ANTHROPOLOGY & PHILOSOPHY

International Multidisciplinary Journal

Questo fascicolo è stato pubblicato con un contributo parziale
del Dipartimento di Studi Storico-Sociali e Filosofici
dell'Università di Siena

A&P

ANTHROPOLOGY & PHILOSOPHY
International Multidisciplinary Journal

Editorial Address:
Università di Siena - Dipartimento di Studi Storico-Sociali e Filosofici – Viale L. Cittadini, 33 - 52100 Arezzo (Italy) – Ph. +39-0575-9261 - Fax +39-0575-926312 - e-mail: bianca@unisi.it

A&P is published as one volume divided in 2 issues

Subscription Rates:
Institutional Subscription: 35 euro
Personal Subscription: 25 euro
To subscribe please write to:
A & P – Università di Siena - Dipartimento di Studi Storico – Sociali e Filosofici, Viale L. Cittadini, 33 – 52100 Arezzo (Italy)

ANTHROPOLOGY & PHILOSOPHY
International Multidisciplinary Journal
Volume 7 (1-2) 2006

Special Issue on

Mental Imagery and Visual Perception

Edited by Mariano L. Bianca and Lucia Foglia

CONTENTS

Foreword .... p. 7

Talis Bachmann — *Imagination as virtual reality and how it can be explained by the perceptual retouch theory of non-specific modulation* .... p. 10

Mariano L. Bianca, Lucia Foglia — *Non-Perceptive Mental Image Generation: a Non-Linear Dynamic Framework* .... p. 28

Anna M. Borghi, Claudia Scorolli — *Object Concepts and Mental Images* .... p. 64

Francesco Ferretti — *Imagery, Perception and Creativity*................. p. 75

James Genone — *Concepts and Imagery in Episodic Memory* .... p. 95

Janet Levin — *Can Mental Images Provide Evidence for What is Possible?* .... p. 108

Marco Mazzone — *Imagery, Language and the Flexibility of Thought* .... p. 120

Michael J. Proulx, Petra Stoerig — *Seeing sounds and tingling tongues: Qualia in synaesthesia and sensory substitution* .... p. 135

Frédérique de Vignemont — *Brainreading of perceptual experiences: a challenge for first-person authority?* .... p. 151

Foreword

In recent years the issue of mental imagery has been addressed from different point of view, for instance, in neuroscience, neurophysiology, neuropsychology and philosophy of mind.

As we know, several hypotheses have been put forward, sometimes very different from one another, whose main aim is, on the one hand, to clarify the nature of imagery and its relationship with visual perception; on the other hand, to discuss the questionable topic of the involvement of conceptual knowledge, or other mental contents, in imagery generation. Up to now, neurophysiological findings have given some relevant hints, but certainly not exhaustive, and so there is the need to investigate further in order to fully explain mental imagery generation and its contents.

This Special Issue of Anthropology & Philosophy, devoted to Mental Imagery and Visual Perception, attempts to offer some more contributions to this current investigation and debate from different perspectives which use specific methods and tools. These approaches, indeed, show that the issue of imagery can be addressed in different ways and so they highlight different aspects of it.

In *Imagination as virtual reality and how it can be explained by the perceptual retouch theory of non-specific modulation* Talis Bachmann deals with mental imagery from its microgenetic approach through the retouch theory. According to Bachmann there are two processes in percept genesis which give rise respectively to a specific representation of perceptual contents and to a non-specific modulation from thalamus, serving the function of revelation of the contents represented by the specific system for subjects consciousness.

In *Non-Perceptive Mental Image Generation: a Non-Linear Dynamic Framework* Mariano L. Bianca and Lucia Foglia show the difference between what they call Mental Imagery Standard Framework, which includes the usual assumptions on imagery, and what they call Mental Imagery Non-linear Framework, by which they analyse perceptive images and the main features of non-perceptive images, their generation, the relationship between them and their semantic reference to the physical objects. They hold that the main result of this approach is that imagery is a configurational setting emerging from different retrieval processes which can modify visual information.

In *Object Contents and Mental Images* Anna M. Borghi and Claudia Scorolli focus on mental imagery and concepts, presenting a perspective, the so-called ‘embodied’ view, which highlights the relevance of perceptual and motor imagery for concepts, and suggest that thinking of an object or comprehending a word implies referring to the perceptual aspects of its referent.

In *Imagery, Perception and Creativity* Francesco Ferretti justifies the role of mental imagery in creativity; in particular, the central idea of his paper is that the justification for the role of mental images in the creative process lies in the analysis of the relationship between vision and imagery. Imagery, indeed, is a process analogous to the perception when an object is absent, as imagery and vision share common characteristics; according to Ferretti the relationship between imagery and perception allows imagery to be a good substitute for reality.

In *Concepts and Imagery in Episodic Memory* James Genone addresses the general topic of the role played by concepts in mental imagery generation and visual perception; to deal with this issue, on the one hand, he maintains that there is a phenomenal difference between mental imagery and perceptual experience, as they involve different concepts; on the other hand, he claims that there must be a connection between the concepts applied to an experience that is recalled in memory and the concepts involved in the original experience.

In *Can Mental Images Provide Evidence for What is Possible?* Janet Levin approaches a very difficult and tricky problem, i.e., the way in which we can provide evidence for knowledge or for what is possible or impossible metaphysically speaking. She argues that mental images provide good evidence for modal claims but she also underlines that, in some cases, we are not able to formulate an image to support what is possible or impossible, for instance, a 'round square'. In this last example, indeed, we can support the possibility of thinking of a 'round square' by employing to non sensory representations, or rather, conceptual resources; thus, according to Levin there are two ways to provide evidence for what is metaphysically possible or impossible.

