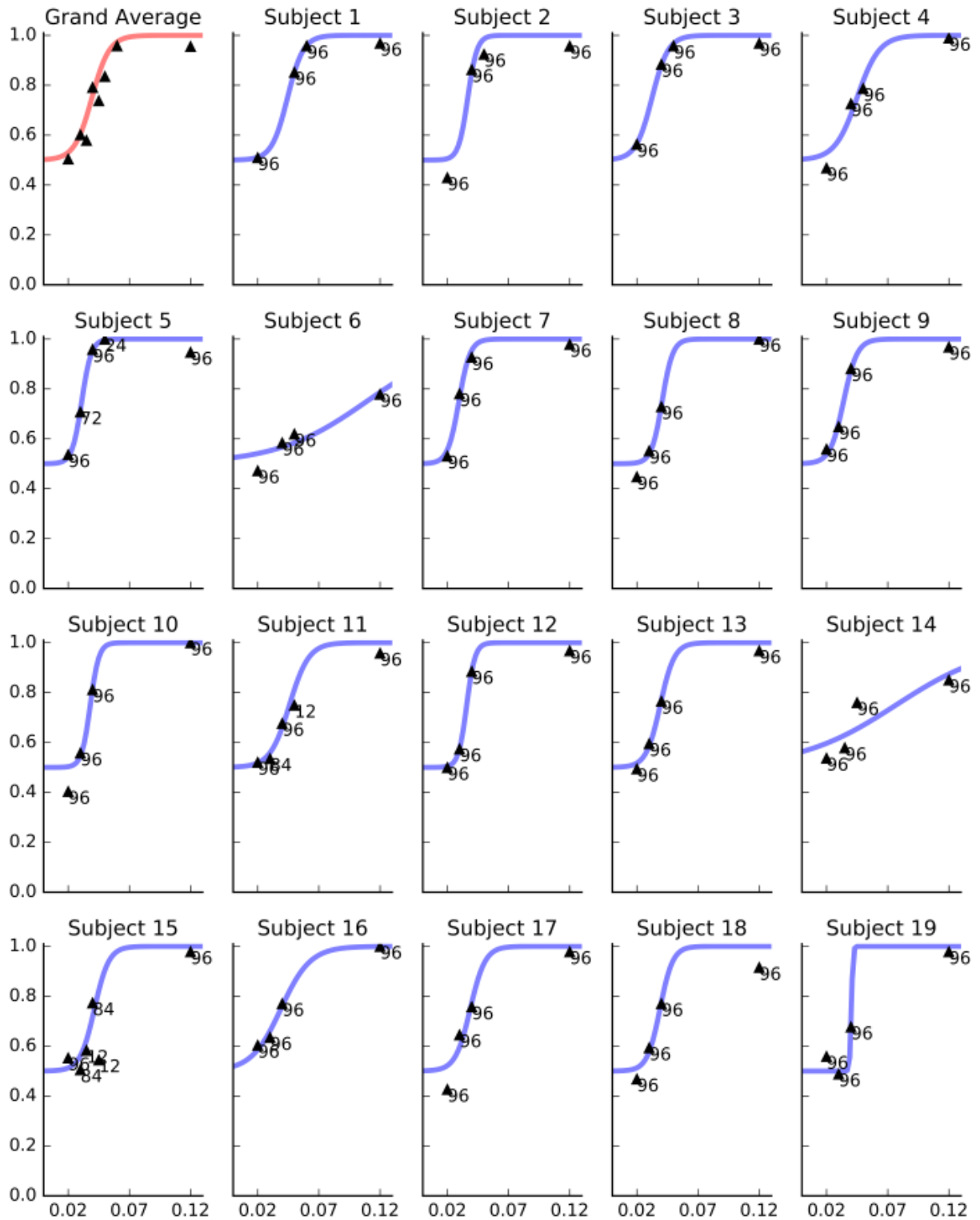


Figure 4.1. Per-subject psychometric curves. Mean hit-rate as a function of target contrast in the main experimental block. For each subject, a psychometric function was fit to the data. For subjects whose contrasts were adjusted in-flight, we pooled trials with neighboring contrasts (see methods).

A linear mixed-effect regression found that the rate of correct responses for the 2AFC orientation-discrimination task increased systematically as a function of contrast ($p < .001$, 95% CI = [.029 .033]), averaging .5 at 2% contrast ($SD=.05$; 19 subjects), .6 at 3% contrast ($SD=.05$; 13 subjects), .58 at 3.5% contrast ($SD=.05$; 2 subjects), .79 at 4% contrast ($SD=.04$; 17 subjects), .65 at 4.5% contrast ($SD=.05$; 2 subjects), .84 at 5% contrast ($SD=.03$, 7 subjects), .96 at 6% contrast ($SD=.02$; 1 subject) and .96 at 12% ($SD=.02$; 19 subjects). Likewise, a main effect was found for mean visibility such that it increased as a function of contrast ($p < .001$, 95% CI = [2.60 2.27]). Mean normalized visibility for each contrast was of .43 at 2% contrast ($SD=.70$), .52 at 3% contrast ($SD=.78$), 2.52 at 3.5% contrast ($SD=2.18$), 1.70 at 4% contrast ($SD=1.32$), 3.44 at 4.5% contrast ($SD=2.43$), 2.82 at 5% contrast ($SD=1.54$), 3.40 at 6% contrast ($SD=1.89$) and 4.91 at 12% contrast ($SD=1.57$).

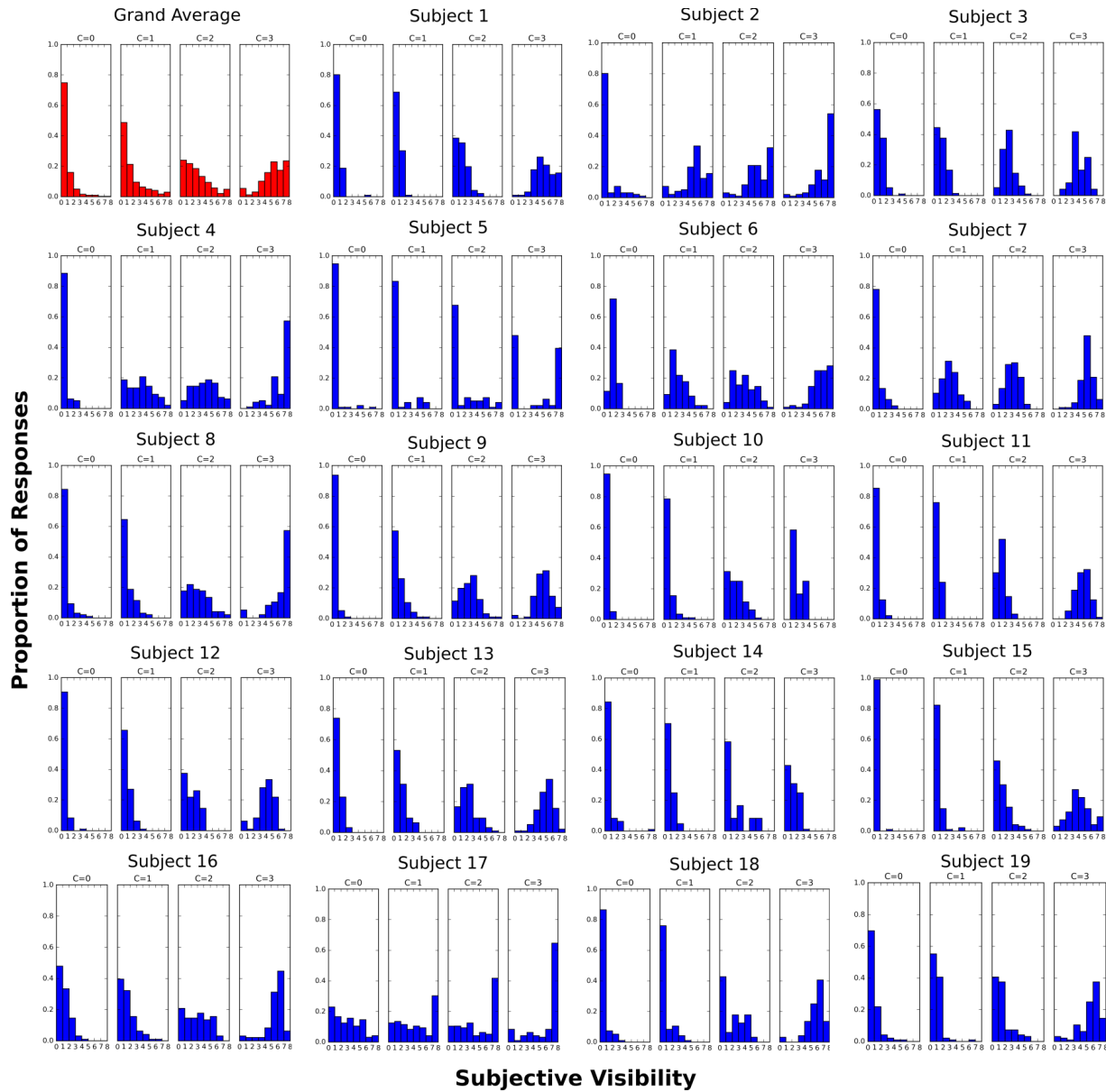


Figure 4.2: Subjective visibility for each contrast level across all subjects.

Imaging

Contrast Analysis

Our first sanity check on the group level confirmed that our visual stimuli, although faint, evoked visual responses in early sensory cortex and that these activations increased as a function of stimulus contrast.