In *Imagery, Language and Flexibility* Marco Mazzone deals with sensory imagination conceived as a form of reasoning, which allows people to connect means with ends; in doing so, he analyses the notions of accessibility and systematicity both in propositional representation and perceptual representation and, presenting different perspectives, he holds that people can make use of sensory imagination as cognitive strategy.

In *Seeing sounds and tingling tongues: Qualia in synaesthesia and sensory substitution* Michael J. Proulx and Petra Stoerig bring together two seemingly independent areas of research: synaesthesia and sensory substitution, i.e., what substitutes a sensory modality that a person has lost by transforming the information it provided so that it can be accessed through another, intact sensory modality. In this paper, they review the literature on synaesthesia, and discuss the issue of qualia for this domain of research, to ask finally whether sensory substitution may induce a 'synthetic' form of synaesthesia by taking advantage of the nervous system's capacity for generating visual images in the absence of retinal input.

In *Brainreading of perceptual experiences: a challenge for first-person authority?* Frédérique de Vignemont deals with the problem of the privacy of the mind, or rather, the privacy of conscious experience; as progress in brain imaging has given brain scientists the possibility to read someone's mind, the main question is to understand how they can actually challenge the authority of the mind. Could they really know better than us our own perceptual experiences? As we can see, what is at stake here is the possibility that somebody else can have access to our conscious experiences.

We believe that all these papers can give a valid contribution to shedding some further light on the matter.

We wish to thank, first of all, the authors of the papers, who accepted our proposal to take part in this project and, thereby, allow the realization of the Special Issue on Mental Imagery and Visual Perception, and, naturally, the Editorial board of A&P who have collaborated in its final drafting and review.

Mariano L. Bianca & Lucia Foglia

Talis Bachmann
University of Tartu

Imagination as virtual reality and how it can be explained by the perceptual retouch theory of non-specific modulation

Abstract Perceptual retouch metatheory of microgenetic conscious perception is presented and adjusted so as to explain mental imagery. In the retouch theory, two principal participating processes involved in percept genesis are characterised and their interaction described. The first one of the processes is specific representation of the perceptual contents, updated continuously with continually impinging sensoriae. The second one of the processes is a general-purpose, unspecific modulation from thalamus serving the function of revelation of the contents represented by the specific system for subject's consciousness. When instead of the actual sensory data the contents of cognition are episodic memory based ('eyes-closed vision'), the non-specific thalamic modulation is applied onto perceptual cortical codes that exemplify these contents. Thereby, virtual-reality mode of perceptual representation can be created for subjective apprehension. However, the control of the modal contents of imagery as the freely permuted individual units of the retouched episodic-memory recombinations is endogenous and symbolic.

Key words Perception, imagery, consciousness, perceptual retouch, thalamic non-specific modulation, virtual reality.

1. Introduction

When I play Giuseppe Tartini's Sarabande on my violin, I am conscious of the magic of this fine piece of music and at the same time aware of the fact that physical sounds produced by the movement of my Tartini's bow over the strings are the link in a causal change of events leading to a rewarding subjective experience. But when I then look at the notes in Tartini's Sonata in G Minor ('The Devil'sTrill' that I can't play myself for obvious reasons and surely not in the right tempo), my experience of music is virtual, not caused directly by physical sounds but only implied by the notation symbols on a page of paper. Nobody, however, would deny that there is an experience present and its contents are largely musical (perhaps mixed with a hint of sinful undertones and devine overtones of a conscientious nature). So some experiences are those of actual perceptual content and some other are imaginary, with music being present in a virtual mode. (If indeed the dramatic legend about the Tartini's dream from where he got the music has some authenticity in it, we could of course extend the discourse to the themes of an intriguing possibility that what becomes actual may have had to be first virtual, but this is another topic not pursued in the present article. One way or another, my hope is that the devil-in-dream theme appears in my writing not as a hidden *primo causa*, but as an insignificant coincident.)

Nowadays, it goes without saying that brains mediate subjective experiences and we even have an accumulating knowledge about the necessary mechanisms involved in generating the capacity for perception and imagination. (Let us leave the problem of mechanismic sufficiency aside for a while – we simply do not know enough.) A productive strategy of scientific research of psychological processes presupposes juxtaposition of some various perceptual phenomena that include some controversy or paradox in them on the one hand and the workings of brain mechanisms carried out at the time of experiencing these phenomena on the other hand. The unexpected may mean that there is something we do not know as yet and a careful analysis of how the unusual could emerge can clarify also the understanding of what is routinely behind the usual. In the present work I will describe my understanding of the mechanisms that constitute the neurobiological basis of perceptual subjectivity, present the core of my perceptual retouch theory and show how it is easy to use this theory for describing not only several phenonmena of actual perception but also mental imagery as a virtual-reality form of subjectivity.

My topic is related to the current scientific debate on the role of binding operations that are carried out by the brain processes of synchronisation between the activities of different specialised pools of neurons. Thereby different, but intimately interrelated domains of binding (as cognition) become dealt with: pre-conscious representational binding of object/event features into pre-conscious perceptual objects; binding of these pre-conscious objects-to-be into conscious apprehension (executed by perceptual retouch operations); binding of activated semi-autonomous episodic-memory units into conscious imagery (executed, again, by perceptual retouch operations).

2. Sensory-based cognition and perceptual retouch

2.1 Understanding of perception as microgenesis

Mental apprehension of an object or event, after a subject has encountered it, occurs after the latent time interval and does not appear in its final, fully formed character as a sudden "all-from-none" percept with infinitesimally small time of formation. Mental states of perception are themselves a subject to development which is characterized by definite successive stages (Lange, 1892; Sander, 1928; Werner, 1940, 1956; Undeutsch, 1942; Flavell & Draguns, 1957; Smith, 1957; Hanlon, 1991; Bachmann, 2000; Riffert, 2004; Rosenthal, 2004). I distinguish between two aspects of perceptual formation as microgenesis. First, *representational microgenesis* (RMG) that stands for formation of an active mental representation of an object, scene or event. This formation is aided by specialized neurobiological processes and constitutes the succession of the necessary stages in mental activity that will guarantee that the information about a particular object becomes a psychological reality that allows a person implicitly or explicitly understand the properties and the meanings of it, to act upon it, and store this information in memory. We can analyse RMG irrespectively of its relation to whether what is represented is apprehended in consciousness or not. A lot of evidence shows that specific information about the physical properties, meanings, emotiogenic qualities, numerical characteristics and action-guiding choices is processed, represented and stored by brain preconsciously (Dixon, 1981; Marcel, 1983; de Gelder et al., 2001; Dehaene & Naccache, 2001; Naccache & Dehaene, 2001; Jaśkowski et al., 2002; Moutoussis & Zeki, 2002; VanRullen & Koch, 2003; Kinoshita & Lupker, 2003). A technical analogue for RMG would be the photographical process where negative image of a scene stands for object and the physical-chemical processes taking part within the emulsion layer of the photographic paper after it has been stimulated by the light pattern from the negative image stand for RMG.

The second aspect of perceptual formation relates to *phenomenal microgenesis* (PMG). This is unfolding or formation of an active and dynamic mental representation of object or event in the directly experienced, phenomenally explicit format (Bachmann, 2000). A technical analogue for PMG would be the photographical process where the photochemically structured information from the negative image within the emulsion layer of the photographic paper becomes developed into the pattern of reflectance gradients on the surface of the paper after it has been immersed into the developer liquid. In photography the observable picture development takes dozens of seconds, but in mental (perceptual) microgenesis the process unfolds within about 0.1-0.3 seconds (Breitmeyer, 1984; Bachmann, 2000).

It is realistic to assume that in microgenesis a representation at the proto-object stage is being formed, which will be used for further updating and individuation at later stages of perception (compare also with "tokens" and "types" – Kanwisher, 1987). Updating of the evolving microgenetic object is faster than creating a new representation for the newly appearing object *ab ovo*. Proto-objects at the early stages of microgenesis are sufficient in order to provide the "gist" of stimulation for the higher cognitive levels before the full-blown PMG has been completed. In microgenesis, gist and coarse characteristics precede the detailed and individuated perceptual object and this regularity has been supported both by psychophysical and neurobiological research data (Bachmann, 2000; Crick & Koch, 2003; Rensink, 2000; Hochstein & Ahissar, 2002; Lamme, 2003). In addition to the stagewise representational activity, one more process is needed: the process by which an explicit, reportable, phenomenal reference could be established, as related to the higher stage of processing "entering" consciousness. Here, we speak about modeling the world also in a phenomenally direct, conscious-experience format.

2.2 Perceptual retouch theory of microgenesis

Research in neurobiology and psychophyhsiology has shown that the systems for content processing are insufficient for explicit perception. There is a distinction between two brain systems: the sensory systems for stimulus-specific *content* and the systems for providing sufficient level and a proper pattern of cortical neuronal activation that is necessary for permitting a particular content to become explicitly represented (become conscious). The latter factor is required for awareness, but does not directly contain specific contents of conscious experiences (Baars, 1995, 1997; Bachmann, 1984; Bogen, 1995; Llinás & Ribary, 2001; Rees, Kreiman, & Koch, 2002). The neurons of the content-specific system are termed "drivers" and the neurons of the conscious state systems belong to the class of "modulators" (Crick & Koch, 1998; Sherman & Guillery, 1998). Drivers that have small receptive fields and that respond to spatially localised stimuli with very short delays are modulated by the facilitatory input from the content-free "modulators" of the so-called non-specific thalamus (Bachmann, 1984, 1994; Crick, 1984; Magoun, 1958; Purpura, 1970; Steriade, 1996; Steriade, Jones, & Llinás, 1990; Steriade, Jones, & McCormick, 1997).

Drivers encode specific stimulus features such as size, orientation, color, motion, etc. The thalamic structures termed "non-specific" (e.g., intralaminar nuclei, pulvinar, nucleus reticularis, etc.) do not participate directly in the operations of encoding of the contents of specific sensory information. Yet, their efferent pathways are projected presynaptically onto specific cortical driver-neurons. Non-specific units modulate the level of activity of the drivers, no matter what were the specific signals that evoked the activity within the specific system. It is the process of becoming conscious of what is represented by the specific neurons.