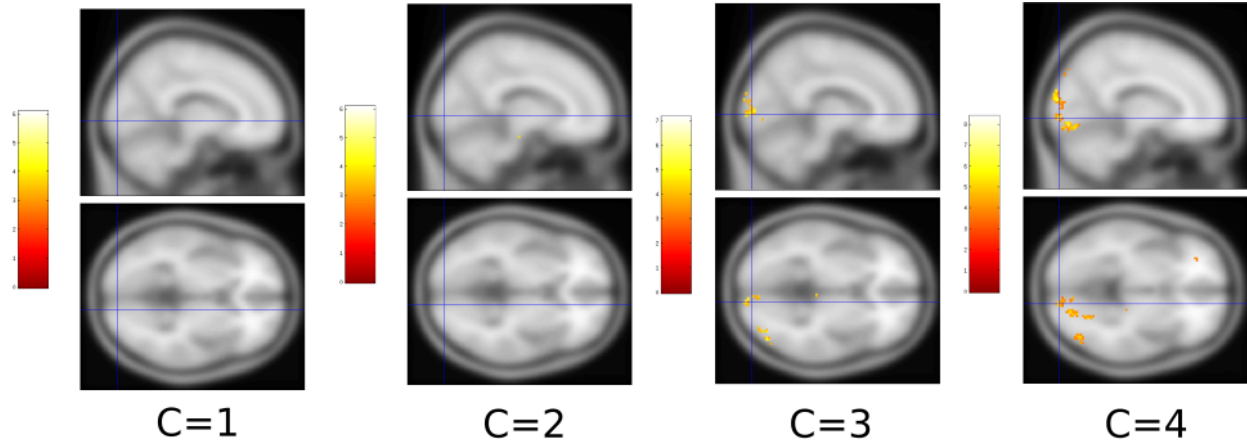


Figure 4.3: Left versus right target position for each of four possible contrast levels, C , across all subjects. Clusters represent uncorrected bold activations where $p < .001$ and BOLD signal is normalized to a standard MNI template.

Satisfied with our qualitative inspection, we proceeded to quantify the degree to which signal strength varies within our region-of-interest, relative to the strength of the stimulus. For each luminance contrast and each target location we examined the activity (beta values) in our left and right ROIs (Fig 4.4). We find main effect in both the left and right hemispheres such that beta values for contralateral targets tended to exhibit larger positive values than their ipsilateral counterparts (see Table 4.1). This suggests a relative increase in localized neural activity that depends on target position, consistent with the seminal fMRI literature that establishes a positive relationship between the contrast of visual stimulation and the magnitude and extent of BOLD response in visual cortex (Goodyear & Menon, 1998). We also report significant interactions between laterality and contrast in each hemisphere, such that an increase in contrast predicts an increase in the magnitude of the difference between the beta values of ipsilateral and contralateral ROIs. Otherwise stated, higher contrasts tend to reduce the beta in ipsilateral visual cortex while increasing it in the contralateral region.

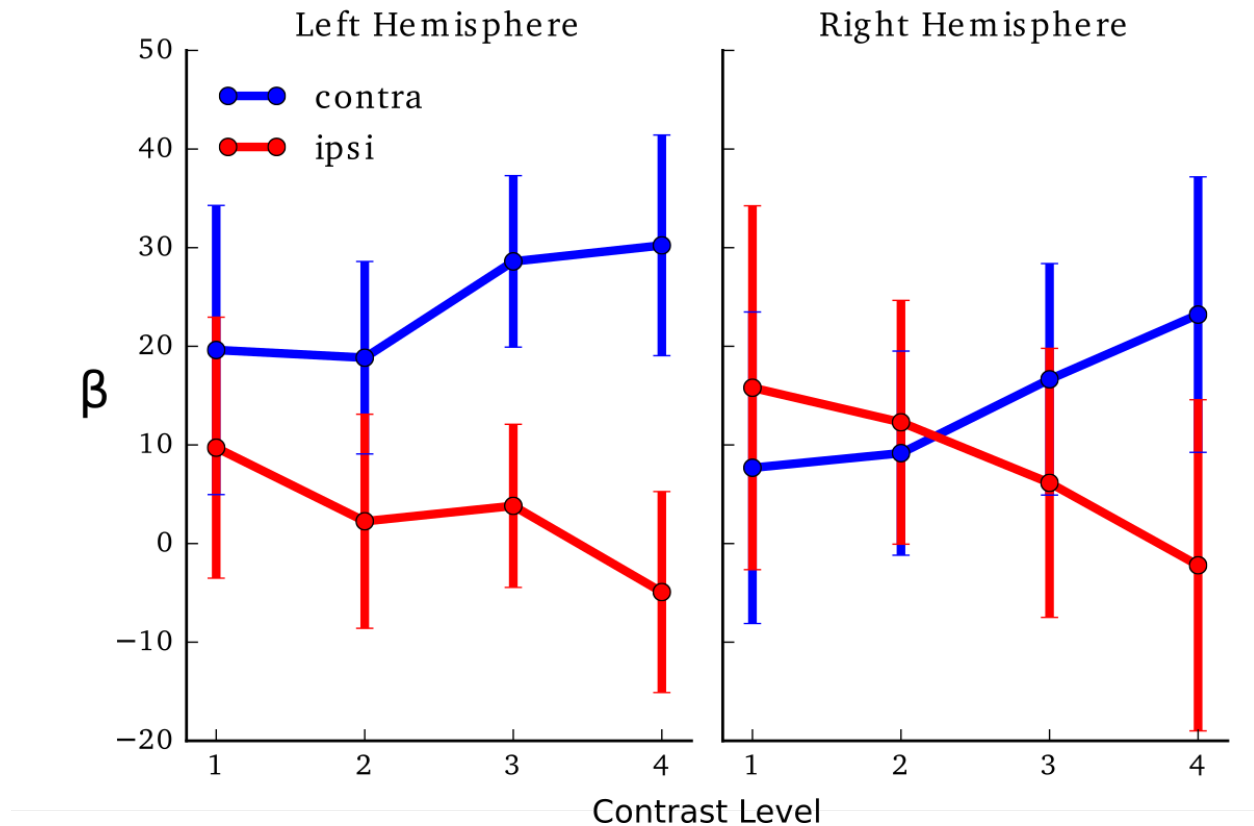


Figure 4.4: Mean beta value in regions-of-interest contralateral and ipsilateral relative to the position of the target. Beta values are in arbitrary units.

Several features of our univariate BOLD analyses are consistent with the results from previous studies that report inhibitory effects of ipsilateral stimuli. In macaques, there is evidence for ipsilateral inhibition of both cortical and retinal origin. Indeed, upon sectioning the corpus callosum and anterior commissure, ipsilateral V4 and V1, respectively, exhibit greatly enhanced responses to visual stimulation, thought to reflect activation of inhibitory neurons. This pattern suggests that the inhibitory surrounds of cortical receptive fields may, to a certain extent, project from one hemisphere to another (Desimone et al., 1993). In humans, studies that examine the lateralization of BOLD response in the visual cortex report that in cortical regions showing significant retinotopic effects (e.g. V1, V2 and V3), the presentation of an identical stimulus in the ipsilateral hemifield has the effect of decreasing the BOLD signal in that same area (Tootell et al. 1998), so it is therefore likely that inhibitory mechanisms similar to those found in monkeys also exist in human visual cortex.

The likely existence of ipsilateral inhibitory responses to visual stimulation suggests the possibility of orientation-specific information in this same ipsilateral cortex. Such ipsilateral signals might be detectable by a machine-learning algorithm, and might aid said classifier in discriminating target orientation on the basis of voxel-wise activation patterns. While Kamitani and Tong report that orientation-classification performance improves proportionally to the number of contralateral voxels considered (Kamitani & Tong, 2005), stimulus orientation has been successfully decoded by analysing voxels which maximally respond to conditions of interest, regardless of their anatomical locations (Haynes & Rees, 2005). As such, these inhibitory effects motivate our decision to decode target orientation based on voxels selected from both the right and the left ROI for each target.

A	ANOVA				
	<i>Left Hemisphere</i>	<i>df</i>	<i>cdfs</i>	<i>F</i>	<i>p</i>
	Lateralization <small>contra / ipsi</small>	1, 17	1, 17	15.11	< .001
	Contrast	3, 51	1.54, 26.11	0.32	< .68
	Lateralization x Contrast	3, 51	2.13, 36.15	23.91	< .001
B	ANOVA				
	<i>Right Hemisphere</i>	<i>df</i>	<i>cdfs</i>	<i>F</i>	<i>p</i>
	Lateralization <small>contra / ipsi</small>	1, 17	1, 17	2.39	< .14
	Contrast	3, 51	1.58, 26.83	0.018	< .96
	Lateralization x Contrast	3, 51	1.57, 26.65	2676	< .001
C	ANOVA				
	<i>Combined</i>	<i>df</i>	<i>cdfs</i>	<i>F</i>	<i>p</i>
	Lateralization <small>contra / ipsi</small>	1, 17	1, 17	47.93	< .001
	Contrast	3, 51	1.36, 23.04	0.14	< .79
	Lateralization x Contrast	3, 51	1.72, 29.24	49.51	< .001
	Hemisphere <small>left / right</small>	1, 17	1, 17	1.02	< .327
	Lateralization x Hemisphere	1, 17	1, 17	3.07	< .098
	Contrast x Position	3, 51	2.55, 43.30	0.34	< .762
	Lateralization x Contrast x Hemisphere	3, 51	1.46, 24.81	1.47	< .246

Table 4.1: ANOVA of betas in the ROIs spanning (A) the left hemisphere, (B) the right hemisphere and (C) the combined left-and-right hemispheres. Bold text denotes statistical significance at $p < .001$.

Classification Analysis

Classifier training — both using localizer data and high-contrast experimental trials — failed to decode above chance at the cross-validation stage. Overall decoding accuracy remained around 50% when decoding from the contralateral ROI, and this performance did not differ significantly from its ipsilateral counterpart.

Discussion

Our failure to adequately train a decoder prevents us from drawing meaningful conclusions about the presence and nature of changes in sensory representations of visual stimuli following their transition from a nonconscious state to a conscious state. What went wrong?