The processing of novel sensory signals from environmental objects is thus serviced by two processes: (1) fast stimulus-specific responses by drivers and (2) a slower, spatially dispersed modulation via the collaterals that pass through the non-specific thalamus. Since the latency of the cortical response to non-specific modulation is considerably slower as compared to the afferent latency of the specific cortical neurons, the driver-neurons, initially activated only by the specific afference, have to wait for the arrival of the stimulus-related modulatory input. This secondary input has been shown to be necessary for explicit perception (awareness) of the stimulus information pertaining initially to preconscious specific representations (Baars, 1997; Bachmann, 1984, 1994; LaBerge, 1997; Llinás & Ribary, 2001; Newman, 1995; Steriade et al., 1997). Because receptive fields of the nonspecific units are coarse, a spatially remote stimulus can affect facilitatory modulation of some other stimulus' specific processing; these specific units while having different receptive fields in terms of the specific driver system share the receptive field of the same nonspecific units in thalamus. Non-specific thalamus receives substantial reentrant input from higher cortical areas.

In terms of the retouch theory (Bachmann, 1984, 1994), the temporally delayed and spatially diffuse modulation from the modulators, targeted at the specific cortical neurons that carry preconscious information about the specific stimulation content serves to "retouch" that content for visual awareness (for explicit representation). Initially, at the onset of stimulation where there has not been any locally preceding visual input, the modulation through non-specific thalamus (or the change in the dynamic characteristics of this modulation such as phase resetting of the oscillatory activity of neurons) takes considerable time to become effective. Consequently, the latency of sampling of the newly presented specific stimulus-signals for explicit representation is initially slow. The slow nonspecific modulation that is necessary for explicit representation and that was evoked by the preceding stimulus-input arrives at cortical specific level of feature representation with a delay during which the subsequent stimulus-input has entered. If the subsequent input is invariant with the previous one (e.g., the same object continued to be present), nothing surprising happens and the same content of sensoriae is continually upgraded for explicit perception. If the subsequent input is much different, such as in case of backward masking or metacontrast (Breitmeyer, 1984; Bachmann, 1994), it will replace the previous one in explicit perception. The extent of the temporal travelling window of dynamic interaction where the modulation initiated by the input at the time **t** modulates the specific activity of the input presented at the time **t+1** is comparable to about 50-150 milliseconds (ms). This is the time difference between the typical delays with which non-specific modulation- and specific sensory driving impulses reach cortex at the sites where the content-representing neurons reside.

Several surprising phenomena can be explained by the retouch theory in a not so surprising way if one takes a closer look at the neurobiology of processing. The above described regularity explains why backward masking of the signals presented at **t**(0) by the signals presented at **t**+1 dominates over forward masking (Bachmann, 1984, 1994, 1997, 1999). At the time when modulation arrives cortex, the newly acquired specific signals of the *mask information* (**t**+1 "fresh" signals that have the highest possible signal-to-noise ratio just at that moment) arrive or are continuously shipped through the fast specific channels and *take advantage of the modulation* that was preset by the preceding stimulation. The mask dominates explicit representation because specific signals of the preceding target are more decayed at this critical epoch of time and specific signals of the mask are less decayed or even are continuously arriving to the specific feature processing sites in cortex as is the case with simultaneous onset, asynchronous offset of target and mask (e.g., Di Lollo, Enns, & Rensink 2000; Enns, 2004; Bachmann, 2005).

It is well known that the different types of visual features (colour, orientation, direction of motion, size, etc) are represented by the specialised neurons in the sensory areas of the back of cortex and are mutually separated by a long distance (Churchland, 2002; Cleeremans, 2003; Crick & Koch, 2003). Yet, conscious experience of the objects that include different features (e.g., red and green colour of a skirt with tilted stripes) involves all those different features as integrated and belonging to the same coherent object. Therefore, some process of feature binding should participate in how the integrated objects become represented as coherent entities (e.g., Treisman, 1998; Cleeremans, 2003). Pre-conscious effects on motor choice responses, priming interpretations and free associations can be based on integrated properties of elementary local sensory input (such as emotional connotation of faces, number information, meanings of words). Therefore, the binding of features into conjoint objects has to be performed pre-consciously. (For the evidence on pre-conscious perceptual effects see Dixon, 1981; Marcel, 1983; de Gelder et al., 2001; Dehaene & Naccache, 2001; Naccache & Dehaene, 2001; Jaśkowski et al., 2002; Moutoussis & Zeki, 2002; VanRullen & Koch, 2003; Kinoshita & Lupker, 2003.)

At this point, we have reached the acceptance of few basic principles. First, pre-conscious microgenetic build-up of perceptual object representations in RMG should necessarily involve feature-binding operations. Secondly, these binding-processes are carried out by the drivers' activity in the specific cortical system and are executed faster than the slow modulation from thalamus arrives. It therefore follows that explicit perception in consciousness that owes to the slow thalamic modulation (perceptual retouch) finds objects and events put in conscious format as coherent, integrated wholes. On the other hand, some special experimental conditions can lead to unusual effects. Where different feature processing can be dissociated artificially, illusory experiences involving dissociated features of the actually non-dissociated featured objects can be found. For example, Rick Cai and John Schlag (e.g., Cai & Schlag, 2005) and several others (e.g., Moutoussis & Zeki, 1997; Arnold, Clifford & Wenderoth, 2001) have produced illusory spatial dissociation of different properties of a gradually changing object. Thus, when an object in motion keeps changing its size and at one particular instant of size change its colour also changes, perceivers usually misalign the locations of shape-change and colour-change so that a wrong colour will be misbound with a wrong-sized object from the later instances of the changing shape (which continually changes its size and location). My interpretation assumes that the factor of fast spatial change permits the non-specific modulation that comes late to "find" the sensory traces of specific feature information in a dissociated state (i.e., not integrated as yet) and mis-binding is what appears in the form of a perceptual illusion.

Assuming that perception of objects presupposes binding of features into coherent perceptual entities (e.g., Cleeremans, 2003), it is strongly hypothesised that gamma-frequency (>40-Hz) oscillations in the frequency of firing of the specific neurons that mediate specific feature processing are an important mechanism of feature binding. A given actual subset of feature-neurons selected out of the endless number of various possible combinations fires in synchrony – this is hypothesised to be the set of units which are representative of the actual specific input (e.g., a red-green tilted-striped skirt and *not* a blue-green horizontally-striped or a black-white vertically-striped skirt).

One of the main effects of modulation from the nonspecific thalamus in upgrading the activity of specific cortical driver neurons (e.g., pyramidal neurons in the visual cortex) up to the level of explicating their contents for conscious perception consists in generating 40-Hz synchrony between the specific representational system activity and the nonspecific intralaminar thalamic-system activity (Llinás & Ribary, 2001; Newman, 1995; Steriade et al., 1997). Because the non-specific counterpart of the synchronizing activity has a long delay after a stimulus onset, formation of consciously perceived objects should be also delayed. Also, the focus of explicit perceptual activity should be related to coherent wholistic objects as the main matter of the retouch process. Therefore, we have now two types of binding processes: (a) pre-conscious binding of single localised features (represented by specific driver neurons) into coherent objects or events with the actual features integrated; (b) binding of the being-binded pre-conscious objects (events) into consciously apprehended visual scene by synchronisation that originates from the non-specific thalamus (the process of "retouch").

2.3. A sample of experimental effects and the retouch theory

If a stimulus-object (B) is presented alone or with long time intervals after a preceding, spatially overlapping stimulus-object (A), it is perceived clearly and described adequately. If B is presented very soon after A (e.g., onset of B trailing the onset of A by about 50-100 ms), it can be well perceived again, but it may be perceived also as having higher luminance contrast compared to its isolated presentation (Bachmann, 1988). Moreover, the delay with which it appears in consciousness can be shortened by about 20-60 ms, a phenomenon termed perceptual latency priming (Bachmann, 1989; Klotz & Wolf 1995; Scharlau, 2004; Scharlau & Neumann, 2003). This effect is present even when the preceding stimulus remains pre-conscious. According to the retouch theory, the perception saliency and speed of B are facilitated because A initiated the retouch process and B "uses" this for facilitating its own processing efficiency.

In real life, isolated, brief objects in an empty field are very rare. Usually there is a flow of sensory input signals. We can experimentally juxtapose perception of isoloated test-objects and the test-objects that appear within streams of sensory input. It appears that target objects (e.g., letter **Z**) within streams of unchanging items (e.g., letter **I**) tend to be perceived before the replicas of the same **Z**-targets that are simultaneously presented out of stream, in isolation (Bachmann & Põder 2001; Bachmann et al., 2003). The perceptual speeding-up of microgenesis of the samples of in-stream stimulation accumulates and sampling obtains its maximum perceptual speed within the first 150 ms from the stream onset (the temporal advantage of the in-stream target perception over the isolated target perception amounting to 60-80 ms). In-stream temporal facilitation somewhat subsides at the subsequent epochs, however persists with the value of about 30 ms up to at least 1 second long epochs. Not only the *speed* of microgenesis depends on where in a stream the critical item is located and whether it is located in stream at all: If two target-objects are successively presented within a stream of invariant **I**s, the first target prevails in conscious perception if presented in the initial stream epoch (50-150 ms), but the second target gets upper hand at later stream epochs (Bachmann & Sikka, 2005). It is evident that the way how an item in stream is perceived is not only a function of the item itself and even not the function of the item's immediate neighbours in stream, but is also a function of the properties of in-stream stimulation over a longer range of intervals.

The time values and the nature of the in-stream facilitation effect allow to consider this effect as a variety of the well-known flash-lag effect (FLE; see Nijhawan, 1994; Sheth et al., 2000; Krekelberg & Lappe, 2001; Eagleman, 2001; Whitney, 2002; Brenner & Smeets, 2000; Kreegipuu & Allik, 2003; Öğmen, Patel, Bedell, & Camuz, 2004). In both cases the isolated presentation of the target is juxtaposed experimentally with the continuously changing stimulus (e.g., a moving object that changes its spatial position or an object that continuously changes its colour or spatial frequency). In both cases, perception of the "flashed", out-of-stream, object lags behind the perception of the object that is undergoing continuous change or is presented within the continuously presented stream of sensory input. Because FLE occurs both in streams with change in the feature values of the streamed sensory input (Sheth, Nijhawan, & Shimojo, 2000) and in spatially invariant streams without change in feature values of the streamed input (Bachmann & Põder, 2001; Bachmann & Oja, 2003), the in-stream facilitation effect appears to result from streamness as such. Therefore, FLE is a phenomenon that does not necessarily require a change in the spatial location of continuous streamed input nor change in the perceptual feature-attributes of the streamed input. I hypothesise that FLE is the result of the workings of the perceptual retouch mechanism(s) where sensory input that is preceded closely in time by a spatially overlapping or adjacent other input will be processed up to the level of explicit visibility (conscious awareness) faster than the same input when it is presented in isolation (is flashed). I call this effect *in-stream facilitation*.