One potential problem is lack of training data. Though the orientation of similar stimuli has been reported using a similarly-sized training set, these stimuli are typically much larger and have a significantly higher contrast ratio (Kamitani & Tong, 2006; Albers et al., 2013; Kok et al., 2012). As such, it is likely that we are ultimately decoding from fewer voxels than our predecessors. We attempted to mitigate this problem by decoding from the combined left and right ROIs, under the hypothesis that there may be some general (that is, non-lateralized) orientation-selective neurons in the ipsilateral hemisphere, and that this could help boost our classification accuracy. Indeed, we report evidence suggesting that some degree of ipsilateral inhibition may be at play as per previous reports (Desimone et al., 1993; Tootell et al., 1998), and these putative signals may show some selectivity for target orientation. Unfortunately the inclusion of the ipsilateral ROI did not significantly improve our decoder's performance, and our classification analysis ultimately remained at chance-level performance.

Support vector classifiers are arguably the most widely used machine-learning algorithm used in fMRI analysis (Sandberg et al., 2014) due to their effectiveness on (relatively) small datasets, conceptual simplicity, low computational costs and suitability for estimating non-linear processes. This having been said, there are other approaches to extracting orientation information from voxel beta-maps which might be more sensitive or otherwise more adapted to

the study at hand. These should be investigated as possible alternatives to the current approach to classification.

Chapter 5: General Discussion

Of Latency and Locality

The present manuscript presents evidence in favor of what we call a “late-and-global” theory of consciousness: an interpretation of the scientific literature according to which awareness emerges from a distinct, secondary phase of perceptual processing that is triggered after the initial buildup of a sensory representation (Baars, 2005; Dehaene et al., 2006). We further argue that attention fulfils this gatekeeping function, with the important corollary that this function can — in principle — be invoked at an arbitrary point in time after the encoding of the stimulus, which in certain cases can forcefully transition a preconscious representation into full-fledged awareness. To wit, we report measurements from two studies that exhibit a discrete, nonlinear transition in the amount of information available to a subject for heterophenomenological report when attention is cued to the location of a previously-displayed item. This constitutes one of the hallmark signatures of conscious access. Numerous electrophysiological and brain imaging studies report that conscious access is accompanied by the sudden apparition of a distributed network of cortical activity that is independent of stimulus strength (Sergent, Baillet & Dehaene, 2005; Marti, Sigman & Dehaene, 2012; Williams, Baker, Op de Beck, Shim, Dang, Triantafyllou & Kanwisher 2008; Roelfsema, Lamme & Spekreijse 1998; Roelfsema, Khayat & Spekreijse 2003; Supèr, Spekreijse & Lamme 2001a; Supèr, Spekreijse & Lamme 2001b; Haynes, Driver & Rees 2005; Williams, Visser, Cunnington & Mattingley 2008; Lamy, Salti & Bar-Haim 2009; Del Cul, Baillet & Dehaene 2007; Donchin & Coles 1988; Bekinshtein, Dehaene, Rothaut, Tadel, Cohen & Naccache 2009; Picton 1992; Melloni, Molina, Pena, Torres, Singer & Rodriguez 2007; Dehaene & Changeux, 2011; Marti, Thibault & Dehaene, 2014). This activity is thought to reflect the encoded stimulus’ accessing of a global routing infrastructure that allows the representation’s informational content to be accessed by various, anatomically-disparate functions, many of which are required for verbal report (Baars 1988; Baars, 1994; Baars, 2005). Such a theory predicts that when a stimulus is

not routed in such a manner, few (if any) bits of information are available for voluntary report. Thus, under properly-controlled circumstances, a subject should respond randomly if forcibly probed for a description of the stimulus' features.

The late-and-global view contrasts sharply with a competing claim that proposes a two-stage process in which consciousness happens early during sensory processing (see Fig 5.1), but in which voluntary report is not possible unless the conscious percept is encoded into memory (Lamme, 2003). Such “early-and-local” theories often point to a general class of effects in which conspicuous changes to a visual scene will escape the observer’s awareness. These effects can be elicited by manipulating wide variety of visual features; an object’s position, color, orientation, or even presence can be altered so long as the experimenter is able to mask the transients elicited by such a change. Typically, this is done by briefly displaying an empty screen between the original and modified versions of the image, sometimes oscillating back and forth between the two (Rensink, 2000; Rensink, 2002; Simons & Levin, 1997; Rizzo et al., 2009). One such experimental paradigm by Becker and colleagues is of particular interest because of its close resemblance to Study 2 (Becker et al., 2000). Given a circular arrangement of letters, each of which might randomly change after a brief blanking of the display, the authors find that pre-cueing one of the letters drastically improves change-detection relative to the absence of any cueing. In addition, the authors report that post-cueing was practically as effective, provided the post-cue appears prior to the onset of the second display of letters (i.e., the one with the changed item). Our colleague Victor Lamme cites this study in support of an early-and-local view of consciousness, arguing not only that the cueing initiates a transfer from iconic into working memory (thus preserving a pre-existing conscious percept), but also that this transition from iconic to working memory can be equated to a qualitative change in the *kind* of consciousness being exhibited (Lamme, 2003). Although we remain cautious in interpreting the results from Study 2, we believe that these offer evidence against this claim. Indeed, a majority of our subjects report upon the stimuli in a manner consistent with flat-out guessing on a selection of trials, and the act of post-cueing seems to grant access to information about a target in an all-or-none fashion (that is, SD does not reliably vary across conditions, suggesting that the same

amount of information is gleaned from post-cueing on average). This is highly suggestive of conscious access, and contrary to the predictions that follow from the early-and-local account of the phenomenon at hand. Our results suggest that improved performance on the aforementioned change-blindness task might, in actuality, stem from a sudden conscious access to the stimulus, which then enables the subject to compare it to the display that follows. We suggest that it may be fruitful to replicate the study performed by Becker and colleagues using the stimuli from Study 2, along with a task amenable to mixture-modelling, with the goal of showing that post-cues decrease the rate at which a subject guesses the orientation difference between a given stimulus across the two displays in each trial. Such results would provide compelling evidence that subjects are not initially aware of the display.

Figure 5.2. Change blindness paradigm with attentional cueing. Change blindness in an abstract scene, and the role of attention. In these change blindness trials (a – c), a scene containing multiple items is presented (Stimulus 1), followed by a gray screen inter-stimulus interval (ISI), after which the same scene (Stimulus 2) is shown again. The subject is then asked whether the cued item (indicated by the orange line) has changed or not. In (a) it has changed orientation. Subjects perform poorly at this task, (60% correct, lower left histogram). Performance can be converted in a ‘capacity’ measure (lower right histogram) indicating how many items the subject had available (in working memory) for change detection, in this case, approximately four items. When the to be changed item is cued in advance (b), subjects perform almost 100% correct (resulting in a virtual capacity of all eight objects). However, when subjects are cued after the disappearance of Stimulus 1 but before the onset of Stimulus 2 (c), they perform almost as well and seem to have stored almost all object (figure and legend reproduced from Lamme, 2003).

It should, of course, be noted that if the VP model were to better describe such a change-blindness experiment, this would not suffice to invalidate the late-and-global interpretation of the consciousness literature. As previously mentioned, we take the relevance of this model for a minority subjects as indicative of differing attentional strategies for completing the experimental task, and this brings up a crucial point in the current debate. In theory, nothing prevents an individual from considering the display as a whole, distributing his attention evenly across the screen and interpreting the scene as a meta-object of sorts. To be sure, such “gist perception” has been extensively documented and considerable evidence has accrued pointing to the involvement of attention in such tasks (Mack & Clarke, 2012; Castelhana & Henderson, 2008; Loschky & Simons, 2004; Vanmarcke & Wagemans, 2014; Cohen, Alvarez & Nakayama, 2011). Our

position clearly calls for a sequel in which we control for such attentional strategies. For instance, it may be possible — through the use of exogenous cues — to initially bias attention towards a subset, S , of targets in the display prior to post-cueing a random target for behavioral report. We would expect that when a member of S is post-cued, the subject would exhibit low guessing rates. Inversely, we would expect high guessing-rates when nonmembers of S are post-cued. And finally, if all targets in the display were pre-cued, it is conceivable that the subject will spontaneously alter his attentional strategy, spreading his attention over the entire display, consistent with a VP model. Note that this “pre-cue all” condition differs from the standard protocol in which there is no pre-cuing. The prediction here is that an absence of a pre-cue results in an arbitrary (and perhaps random) attentional strategy on the part of the subject, whereas employing pre-cues similar to those in Study 1 would exogenously attract attention to all targets at once, in effect “forcing” the subject to deploy an attentional strategy in which he/she encodes some information about all targets.

Additional evidence in favor of the late-and-global interpretation of post-cueing effects can potentially be found in the time-course of retroperception. Observing that a post-cue delays the subjective moment at which a target appears would constitute strong evidence in favor of a latent phase of processing that is distinct from the perceptual encoding phase. In late 2013, we attempted such an experiment using a modified form of the Libet clock paradigm (Libet et al., 1983), but eventually abandoned the endeavor due to the poor temporal resolution of this measure (Buehner, 2009). The principle, however, remains intriguing. During retro-perceived trials, do subjects identify the veridical point in time at which the target was presented, or is the target’s apparition delayed, possibly to the point of being concomitant with the cue? If a suitably-sensitive means of reporting perceived target onset could be found, the relationship between subjective time of apparition and the time-course of various NCCs might provide a means to further dissect the role of these signatures, possibly even ruling out epiphenomenal components.

On the Contents of Consciousness

Returning to the couple from Chapter 1, their argument might be rephrased as follows: how rich is conscious experience, and when in the perceptual pipeline does subjective experience first emerge? By now it should be clear that we intend to answer this question from the perspective of heterophenomenology, and in so doing we must lend credence both to the young woman's claim about the simultaneous, vibrant quality of her conscious experience, and the young man's impression of impoverished perception, which allows him to experience that to which he attends. Crucially, the casual eavesdropper might be tempted to dismiss the first of the two claims on the basis that the young woman is unable to report any information about these wispy percepts she claims to subjectively experience, yet to do so would constitute a foul. Recall that the heterophenomenologist must account for the presence of indescribable-yet-conscious percepts, and more still, he/she must ultimately account for their ineffable quality! The argument, therefore, will not be won or lost on a technicality but on the basis of epistemically-objective evidence.