Let us see how the perceptual retouch account where specific processing of streamed sensory contents is supplemented by non-specific thalamic modulation can help to explain all varieties of FLE. Initially, at the onset of stimulation where there has not been any locally preceding visual input, such as with flashed target or the beginning of the stream, the modulation through NSP (or the change in the dynamic characteristics of this modulation) takes considerable time to become effective. Consequently, the latency of sampling of the specific signals for explicit representation is slow. With time passing, the already presented input has had progressively more time for initiating and setting the effective non-specific modulation. Each new shipment of the specific signals, sent via the fast afferent pathways to the cortical driver-units that carry stimulus-specific information (including precise spatial position) will be therefore upgraded for explicit representation with progressively smaller visual latency. There is acceleration of the speed of conscious perception online with accumulating perceptual stimulation. After the time interval that is typical for the duration of the whole effective cycle of specific plus modulatory afferent activity has passed (i.e., about 100-200 ms), the acceleration stage is over. From now on, the information-sampling speed for explicit representation stays at a more or less stable level, but higher than it was initially. This explains why target within stream comes to awareness sooner than the flashed target presented without the locally preceding input out of the stream.

Thus, summa summarum, explicit perception of streamed sensory input combines two binding processes – fast binding of specific features into pre-conscious object representations and a slower binding of the bound pre-conscious objects with the non-specific retouch process allowing these objects gradually enter the conscious mode of representation.

Recent approaches to feature binding and explicit perception have found that a substantial role in mediating binding operations may be played by the gamma-frequency (>40-Hz) synchronised oscillations in the neuronal activity of the brain. This fits well with perceptual retouch theory standpoints. In numerous studies the individual consciousness-generating and feature-binding functions of the non-specific thalamus (the NSP) are made to co-exist (e.g., Llinás, Ribary, Contreras, & Pedroarena, 1998; Llinás & Ribary, 2001; Llinás, Urbano, Leznik, Ramirez, & van Marle, 2005; Singer, 1998; Steriade, 2000). The specific classical afferent pathways pass on specific environmental information and establish gamma-frequency synchronised activity within the neuronal assemblies that stand for the actual sensory input (including objects). For this information to become integrated into conscious perceptual image, the respective activity should be synchronised with another gamma-band synchroneous oscillatory activity generated by the NSP. Because the set of specific neurons which are participating in the oscillatory compound may be recruited as varying neuronal pools (that represent different objects and/or their features) at different epochs of time, illusory dissociations in the specific characteristics of the actually coherent and integrated objects and events can be expected. If at the moment when nonspecific synchronising activity arrives, only part of the specific information pertaining to target object has been encoded by the specific neurons, only that part will be integrated into conscious representation at that moment. Specific information that is slower in being established as a specific representation "enters" integrated dynamic core at a later moment. Synchronisation with NSP-modulation need not be localised in time at one narrow instant. The very concept of synchronised oscillations is a process-concept *par excellence*. It is a *process* of representing, not a "thing" or "place". It is a continuous, temporally extended entity. Thus, naturally, dissociation in time of separate features of an object can be possible. But to obtain this, special artificial conditions have to be used where the natural speed of specific-plus-nonspecific integration through gamma-range synchrony cannot cope with the spatiotemporal physical properties of experimental stimulation. Visual masking, Cai and Schlag effect, and FLE paradigms – all these nicely exemplify this.

Despite that NSP-mediated conscious experience comes slowly, specific brain areas are processing specific sensory information in the meantime. Experiments by Mitroff and Scholl (2004, 2005) demonstrate that even if sensory objects are not represented in consciousness (for instance, this can be guaranteed by the phenomenon of motion-induced blindness – see Bonneh et al., 2001), their specific characteristics continue to be processed pre-consciously. Thus, when a bar gradually changes its orientation while out of awareness and is switched off abruptly, it re-enters conscious perception. But it does this with having feature value (orientation) it acquired during the preconscious time of processing, but not the last value of when it was consciously perceived before disappearing from awareness. Updating functions can be well performed by the specific system preconsciously. Perceptual retouch is applied to the updated, not to the old contents even if updating itself proceeds on out of awareness. The content system for cognition is autonomous from the consciousness-property, system, of cognition.

3. Perceptual retouch and mental imagery

For many years, specialists have argued over whether perception and imagery are distinct, separate processes or largely overlapping cognitive processes (Kosslyn, 1994; Pylyshyn, 2003). Evidence showing activation of same visual cortical areas in perception and in producing mental images and facts about mutual interference between visual perception and actively generated mental images seem to support the notion that perception centers and imagery centers in brain are the same or largely overlapping, except the earliest (receptor-linked) levels of perceptual channels (Kosslyn, 1994). Sometimes this approach has been named 'picture-in-the-head', approach. If visual imagery would be just another, although a quantitatively slightly weaker (less detailed or salient) instance of activation of the sensory-perceptual centers in the brain which work according to the analog-format representational principles, many of the properties of manipulation and interrelation encountered between real-world objects would be repreated in the mental representational format. Thus, if subjects who well command chess should visualize moving chess pieces up and/or right and/or left and/or down within a grid (following the experimenter's commands) and over many successive moves, it should make no difference whether the piece which the player mentally moves is a piece for which the moves are legal (e.g., rook) or illegal (e.g., bishop). In physical reality, the task of moving game pieces between grid squares should be accomplished with the equal precision despite of the symbolic meaning of the piece. However, if subjects would make more mistakes when moving illegally in their imagination, the analog-format simple theory would be in trouble.

Empirical evidence supports the notion that imagery is not just another, 'internal' sensorimotor activity within an analog-format problem space: because in the experiment there are more mistakes with illegal moves, the imagery process should necessarily include propositional factors that derive from the symbolic meaning of the imaged objects (Bachmann & Oit, 1992).

There seems to be a controversy between analog-format depiction's quality and structure in mental images on the one hand (similarly to these properties of actual perception and related to the activity of the brain centers subserving actual perception) and propositional sensitivity and symbolic dependence of mental images on the other hand. Imagery provides us with an exceptional psychological freedom: being here in chilly Estonia right now surrounded by the views of my computer, items on the desk, my wristwatch, *et cetera*, I can easily get free from the constraints the actual environment sets on me and mentally travel to lake Como or La Scala, or Ischia Island (or even to Ferrari headquarters) and experience memory images (e.g. Como or Vesuvius) or phantasy images (Ferrari garages). I am free to mentally experience whatever I like (provided the constraints of my episodic memory contents and inventiveness) both from distant familiar reality and from the realm of that kind of virtual reality which is put together in a way of unprecedented combinations of the familiar elementary items. Say, so as to create a preposterous phantasy image such as myself playing the Devil'sTrill on a mountaintop in Antarctic and using contrabass for this purpose.

How the retouch theory can deal with the task to explain how the experience of mental images may be produced? Actual sensory input is absent, so the non-specific modulating system has to be fed by some signals from other sources. Thus there is the possibility for choosing symbolic reference and propositional and predispositional stimulation that comes endogenously from memory and from the currently activated associative activity. Most probably the control functions for what becomes the contents of mental images are predetermined by prefrontal and (especially left) parietal cortical activity and more wide-spread associative activity in distributed cortical locations related to analog-format and symbolic representations (Kosslyn, 1994; Ganis et al., 2004; Sack at al., 2005; Saariluoma et al., 2004). This level of mental activity that can operate also preconsciously (see, e.g., Owen et al., 2006) instigates cortico-thalamic feedback loops that in turn activate retouch processes that reverse the facilitative flow back to cortex. As soon as this interactive looping system consisting in content-specific oscillatory activity and non-specific thalamic oscillatory activity is becoming synchronized, a combination of specific representational units in cortex becomes singled out as the actually bound permutation variant of the perceptual elements. This elemental combination makes up the contents for the currently active mental image. Insofar as (and because) sensory flow from receptors is weak or irrelevant for the current attentional focus, mental images can compete with actual sensory input, being salient enough to be experienced, followed, understood and described. They become even relatively stronger when sensory input is weak (e.g., darkness, eyes closed).

Whereas the cortical specific areas in case of perception and imagery that are modulated by retouch activity from thalamus are overlapping and because mental images are strongly memory-dependent, we should conclude that the level of analog-format implementation of the content that is derived from long term memory is residing at the locations in brain where also the perception content is represented. One's own thoughts, backed up by long-term memory codes and by currently ongoing pre-conscious mental activity steer the feedback to thalamic modulators which in turn generate consciousness for the selected specific contents instantiated by cortical driver-neuron units. But as the actual input channels from the receptive periphery are not involved in representing the mental image content, it can almost never be as stable, salient, detailed and rich as the real perceptual image. (Except in some unusual cases of eidetic imagery and influence of psychotropic substances.)

While the control functions in processing perceptual and mental images seem to be carried out by the same high-level cortical centers, there appear to be differences in how the reentrant activity is applied at the lower level cortical areas (Ganis et al., 2004): the spatial overlap between perception-related and imagery-related active areas in the temporal and occipital cortex is far from perfect, suggesting that retouch modulation in case of imagery may be targeted at the specific representational units that have a coarser and less distinct resolution in terms of representing detail. From the adaptive point of view, some differences in qualitative appearance between imagery and perception are parsimonious and advantageous because it should be important for a subject to understand what is reality (supported by a detailed and vivid and temporally coherently changing -- vis-à-vis real-time physical-world changes – retouched images) and what is just imagination (supported by a more vague, less seamless and less stable retouched image as the virtual reality). (See also O'Connor and Aardema, 2005, or Revonsuo, 2006, on related issues.)

Scientific facts that cortical areas which provide content-representations for dreaming consciousness as well as for visual environmental consciousness, overlap, necessitate that some minimum range of duality in brain mechanisms subserving consciousness is a *sine qua non*. Imagery and dreams as virtual reality that is relatively free from the actual environmental input (and despite of the objective presence of this input) are realities that require an autonomous mechanism that is relatively free to upgrade optional parts of the specific representational field onto conscious level and leave some other parts at a less well modulated saliency level. Otherwise there would not be mental freedom from the slavery of the actual environment. The perceptual mechanism described in the present article appears to be capable of fulfilling this autonomous function as the neurobiological mechanism that mediates consciousness. This position is not at odds with the views that conscious perception supposes processual unity between the activity in the brain and the perceived event in the external world (Manzotti, 2006). What matters is how substantial is the relative share of memory-based episodic information and receptor-mediated actual information in the retouched conscious image, generated by the help of the non-specific thalamus. I share with my reader the hope that my theoretical vision has not been caused in the same way as Tartini's Sonata G-Minor has been reputedly created. Emotions can be thrilling and trilling and our imagination can play tricks. Whether the calm bright truth is not different from my conceptual understanding or not remains to be seen.

Acknowledgments

I am grateful to Professor Mariano Bianca and his colleagues for the honour to be invited as a contributor to this special issue. The support from Estonian

Science foundation grant #5778 is also highly appreciated.

References

Arnold, D.H., Clifford, C.W.G., & Wenderoth, P. (2001). Asynchronous processing in vision. Color leads motion. *Current Biology*, 11, 596-600.

Baars, B. (1995). Surprisingly small subcortical structures are needed for the *state* of waking consciousness, while cortical projection areas seem to provide perceptual *contents* of consciousness. *Consciousness and Cognition,* 4, 159-162.