All-in-all, the current retroperception literature strongly supports the notion that attention is involved in the emergence of conscious perception, to the point that it appears able to causally trigger consciousness. It would appear, as such, that we are for the most part conscious of the objects of our attention, but what of the stimuli that are not attended? Are we conscious of those too? We argue that that we are not on the grounds that subjects *tell us so*, and that these assertions are consistent with behavioral models that take visibility into account on a theoretical level. In the original retroperception study by Sergent and colleagues (2003), valid post-cue had a gross effect of reducing the number of zero-level responses in the subjective visibility scale. Importantly, subjects were instructed to submit an answer of zero if (and only if) they had not seen the target at all, and this result was corroborated by a modelling analysis of the response distribution on the subjective visibility scale (see Fig 5.3).

Figure 5.3. Response Distributions on the Subjective Visibility Scale. (A) Response distributions on the subjective visibility scale (0% is “not seen”; 100% is “maximal visibility”) when the target was present for congruent versus incongruent cues at the three tested stimulus onset asynchronies (SOAs) (in columns). Distributions for congruent cues are transparent blue areas; distributions for incongruent cues are transparent red areas. The area is purple where the two distributions overlap. (B) Response distributions in the same experimental conditions when the target was absent. Congruent and incongruent cueing conditions can still be distinguished in that case, thanks to the use of a response cue. (C–E) Modeling of the response distributions. (C) Schematic representation of the model. For each subject, we modeled the visibility distributions in each experimental condition as a balance between the distributions obtained—for this subject—when the target was absent (bottom row) and when the target was present with a good visibility (top row; precued targets with a correct response on orientation; see Supplemental Experimental Procedures). Shown here is the average of these two template distributions across all participants. The model was tested using a simple regression with one parameter: b . (D) Illustration of the model fit for one experimental condition. In black is the averaged distribution across participants for congruent postcue at SOA = +100 ms (reproduced from A, second column). In green is the average of the modeled distribution across participants for this condition. (E) Parameter estimates. The average b value across participants, which estimates the contribution of the “seen” trials as modeled in (C) top row, was significantly increased for congruent versus incongruent cues, both before and after target presentation (figure reproduced and legend adapted from Sergent et al., 2013).

Study 1 directly extends these findings by providing unambiguous evidence against an alternative account of the retroperception effect according to which objective and subjective metrics are improved by means of increased perceptual precision. Rather, the discrete jump in guessing-rate between valid and invalid cues supports the notion that post-cueing helps subjects see a target *more often* as opposed to seeing a reliably-discernable target *better*.

Although we find this evidence convincing, it is good scientific practice to consider the landscape at a distance. If retroperception indeed reflects the late emergence of consciousness following the intervention of attention, what else might we expect? Otherwise stated, what else might we predict from the hypothesis that attentional-gatekeeping is a general feature of conscious perception?

One obvious prediction is that retroperception should generalize across experimental paradigms and sensory modalities. While the above studies focus on low-level visual features, complex stimuli such as written words should display some sensitivity to retroperception. One such approach is currently being investigated wherein we predict that backwards-masked words should increase in visibility if a congruent semantic cue is presented after the mask. With respect

to generalizing retroperception across sensory modalities, two colleagues from our laboratory, Mark Wexler and Lucile Dupin, recently published an experiment employing a novel method of tactile stimulation, with significant potential for retroperception (Dupin et al., 2014). In short, a square matrix of independently-vibrating piezoelectric pins is used to simulate the running of one's finger over a textured surface. Because the pins can move independently, an experimenter could inject an arbitrary amount of noise, and thus implement a perceptual staircase as in Study 1. Using two such stimulation arrays, it should be possible to implement the haptic equivalent of Study 1 and test for the existence of retroperception in the haptic modality.

The second major prediction is none other than the one that motivated our fMRI study. We initially hypothesized that in times of retroperception, attention acts on the mnemonic trace of a preconscious stimulus to improve the signal-to-noise ratio of the representation and to stabilize it throughout the duration of awareness. This prediction was partially refuted by the results from Study 1, which found that retroperception was governed by a change in conscious access, but not recall precision, thus suggesting that the fidelity of the cortical representation of the target was unchanged. It should be noted, however, that this does not preclude attention from otherwise modifying the encoded percept. Attention might, for instance, ensure that the mnemonic trace persists through time, thereby supplying a machine-learning classifier with additional information stemming from a more accurate beta estimation. While our attempt at MVPA-driven decoding of stimulus orientation is thus far deficient, a successful remedy would pave the way for a direct study of attention influence on the structure of internal perceptual representations. Indeed, should the orientation of the stimuli employed in Study 3 ultimately prove decodable, our next step would be to perform a retroperception experiment in fMRI. Given our hypotheses about the influence of attention on the representation of a stimulus in sensory cortex, we would expect trials with congruent post-cues to exhibit better decoding accuracy relative to those incongruent post-cues or no cue at all. Likewise, we would expect increasing SOA to suppress decoder performance, mirroring the behavioral effects reported in the original retroperception publication.

To conclude, retroperception is a promising new effect in the study of consciousness because it provides a means by which to test for causal involvement of attention in a variety of paradigms. The present manuscript builds upon the original findings, offering confirmatory evidence that attention is capable of triggering awareness, even after the physical source of the stimulus has disappeared. Further, we offer exploratory analyses that invite us to update existing models of foundational effects such as iconic-memory post-cueing and change blindness. We fall short of revealing evidence that attention modifies the informational content of representations in early sensory cortex, but suggest several potential and alternative avenues for exploring this question.

References

- Albers, Anke Marit, Peter Kok, Ivan Toni, H. Chris Dijkerman, and Floris P. de Lange. "Shared Representations for Working Memory and Mental Imagery in Early Visual Cortex." *Current Biology* 23, no. 15 (August 2013): 1427–31. doi:10.1016/j.cub.2013.05.065.
- Albrecht, Thorsten, Susan Klapötke, and Uwe Mattler. "Individual Differences in Metacontrast Masking Are Enhanced by Perceptual Learning." *Consciousness and Cognition* 19, no. 2 (June 2010): 656–66. doi:10.1016/j.concog.2009.12.002.
- Appelle, Stuart. "Perception and Discrimination as a Function of Stimulus Orientation: The 'Oblique Effect' in Man and Animals." *Psychological Bulletin* 78, no. 4 (1972): 266–78. doi:10.1037/h0033117.
- Aru, J., N. Axmacher, A. T. A. Do Lam, J. Fell, C. E. Elger, W. Singer, and L. Melloni. "Local Category-Specific Gamma Band Responses in the Visual Cortex Do Not Reflect Conscious Perception." *Journal of Neuroscience* 32, no. 43 (October 24, 2012): 14909–14. doi:10.1523/JNEUROSCI.2051-12.2012.
- Asplund, Christopher L., Daryl Fougny, Samir Zughni, Justin W. Martin, and René Marois. "The Attentional Blink Reveals the Probabilistic Nature of Discrete Conscious Perception." *Psychological Science* 25, no. 3 (March 2014): 824–31. doi:10.1177/0956797613513810.
- Astle, Duncan Edward, Jennifer Summerfield, Ivan Griffin, and Anna Christina Nobre. "Orienting Attention to Locations in Mental Representations." *Attention, Perception & Psychophysics* 74, no. 1 (January 2012): 146–62. doi:10.3758/s13414-011-0218-3.
- Baars, Bernard J. *A Cognitive Theory of Consciousness*. Cambridge [England] ; New York: Cambridge University Press, 1988.
- Baars, Bernard J. "A Thoroughly Empirical Approach to Consciousness." *Psyche* 1, no. 6 (1994): 1–18.
- Baars, Bernard J. "Global Workspace Theory of Consciousness: Toward a Cognitive Neuroscience of Human Experience." *Progress in Brain Research* 150 (2005): 45–53. doi:10.1016/S0079-6123(05)50004-9.
- Barlow, H. B. "Inductive Inference, Coding, Perception, and Language." *Perception* 3, no. 2 (1974): 123–34.
- Becker, M. W., H. Pashler, and S. M. Anstis. "The Role of Iconic Memory in Change-Detection Tasks." *Perception* 29, no. 3 (2000): 273–86.

- Bekinschtein, T. A., S. Dehaene, B. Rohaut, F. Tadel, L. Cohen, and L. Naccache. "Neural Signature of the Conscious Processing of Auditory Regularities." *Proceedings of the National Academy of Sciences* 106, no. 5 (February 3, 2009): 1672–77. doi:10.1073/pnas.0809667106.
- Block, Ned. "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience." *The Behavioral and Brain Sciences* 30, no. 5–6 (December 2007): 481–99; discussion 499–548. doi:10.1017/S0140525X07002786.
- . "Two Neural Correlates of Consciousness." *Trends in Cognitive Sciences* 9, no. 2 (February 2005): 46–52. doi:10.1016/j.tics.2004.12.006.
- Brainard, D. H. "The Psychophysics Toolbox." *Spatial Vision* 10, no. 4 (1997): 433–36.
- Breitmeyer, Bruno G., and Haluk Ögmen. *Visual Masking: Time Slices through Conscious and Unconscious Vision*. 2nd ed. Oxford Psychology Series, no. 41. Oxford ; New York: Oxford University Press, 2006.
- Broadbent, D. E., and M. H. Broadbent. "From Detection to Identification: Response to Multiple Targets in Rapid Serial Visual Presentation." *Perception & Psychophysics* 42, no. 2 (August 1987): 105–13.
- Castelhano, Monica S., and John M. Henderson. "The Influence of Color on the Perception of Scene Gist." *Journal of Experimental Psychology: Human Perception and Performance* 34, no. 3 (2008): 660–75. doi:10.1037/0096-1523.34.3.660.
- Chalmers, David J. "Facing up to the Problem of Consciousness." *Journal of Consciousness Studies* 2, no. 3 (1995): 200–219.
- Cheal, Mary Lou, Don R. Lyon, and Lawrence R. Gottlob. "A Framework for Understanding the Allocation of Attention in Location-Precued Discrimination." *The Quarterly Journal of Experimental Psychology Section A* 47, no. 3 (August 1994): 699–739. doi:10.1080/14640749408401134.
- Cheal, M., D. R. Lyon, and D. C. Hubbard. "Does Attention Have Different Effects on Line Orientation and Line Arrangement Discrimination?" *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology* 43, no. 4 (November 1991): 825–57.
- Churchland, Paul M. "Reduction, Qualia, and the Direct Introspection of Brain States." *The Journal of Philosophy* 82, no. 1 (January 1985): 8. doi:10.2307/2026509.
- Cohen, M. A., G. A. Alvarez, and K. Nakayama. "Natural-Scene Perception Requires Attention." *Psychological Science* 22, no. 9 (September 1, 2011): 1165–72. doi:10.1177/0956797611419168.

- Cohen, Michael A., Patrick Cavanagh, Marvin M. Chun, and Ken Nakayama. "The Attentional Requirements of Consciousness." *Trends in Cognitive Sciences* 16, no. 8 (August 2012): 411–17. doi:10.1016/j.tics.2012.06.013.
- Cohen, Michael A., and Daniel C. Dennett. "Consciousness Cannot Be Separated from Function." *Trends in Cognitive Sciences* 15, no. 8 (August 2011): 358–64. doi:10.1016/j.tics.2011.06.008.
- Coles, M. G., G. Gratton, and E. Donchin. "Detecting Early Communication: Using Measures of Movement-Related Potentials to Illuminate Human Information Processing." *Biological Psychology* 26, no. 1–3 (June 1988): 69–89.
- Crick, F., and C. Koch. "Consciousness and Neuroscience." *Cerebral Cortex (New York, N.Y.: 1991)* 8, no. 2 (March 1998): 97–107.
- Crick, Francis, and Christof Koch. "A Framework for Consciousness." *Nature Neuroscience* 6, no. 2 (February 2003): 119–26. doi:10.1038/nn0203-119.
- Dehaene, S., L. Naccache, L. Cohen, D. L. Bihan, J. F. Mangin, J. B. Poline, and D. Rivière. "Cerebral Mechanisms of Word Masking and Unconscious Repetition Priming." *Nature Neuroscience* 4, no. 7 (July 2001): 752–58. doi:10.1038/89551.
- Dehaene, S., L. Naccache, G. Le Clec'H, E. Koechlin, M. Mueller, G. Dehaene-Lambertz, P. F. van de Moortele, and D. Le Bihan. "Imaging Unconscious Semantic Priming." *Nature* 395, no. 6702 (October 8, 1998): 597–600. doi:10.1038/26967.
- Dehaene, Stanislas, and Jean-Pierre Changeux. "Experimental and Theoretical Approaches to Conscious Processing." *Neuron* 70, no. 2 (April 2011): 200–227. doi:10.1016/j.neuron.2011.03.018.
- Dehaene, Stanislas, Jean-Pierre Changeux, Lionel Naccache, Jérôme Sackur, and Claire Sergent. "Conscious, Preconscious, and Subliminal Processing: A Testable Taxonomy." *Trends in Cognitive Sciences* 10, no. 5 (May 2006): 204–11. doi:10.1016/j.tics.2006.03.007.
- Dehaene, Stanislas, Claire Sergent, and Jean-Pierre Changeux. "A Neuronal Network Model Linking Subjective Reports and Objective Physiological Data during Conscious Perception." *Proceedings of the National Academy of Sciences of the United States of America* 100, no. 14 (July 8, 2003): 8520–25. doi:10.1073/pnas.1332574100.
- Del Cul, Antoine, Sylvain Baillet, and Stanislas Dehaene. "Brain Dynamics Underlying the Nonlinear Threshold for Access to Consciousness." Edited by Michael Posner. *PLoS Biology* 5, no. 10 (September 25, 2007): e260. doi:10.1371/journal.pbio.0050260.
- Dennett, Daniel C. *Consciousness Explained*. Penguin Science, Philosophy. London: Penguin Books, 1993.
- . "Intentional Systems." *Journal of Philosophy* 68, no. February (1971): 87–106.

- Dennett, Daniel Clement. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. A Touchstone Book. New York: Touchstone, 1996.
- . *The Intentional Stance*. 7. printing. A Bradford Book. Cambridge, Mass.: MIT Press, 1998.
- Dennett, D. C. *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. The Jean Nicod Lectures. Cambridge, Mass: MIT Press, 2005.
- Descartes, René. *Méditations: traduction française*. Translated by Charles Adam and Paul Tannery. Vol. 11. 13 vols. Oeuvres de Descartes 9.1. Paris: Vrin [u.a.], 1982.
- Desimone, Robert, Jeffrey Moran, Stanley J. Schein, and Mortimer Mishkin. "A Role for the Corpus Callosum in Visual Area V4 of the Macaque." *Visual Neuroscience* 10, no. 01 (January 1993): 159. doi:10.1017/S095252380000328X.
- Dupin, Lucile, Vincent Hayward, and Mark Wexler. "Direct Coupling of Haptic Signals between Hands." *Proceedings of the National Academy of Sciences* 112, no. 2 (January 13, 2015): 619–24. doi:10.1073/pnas.1419539112.
- Egeth, Howard E., and Steven Yantis. "VISUAL ATTENTION: Control, Representation, and Time Course." *Annual Review of Psychology* 48, no. 1 (February 1997): 269–97. doi:10.1146/annurev.psych.48.1.269.
- Fahrenfort, J. J., H. S. Scholte, and V. a. F. Lamme. "Masking Disrupts Reentrant Processing in Human Visual Cortex." *Journal of Cognitive Neuroscience* 19, no. 9 (September 2007): 1488–97. doi:10.1162/jocn.2007.19.9.1488.
- Fisch, Lior, Eran Privman, Michal Ramot, Michal Harel, Yuval Nir, Svetlana Kipervasser, Fani Andelman, et al. "Neural 'Ignition': Enhanced Activation Linked to Perceptual Awareness in Human Ventral Stream Visual Cortex." *Neuron* 64, no. 4 (November 2009): 562–74. doi:10.1016/j.neuron.2009.11.001.
- Fougnie, Daryl, Jordan W. Suchow, and George A. Alvarez. "Variability in the Quality of Visual Working Memory." *Nature Communications* 3 (2012): 1229. doi:10.1038/ncomms2237.
- Fries, Pascal. "A Mechanism for Cognitive Dynamics: Neuronal Communication through Neuronal Coherence." *Trends in Cognitive Sciences* 9, no. 10 (October 2005): 474–80. doi:10.1016/j.tics.2005.08.011.
- Gaillard, Raphaël, Stanislas Dehaene, Claude Adam, Stéphane Clémenceau, Dominique Hasboun, Michel Baulac, Laurent Cohen, and Lionel Naccache. "Converging Intracranial Markers of Conscious Access." Edited by Leslie Ungerleider. *PLoS Biology* 7, no. 3 (March 17, 2009): e1000061. doi:10.1371/journal.pbio.1000061.
- Gilbert, C., M. Ito, M. Kapadia, and G. Westheimer. "Interactions between Attention, Context and Learning in Primary Visual Cortex." *Vision Research* 40, no. 10–12 (2000): 1217–26.

- Giordano, Anna Marie, Brian McElree, and Marisa Carrasco. "On the Automaticity and Flexibility of Covert Attention: A Speed-Accuracy Trade-off Analysis." *Journal of Vision* 9, no. 3 (2009): 30.1–10. doi:10.1167/9.3.30.
- Goodyear, B. G., and R. S. Menon. "Effect of Luminance Contrast on BOLD fMRI Response in Human Primary Visual Areas." *Journal of Neurophysiology* 79, no. 4 (April 1998): 2204–7.
- Gregory, R. L. "Knowledge in Perception and Illusion." *Philosophical Transactions of the Royal Society B: Biological Sciences* 352, no. 1358 (August 29, 1997): 1121–27. doi:10.1098/rstb.1997.0095.
- Griffin, Ivan C., and Anna C. Nobre. "Orienting Attention to Locations in Internal Representations." *Journal of Cognitive Neuroscience* 15, no. 8 (November 15, 2003): 1176–94. doi:10.1162/089892903322598139.
- Gross, Joachim, Frank Schmitz, Irmtraud Schnitzler, Klaus Kessler, Kimron Shapiro, Bernhard Hommel, and Alfons Schnitzler. "Modulation of Long-Range Neural Synchrony Reflects Temporal Limitations of Visual Attention in Humans." *Proceedings of the National Academy of Sciences of the United States of America* 101, no. 35 (August 31, 2004): 13050–55. doi:10.1073/pnas.0404944101.
- Haynes, John-Dylan. "Decoding Visual Consciousness from Human Brain Signals." *Trends in Cognitive Sciences* 13, no. 5 (May 2009): 194–202. doi:10.1016/j.tics.2009.02.004.
- Haynes, John-Dylan, Jon Driver, and Geraint Rees. "Visibility Reflects Dynamic Changes of Effective Connectivity between V1 and Fusiform Cortex." *Neuron* 46, no. 5 (June 2005): 811–21. doi:10.1016/j.neuron.2005.05.012.
- Haynes, John-Dylan, and Geraint Rees. "Decoding Mental States from Brain Activity in Humans." *Nature Reviews Neuroscience* 7, no. 7 (July 2006): 523–34. doi:10.1038/nrn1931.
- He, Sheng, Patrick Cavanagh, and James Intriligator. "Attentional Resolution and the Locus of Visual Awareness." *Nature* 383, no. 6598 (September 26, 1996): 334–37. doi:10.1038/383334a0.
- Hill, Anne. "Phantom Limb Pain." *Journal of Pain and Symptom Management* 17, no. 2 (February 1999): 125–42. doi:10.1016/S0885-3924(98)00136-5.
- Humphreys, Gruffydd R., and Marc J. Buehner. "Magnitude Estimation Reveals Temporal Binding at Super-Second Intervals." *Journal of Experimental Psychology: Human Perception and Performance* 35, no. 5 (2009): 1542–49. doi:10.1037/a0014492.
- Iwasaki, S. "Spatial Attention and Two Modes of Visual Consciousness." *Cognition* 49, no. 3 (December 1993): 211–33.
- Jackson, Frank. "Epiphenomenal Qualia." *Philosophical Quarterly* 32 (1982): 127–36.
- . "What Mary Didn't Know." *The Journal of Philosophy* 83, no. 5 (1986): 291–95.

- Jiang, Y., P. Costello, F. Fang, M. Huang, and S. He. "A Gender- and Sexual Orientation-Dependent Spatial Attentional Effect of Invisible Images." *Proceedings of the National Academy of Sciences* 103, no. 45 (November 7, 2006): 17048–52. doi:10.1073/pnas.0605678103.
- Jonides, John. "Voluntary versus Automatic Control over the Mind's Eye's Movement." *Attention and Performance IX* 9 (1981): 187–203.
- Kaernbach, C. "Simple Adaptive Testing with the Weighted up-down Method." *Perception & Psychophysics* 49, no. 3 (March 1991): 227–29.
- Kamitani, Yukiyasu, and Frank Tong. "Decoding Seen and Attended Motion Directions from Activity in the Human Visual Cortex." *Current Biology: CB* 16, no. 11 (June 6, 2006): 1096–1102. doi:10.1016/j.cub.2006.04.003.
- . "Decoding the Visual and Subjective Contents of the Human Brain." *Nature Neuroscience* 8, no. 5 (May 2005): 679–85. doi:10.1038/nn1444.
- Kapadia, M. K., G. Westheimer, and C. D. Gilbert. "Spatial Distribution of Contextual Interactions in Primary Visual Cortex and in Visual Perception." *Journal of Neurophysiology* 84, no. 4 (October 2000): 2048–62.
- Kentridge, R. "Spatial Attention Speeds Discrimination without Awareness in Blindsight." *Neuropsychologia* 42, no. 6 (2004): 831–35. doi:10.1016/j.neuropsychologia.2003.11.001.
- Kim, Chai-Youn, and Randolph Blake. "Psychophysical Magic: Rendering the Visible 'invisible.'" *Trends in Cognitive Sciences* 9, no. 8 (August 2005): 381–88. doi:10.1016/j.tics.2005.06.012.
- Kinchla, R.A. "The Measurement of Attention." *Attention and Performance VIII*, 1980, 213–38.
- Kinchla, R. A., Z. Chen, and D. Evert. "Precue Effects in Visual Search: Data or Resource Limited?" *Perception & Psychophysics* 57, no. 4 (May 1995): 441–50.
- Kjaer, T.W., M. Nowak, K.W. Kjaer, A.R. Lou, and H.C. Lou. "Precuneus–Prefrontal Activity during Awareness of Visual Verbal Stimuli." *Consciousness and Cognition* 10, no. 3 (September 2001): 356–65. doi:10.1006/ccog.2001.0509.
- Knierim, J. J., and D. C. van Essen. "Neuronal Responses to Static Texture Patterns in Area V1 of the Alert Macaque Monkey." *Journal of Neurophysiology* 67, no. 4 (April 1992): 961–80.
- Koch, Christof. *The Quest for Consciousness: A Neurobiological Approach*. Denver, Colo.: Roberts and Co., 2004.
- Koch, Christof, and Naotsugu Tsuchiya. "Attention and Consciousness: Two Distinct Brain Processes." *Trends in Cognitive Sciences* 11, no. 1 (January 2007): 16–22. doi:10.1016/j.tics.2006.10.012.

- Koivisto, Mika, Mikko Lähteenmäki, Thomas Alrik Sørensen, Signe Vangkilde, Morten Overgaard, and Antti Revonsuo. "The Earliest Electrophysiological Correlate of Visual Awareness?" *Brain and Cognition* 66, no. 1 (February 2008): 91–103. doi:10.1016/j.bandc.2007.05.010.
- Kok, Peter, Janneke F.M. Jehee, and Floris P. de Lange. "Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex." *Neuron* 75, no. 2 (July 2012): 265–70. doi:10.1016/j.neuron.2012.04.034.
- Kooistra, C. A., and K. M. Heilman. "Hemispatial Visual Inattention Masquerading as Hemianopia." *Neurology* 39, no. 8 (August 1989): 1125–27.
- Kouider, Sid, Vincent de Gardelle, Jérôme Sackur, and Emmanuel Dupoux. "How Rich Is Consciousness? The Partial Awareness Hypothesis." *Trends in Cognitive Sciences* 14, no. 7 (July 2010): 301–7. doi:10.1016/j.tics.2010.04.006.
- Kouider, Sid, Stanislas Dehaene, Antoinette Jobert, and Denis Le Bihan. "Cerebral Bases of Subliminal and Supraliminal Priming during Reading." *Cerebral Cortex (New York, N.Y.: 1991)* 17, no. 9 (September 2007): 2019–29. doi:10.1093/cercor/bhl110.
- Lamme, V. A., and P. R. Roelfsema. "The Distinct Modes of Vision Offered by Feedforward and Recurrent Processing." *Trends in Neurosciences* 23, no. 11 (November 2000): 571–79.
- Lamme, Victor A. F. "Towards a True Neural Stance on Consciousness." *Trends in Cognitive Sciences* 10, no. 11 (November 2006): 494–501. doi:10.1016/j.tics.2006.09.001.
- . "Why Visual Attention and Awareness Are Different." *Trends in Cognitive Sciences* 7, no. 1 (January 2003): 12–18.
- Lamy, Dominique, Moti Salti, and Yair Bar-Haim. "Neural Correlates of Subjective Awareness and Unconscious Processing: An ERP Study." *Journal of Cognitive Neuroscience* 21, no. 7 (July 2009): 1435–46. doi:10.1162/jocn.2009.21064.
- Leopold, D. A., and N. K. Logothetis. "Activity Changes in Early Visual Cortex Reflect Monkeys' Percepts during Binocular Rivalry." *Nature* 379, no. 6565 (February 8, 1996): 549–53. doi:10.1038/379549a0.
- Leopold, David A., and Nikos K. Logothetis. "Multistable Phenomena: Changing Views in Perception." *Trends in Cognitive Sciences* 3, no. 7 (July 1999): 254–64. doi:10.1016/S1364-6613(99)01332-7.
- Libet, Benjamin, Curtis A. Gleason, Elwood W. Wright, and Dennis K. Pearl. "Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (readiness-Potential): The Unconscious Initiation of a Freely Voluntary Act." *Brain* 106, no. 3 (1983): 623–42. doi:10.1093/brain/106.3.623.

- Linke, A.C., A. Vicente-Grabovetsky, D.J. Mitchell, and R. Cusack. "Encoding Strategy Accounts for Individual Differences in Change Detection Measures of VSTM." *Neuropsychologia* 49, no. 6 (May 2011): 1476–86. doi:10.1016/j.neuropsychologia.2010.11.034.
- Logothetis, N. K., D. A. Leopold, and D. L. Sheinberg. "What Is Rivalling during Binocular Rivalry?" *Nature* 380, no. 6575 (April 18, 1996): 621–24. doi:10.1038/380621a0.
- Loschky, L. C., and D. J. Simons. "The Effects of Spatial Frequency Content and Color on Scene Gist Perception." *Journal of Vision* 4, no. 8 (August 1, 2004): 881–881. doi:10.1167/4.8.881.
- Luck, S. J., and E. K. Vogel. "The Capacity of Visual Working Memory for Features and Conjunctions." *Nature* 390, no. 6657 (November 20, 1997): 279–81. doi:10.1038/36846.
- Luck, Steven J., and Edward K. Vogel. "Visual Working Memory Capacity: From Psychophysics and Neurobiology to Individual Differences." *Trends in Cognitive Sciences* 17, no. 8 (August 2013): 391–400. doi:10.1016/j.tics.2013.06.006.
- Mack, Arien, and Jason Clarke. "Gist Perception Requires Attention." *Visual Cognition* 20, no. 3 (March 2012): 300–327. doi:10.1080/13506285.2012.666578.
- Maloney, J. Christopher. "About Being a Bat." *Australasian Journal of Philosophy* 63, no. 1 (March 1985): 26–49. doi:10.1080/00048408512341671.
- Mangun, G. R., and S. A. Hillyard. "Allocation of Visual Attention to Spatial Locations: Tradeoff Functions for Event-Related Brain Potentials and Detection Performance." *Perception & Psychophysics* 47, no. 6 (June 1990): 532–50.
- Marois, René, Do-Joon Yi, and Marvin M. Chun. "The Neural Fate of Consciously Perceived and Missed Events in the Attentional Blink." *Neuron* 41, no. 3 (February 5, 2004): 465–72.
- Marti, Sébastien, Mariano Sigman, and Stanislas Dehaene. "A Shared Cortical Bottleneck Underlying Attentional Blink and Psychological Refractory Period." *NeuroImage* 59, no. 3 (February 2012): 2883–98. doi:10.1016/j.neuroimage.2011.09.063.
- Marti, Sébastien, Louis Thibault, and Stanislas Dehaene. "How Does the Extraction of Local and Global Auditory Regularities Vary with Context?" Edited by Jyrki Ahveninen. *PLoS ONE* 9, no. 9 (September 8, 2014): e107227. doi:10.1371/journal.pone.0107227.
- Matthew Brett, Michael Hanke, Ben Cipollini, Marc-Alexandre Côté, Chris Markiewicz, Stephan Gerhard, Eric Larson, et al. "Nibabel 2.0.2," 2015. doi:10.5281/zenodo.60846.
- Ma, Wei Ji, Masud Husain, and Paul M Bays. "Changing Concepts of Working Memory." *Nature Neuroscience* 17, no. 3 (February 25, 2014): 347–56. doi:10.1038/nn.3655.
- Ma, Wei Ji, Masud Husain, and Paul M. Bays. "Changing Concepts of Working Memory." *Nature Neuroscience* 17, no. 3 (March 2014): 347–56. doi:10.1038/nn.3655.

- McIntosh, A. R., M. N. Rajah, and N. J. Lobaugh. "Interactions of Prefrontal Cortex in Relation to Awareness in Sensory Learning." *Science (New York, N.Y.)* 284, no. 5419 (May 28, 1999): 1531–33.
- Melloni, L., C. Molina, M. Pena, D. Torres, W. Singer, and E. Rodriguez. "Synchronization of Neural Activity across Cortical Areas Correlates with Conscious Perception." *Journal of Neuroscience* 27, no. 11 (March 14, 2007): 2858–65. doi:10.1523/JNEUROSCI.4623-06.2007.
- Melzack, Ronald. "Phantom Limbs, the Self and the Brain (the D. O. Hebb Memorial Lecture)." *Canadian Psychology/Psychologie Canadienne* 30, no. 1 (1989): 1–16. doi:10.1037/h0079793.
- Merikle, Philip M., and Steve Joordens. "Parallels between Perception without Attention and Perception without Awareness." *Consciousness and Cognition* 6, no. 2–3 (June 1997): 219–36. doi:10.1006/ccog.1997.0310.
- Mitroff, Stephen R., and Brian J. Scholl. "Forming and Updating Object Representations without Awareness: Evidence from Motion-Induced Blindness." *Vision Research* 45, no. 8 (April 2005): 961–67. doi:10.1016/j.visres.2004.09.044.
- Most, S. B., D. J. Simons, B. J. Scholl, R. Jimenez, E. Clifford, and C. F. Chabris. "How Not to Be Seen: The Contribution of Similarity and Selective Ignoring to Sustained Inattentional Blindness." *Psychological Science* 12, no. 1 (January 2001): 9–17.
- Most, Steven B. "What's 'inattentional' about Inattentional Blindness?." *Consciousness and Cognition* 19, no. 4 (December 2010): 1102–4. doi:10.1016/j.concog.2010.01.011.
- Most, Steven B., Brian J. Scholl, Erin R. Clifford, and Daniel J. Simons. "What You See Is What You Set: Sustained Inattentional Blindness and the Capture of Awareness." *Psychological Review* 112, no. 1 (January 2005): 217–42. doi:10.1037/0033-295X.112.1.217.
- Naccache, Lionel, Elise Blandin, and Stanislas Dehaene. "Unconscious Masked Priming Depends on Temporal Attention." *Psychological Science* 13, no. 5 (September 2002): 416–24.
- Nakayama, K., and M. Mackeben. "Sustained and Transient Components of Focal Visual Attention." *Vision Research* 29, no. 11 (1989): 1631–47.
- Newby, E. A., and I. Rock. "Inattentional Blindness as a Function of Proximity to the Focus of Attention." *Perception* 27, no. 9 (1998): 1025–40.
- Norman, Kenneth A., Sean M. Polyn, Greg J. Detre, and James V. Haxby. "Beyond Mind-Reading: Multi-Voxel Pattern Analysis of fMRI Data." *Trends in Cognitive Sciences* 10, no. 9 (September 2006): 424–30. doi:10.1016/j.tics.2006.07.005.

- Oram, M. W., and D. I. Perrett. "Time Course of Neural Responses Discriminating Different Views of the Face and Head." *Journal of Neurophysiology* 68, no. 1 (July 1992): 70–84.
- O'Regan, J. K., and A. Noë. "A Sensorimotor Account of Vision and Visual Consciousness." *The Behavioral and Brain Sciences* 24, no. 5 (October 2001): 939–73; discussion 973–1031.
- Palva, Satu, Klaus Linkenkaer-Hansen, Risto Näätänen, and J. Matias Palva. "Early Neural Correlates of Conscious Somatosensory Perception." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 25, no. 21 (May 25, 2005): 5248–58. doi:10.1523/JNEUROSCI.0141-05.2005.
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, et al. "Scikit-Learn: Machine Learning in Python." *Journal of Machine Learning Research* 12 (2011): 2825–30.
- Pertsov, Yoni, Paul M. Bays, Sabine Joseph, and Masud Husain. "Rapid Forgetting Prevented by Retrospective Attention Cues." *Journal of Experimental Psychology. Human Perception and Performance* 39, no. 5 (October 2013): 1224–31. doi:10.1037/a0030947.
- Pestilli, Franco, and Marisa Carrasco. "Attention Enhances Contrast Sensitivity at Cued and Impairs It at Uncued Locations." *Vision Research* 45, no. 14 (June 2005): 1867–75. doi:10.1016/j.visres.2005.01.019.
- Petersen, S. E., H. van Mier, J. A. Fiez, and M. E. Raichle. "The Effects of Practice on the Functional Anatomy of Task Performance." *Proceedings of the National Academy of Sciences of the United States of America* 95, no. 3 (February 3, 1998): 853–60.
- Picton, T. W. "The P300 Wave of the Human Event-Related Potential." *Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society* 9, no. 4 (October 1992): 456–79.
- Pins, Delphine, and D. H. Ffytche. "The Neural Correlates of Conscious Vision." *Cerebral Cortex* 13, no. 5 (2003): 461–74.
- Pons, T. P., P. E. Garraghty, A. K. Ommaya, J. H. Kaas, E. Taub, and M. Mishkin. "Massive Cortical Reorganization after Sensory Deafferentation in Adult Macaques." *Science (New York, N.Y.)* 252, no. 5014 (June 28, 1991): 1857–60.
- Posner, Michael I., and Marcus E. Raichle. *Images of Mind*. New York: Scientific American Library, 1994.
- Prinzmetal, William, Virginia Long, and James Leonhardt. "Involuntary Attention and Brightness Contrast." *Perception & Psychophysics* 70, no. 7 (October 2008): 1139–50. doi:10.3758/PP.70.7.1139.

- Railo, Henry, and Mika Koivisto. "Reply to Bachmann on ERP Correlates of Visual Awareness." *Consciousness and Cognition* 18, no. 3 (September 2009): 809–10. doi:10.1016/j.concog.2009.05.005.
- Ramachandran, V. S. "Perception of Shape from Shading." *Nature* 331, no. 6152 (January 14, 1988): 163–66. doi:10.1038/331163a0.
- Ramachandran, V. S., and W. Hirstein. "The Perception of Phantom Limbs. The D. O. Hebb Lecture." *Brain: A Journal of Neurology* 121 (Pt 9) (September 1998): 1603–30.
- Raymond, J. E., K. L. Shapiro, and K. M. Arnell. "Temporary Suppression of Visual Processing in an RSVP Task: An Attentional Blink?" *Journal of Experimental Psychology. Human Perception and Performance* 18, no. 3 (August 1992): 849–60.
- Rensink, R. A. "Seeing, Sensing, and Scrutinizing." *Vision Research* 40, no. 10–12 (2000): 1469–87.
- Rensink, Ronald A. "Change Detection." *Annual Review of Psychology* 53 (2002): 245–77. doi:10.1146/annurev.psych.53.100901.135125.
- Rizzo, Matthew, JonDavid Sparks, Sean McEvoy, Sarah Viamonte, Ida Kellison, and Shaun P. Vecera. "Change Blindness, Aging, and Cognition." *Journal of Clinical and Experimental Neuropsychology* 31, no. 2 (February 3, 2009): 245–56. doi:10.1080/13803390802279668.
- Rock, I., and D. Gutman. "The Effect of Inattention on Form Perception." *Journal of Experimental Psychology. Human Perception and Performance* 7, no. 2 (April 1981): 275–85.
- Rodriguez, E., N. George, J. P. Lachaux, J. Martinerie, B. Renault, and F. J. Varela. "Perception's Shadow: Long-Distance Synchronization of Human Brain Activity." *Nature* 397, no. 6718 (February 4, 1999): 430–33. doi:10.1038/17120.
- Roelfsema, P. R., P. S. Khayat, and H. Spekreijse. "Subtask Sequencing in the Primary Visual Cortex." *Proceedings of the National Academy of Sciences* 100, no. 9 (April 29, 2003): 5467–72. doi:10.1073/pnas.0431051100.
- Roelfsema, P. R., V. A. Lamme, and H. Spekreijse. "Object-Based Attention in the Primary Visual Cortex of the Macaque Monkey." *Nature* 395, no. 6700 (September 24, 1998): 376–81. doi:10.1038/26475.
- Ryle, Gilbert. *The Concept of Mind*. New University of Chicago Press ed. Chicago: University of Chicago Press, 2002.
- Sandberg, Kristian, Lau M. Andersen, and Morten Overgaard. "Using Multivariate Decoding to Go beyond Contrastive Analyses in Consciousness Research." *Frontiers in Psychology* 5 (October 30, 2014). doi:10.3389/fpsyg.2014.01250.

- Scholl, B. J., N. S. Noles, V. Pasheva, and R. Sussman. "Talking on a Cellular Telephone Dramatically Increases 'Sustained Inattentional Blindness.'" *Journal of Vision* 3, no. 9 (March 16, 2010): 156–156. doi:10.1167/3.9.156.
- Schwarz, Gideon. "Estimating the Dimension of a Model." *The Analns of Statistics*, no. 6 (1978): 461–64.
- Sergent, Claire, Sylvain Baillet, and Stanislas Dehaene. "Timing of the Brain Events Underlying Access to Consciousness during the Attentional Blink." *Nature Neuroscience* 8, no. 10 (October 2005): 1391–1400. doi:10.1038/nn1549.
- Sergent, Claire, and Stanislas Dehaene. "Is Consciousness a Gradual Phenomenon? Evidence for an All-or-None Bifurcation during the Attentional Blink." *Psychological Science* 15, no. 11 (November 2004): 720–28. doi:10.1111/j.0956-7976.2004.00748.x.
- Sergent, Claire, and Lionel Naccache. "Imaging Neural Signatures of consciousness: 'What,' 'when,' 'where,' and 'how' Does It Work." *Archives Italiennes de Biologie* 150, no. (2/3) (2012): 91–106.
- Sergent, Claire, Valentin Wyart, Mariana Babo-Rebelo, Laurent Cohen, Lionel Naccache, and Catherine Tallon-Baudry. "Cueing Attention after the Stimulus Is Gone Can Retrospectively Trigger Conscious Perception." *Current Biology: CB* 23, no. 2 (January 21, 2013): 150–55. doi:10.1016/j.cub.2012.11.047.
- Simons, Daniel J., and Ronald A. Rensink. "Change Blindness: Past, Present, and Future." *Trends in Cognitive Sciences* 9, no. 1 (January 2005): 16–20. doi:10.1016/j.tics.2004.11.006.
- Simons, D. J., and C. F. Chabris. "Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events." *Perception* 28, no. 9 (1999): 1059–74.
- Simons, D. J., and D. T. Levin. "Change Blindness." *Trends in Cognitive Sciences* 1, no. 7 (October 1997): 261–67. doi:10.1016/S1364-6613(97)01080-2.
- Simons, null. "Attentional Capture and Inattentional Blindness." *Trends in Cognitive Sciences* 4, no. 4 (April 2000): 147–55.
- Sims, Chris R., Robert A. Jacobs, and David C. Knill. "An Ideal Observer Analysis of Visual Working Memory." *Psychological Review* 119, no. 4 (October 2012): 807–30. doi:10.1037/a0029856.
- Sligte, Ilja G., H. Steven Scholte, and Victor A. F. Lamme. "Are There Multiple Visual Short-Term Memory Stores?" *PloS One* 3, no. 2 (2008): e1699. doi:10.1371/journal.pone.0001699.
- Soto, David, and Glyn W. Humphreys. "Automatic Selection of Irrelevant Object Features Through Working Memory: Evidence for Top-Down Attentional Capture." *Experimental Psychology* 56, no. 3 (January 2009): 165–72. doi:10.1027/1618-3169.56.3.165.

- Southall, James P. C. *Helmholtz's treatise on psychological optics*. New York: Dover, 1962.
- Sperling, George. "The Information Available in Brief Visual Presentations." *Psychological Monographs: General and Applied* 74, no. 11 (1960): 1–29. doi:10.1037/h0093759.
- Sperling, G., and M. J. Melchner. "The Attention Operating Characteristic: Examples from Visual Search." *Science (New York, N.Y.)* 202, no. 4365 (October 20, 1978): 315–18.
- Stephan, K. M., M. H. Thaut, G. Wunderlich, W. Schicks, B. Tian, L. Tellmann, T. Schmitz, et al. "Conscious and Subconscious Sensorimotor Synchronization--Prefrontal Cortex and the Influence of Awareness." *NeuroImage* 15, no. 2 (February 2002): 345–52. doi:10.1006/nimg.2001.0929.
- Sterzer, Philipp, Andreas Kleinschmidt, and Geraint Rees. "The Neural Bases of Multistable Perception." *Trends in Cognitive Sciences* 13, no. 7 (July 2009): 310–18. doi:10.1016/j.tics.2009.04.006.
- Subedi, Bishnu, and George T. Grossberg. "Phantom Limb Pain: Mechanisms and Treatment Approaches." *Pain Research and Treatment* 2011 (2011): 1–8. doi:10.1155/2011/864605.
- Supèr, H., H. Spekreijse, and V. A. Lamme. "A Neural Correlate of Working Memory in the Monkey Primary Visual Cortex." *Science* 293, no. 5527 (July 6, 2001): 120–24. doi:10.1126/science.1060496.
- . "Two Distinct Modes of Sensory Processing Observed in Monkey Primary Visual Cortex (V1)." *Nature Neuroscience* 4, no. 3 (March 2001): 304–10. doi:10.1038/85170.
- Tanaka, Y., and D. Sagi. "A Perceptual Memory for Low-Contrast Visual Signals." *Proceedings of the National Academy of Sciences of the United States of America* 95, no. 21 (October 13, 1998): 12729–33.
- Tanney, Julia. "Review: Speaking My Mind: Expression and Self-Knowledge." *Mind* 116, no. 463 (2007): 727–32.
- Theeuwes, Jan. "Perceptual Selectivity for Color and Form." *Perception & Psychophysics* 51, no. 6 (November 1992): 599–606. doi:10.3758/BF03211656.
- Tootell, R. B. H., J. D. Mendola, N. K. Hadjikhani, A. K. Liu, and A. M. Dale. "The Representation of the Ipsilateral Visual Field in Human Cerebral Cortex." *Proceedings of the National Academy of Sciences* 95, no. 3 (February 3, 1998): 818–24. doi:10.1073/pnas.95.3.818.
- Tovée, Martin J. "Neuronal Processing: How Fast Is the Speed of Thought?" *Current Biology* 4, no. 12 (December 1994): 1125–27. doi:10.1016/S0960-9822(00)00253-0.
- Tsuchiya, Naotsugu, and Christof Koch. "Continuous Flash Suppression Reduces Negative Afterimages." *Nature Neuroscience* 8, no. 8 (August 2005): 1096–1101. doi:10.1038/nn1500.

- Van Aalderen-Smeets, Sandra, Robert Oostenveld, and Jens Schwarzbach. "INVESTIGATING NEUROPHYSIOLOGICAL CORRELATES OF METACONTRAST MASKING WITH MAGNETOENCEPHALOGRAPHY." *Advances in Cognitive Psychology*, 2006.
- Van den Berg, Ronald, Edward Awh, and Wei Ji Ma. "Factorial Comparison of Working Memory Models." *Psychological Review* 121, no. 1 (January 2014): 124–49. doi:10.1037/a0035234.
- Van den Berg, Ronald, Hongsup Shin, Wen-Chuang Chou, Ryan George, and Wei Ji Ma. "Variability in Encoding Precision Accounts for Visual Short-Term Memory Limitations." *Proceedings of the National Academy of Sciences of the United States of America* 109, no. 22 (May 29, 2012): 8780–85. doi:10.1073/pnas.1117465109.
- Vanmarcke, Steven, and Johan Wagemans. "Rapid Gist Perception of Meaningful Real-Life Scenes: Exploring Individual and Gender Differences in Multiple Categorization Tasks." *I-Perception* 6, no. 1 (2015): 19–37. doi:10.1068/i0682.
- Varela, F., J. P. Lachaux, E. Rodriguez, and J. Martinerie. "The Brainweb: Phase Synchronization and Large-Scale Integration." *Nature Reviews. Neuroscience* 2, no. 4 (April 2001): 229–39. doi:10.1038/35067550.
- Vogel, Edward K., Geoffrey F. Woodman, and Steven J. Luck. "Pushing around the Locus of Selection: Evidence for the Flexible-Selection Hypothesis." *Journal of Cognitive Neuroscience* 17, no. 12 (December 2005): 1907–22. doi:10.1162/089892905775008599.
- Walker, Robin, John M. Findlay, Andrew W. Young, and John Welch. "Disentangling Neglect and Hemianopia." *Neuropsychologia* 29, no. 10 (January 1991): 1019–27. doi:10.1016/0028-3932(91)90065-G.
- Ward, Robert, Susan Goodrich, and Jon Driver. "Grouping Reduces Visual Extinction: Neuropsychological Evidence for Weight-Linkage in Visual Selection." *Visual Cognition* 1, no. 1 (January 1994): 101–29. doi:10.1080/13506289408402295.
- Watson, John B. "Psychology as the Behaviorist Views It." *Psychological Review* 20, no. 2 (1913): 158–77. doi:10.1037/h0074428.
- Wiesmann, U. N., S. DiDonato, and N. N. Herschkowitz. "Effect of Chloroquine on Cultured Fibroblasts: Release of Lysosomal Hydrolases and Inhibition of Their Uptake." *Biochemical and Biophysical Research Communications* 66, no. 4 (October 27, 1975): 1338–43.
- Williams, Mark A., Chris I. Baker, Hans P. Op de Beeck, Won Mok Shim, Sabin Dang, Christina Triantafyllou, and Nancy Kanwisher. "Feedback of Visual Object Information to Foveal Retinotopic Cortex." *Nature Neuroscience* 11, no. 12 (December 2008): 1439–45. doi:10.1038/nn.2218.

- Williams, M. A., T. A. W. Visser, R. Cunnington, and J. B. Mattingley. "Attenuation of Neural Responses in Primary Visual Cortex during the Attentional Blink." *Journal of Neuroscience* 28, no. 39 (September 24, 2008): 9890–94. doi:10.1523/JNEUROSCI.3057-08.2008.
- Womelsdorf, Thilo, Jan-Mathijs Schoffelen, Robert Oostenveld, Wolf Singer, Robert Desimone, Andreas K. Engel, and Pascal Fries. "Modulation of Neuronal Interactions through Neuronal Synchronization." *Science (New York, N.Y.)* 316, no. 5831 (June 15, 2007): 1609–12. doi:10.1126/science.1139597.
- Woodman, Geoffrey F., and Steven J. Luck. "Dissociations among Attention, Perception, and Awareness during Object-Substitution Masking." *Psychological Science* 14, no. 6 (November 2003): 605–11. doi:10.1046/j.0956-7976.2003.psci_1472.x.
- Yeshurun, Yaffa, Barbara Montagna, and Marisa Carrasco. "On the Flexibility of Sustained Attention and Its Effects on a Texture Segmentation Task." *Vision Research* 48, no. 1 (January 2008): 80–95. doi:10.1016/j.visres.2007.10.015.
- Zhang, Weiwei, and Steven J. Luck. "Discrete Fixed-Resolution Representations in Visual Working Memory." *Nature* 453, no. 7192 (May 8, 2008): 233–35. doi:10.1038/nature06860.
- Zokaei, Nahid, Maike Heider, and Masud Husain. "Attention Is Required for Maintenance of Feature Binding in Visual Working Memory." *The Quarterly Journal of Experimental Psychology* 67, no. 6 (June 3, 2014): 1191–1213. doi:10.1080/17470218.2013.852232.