Baars, B. J. (1997). *In the Theater of Consciousness. The Workspace of the Mind*. Oxford. Oxford University Press.

Bachmann, T. (1984). The process of perceptual retouch. Nonspecific afferent activation dynamics in explaining visual masking. *Perception & Psychophysics*, 35, 69-84.

Bachmann, T. (1988). Time course of the subjective contrast enhancement for a second stimulus in successively paired above-threshold transient forms. perceptual retouch instead of forward masking. *Vision Research*, 28, 1255-1261.

Bachmann, T. (1989). Microgenesis as traced by the transient paired-forms paradigm. *Acta Psychologica*, 70, 3-17.

Bachmann, T. (1994). *Psychophysiology of Visual Masking. The Fine Structure of Conscious Experience*. Commack, NY. Nova Science Publishers.

Bachmann, T. (1997). Visibility of brief images. The dual-process approach. *Consciousness and Cognition*, 6, 491-518.

Bachmann, T. (1999). Twelve Spatiotemporal Phenomena, and One Explanation. In. G. Ascersleben, T. Bachmann, and J. Müsseler (Eds.). *Cognitive Contributions to the Perception of Spatial and Temporal Events.* Amsterdam. Elsevier, 173-206.

Bachmann, T. (2000). *Microgenetic Approach to the Conscious Mind*. Amsterdam / Philadelphia. John Benjamins.

Bachmann, T. (2005). Object substitution and its relation to other forms of visual masking. reply to James Enns. *Vision Research*, 45, 381-385.

Bachmann, T., & Põder, E. (2001). Change in feature space is not necessary for the flash-lag effect. *Vision Research*, 41, 1103-1106.

Bachmann, T., & Oit, M. (1992). Stroop-like interference in chess-players' imagery. An unexplored possibility to be revealed by the adapted moving spot task. *Psychological Research/Psychologische Forschung*, 54, 27-31.

Bachmann, T., & Oja, A. (2003). Flash-lag without change in feature space is alive and well at late intervals after stream onset. *Perception*, 32S, 126-127.

Bachmann, T., & Sikka, P. (2005). Perception of successive targets presented in invariant-item streams. *Acta Psychologica*, 120, 19-34.

Bachmann, T., Luiga, I., Põder, E., & Kalev, K. (2003). Perceptual acceleration of objects in stream. Evidence from flash-lag displays. *Consciousness and Cognition*, 12, 279-297.

Bogen, J. E. (1995). On the neurophysiology of consciousness. I. An overview. *Consciousness and Cognition*, 4, 52-62.

Bonneh, Y., Cooperman, A., & Sagi, D. (2001). Motion-induced blindness in normal observers. *Nature*, 411, 798-801.

Breitmeyer, B. G. (1984). *Visual Masking. An Integrative Approach*. Oxford. Clarendon.

Brenner, E., & Smeets, J.B.J. (2000). Motion extrapolation is not responsible for the flash-lag effect. *Vision Research*, 40, 1645-1648.

Cai, R., & Schlag, J. (2005). A new form of illusory conjunction between color and shape. (paper submitted to *Issues of Space and Time in Perception and Action*, R.Nijhawan (ed.))

Churchland, P. S. (2002). *Brain-Wise. Studies in Neurophilosophy*. London. MIT Press / Bradford.

Cleeremans, A. (Ed.) (2003). *The Unity of Consciousness. Binding, Integration, and Dissociation*. Oxford. Oxford University Press.

Crick, F. (1984). Function of the thalamic reticular complex. The searchlight hypothesis. *Proceedings of the National Academy of Sciences USA*, 81, 4586-4590.

Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, 8, 97-107.

Crick, F., & Koch, C. (1998). Constraints on cortical and thalamic projections. the no-strong-loops hypothesis. *Nature*, 391, 245-250.

Crick, F., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, 6, 119-126.

De Gelder, B., de Haan, E., & Heywood, C. (Eds.) (2001). *Out of Mind. Varieties of Unconscious Processes*. Oxford. Oxford University Press.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness. basic evidence and a workspace framework. *Cognition*, 79, 1-37.

Di Lollo, V., Enns, J.T., & Rensink, R.A. (2000). Competition for consciousness among visual events. the psychophysics of re-entrant visual processes. *Journal of Experimental Psychology. General,* 129, 481-507.

Dixon, N. F. (1981). *Preconscious Processing*. Chichester. Wiley.

Eagleman, D.M., & Sejnowski, T.J. (2000). Motion integration and postdiction in visual awareness. *Science*, 287, 2036-2038.

Edelman, G. M., and Tononi, G. (2000). *A Universe of Consciousness. How Matter Becomes Imagination*. New York. Basic Books.

Enns, J.T. (2004). Object substitution and its relation to other forms of visual masking. *Vision Research*, 44, 1321-1331.

Flavell, J. H., & Draguns, J. G. (1957). A microgenetic approach to perception and thought. *Psychological Bulletin*, 54, 197-217.

Ganis, G., Thompson, W.L., & Kosslyn, S.M. (2004). Brain areas underlying visual mental imagery and visual perception. An fMRI study. *Cognitive Brain Research*, 20, 226-241.

Hanlon, R. E. (Ed.) (1991). *Cognitive Microgenesis. A Neuropsychological Perspective.* New York. Springer-Verlag.

Hochstein, S., & Ahissar, M. (2002). View from the top. Hierarchies and reverse hierarchies in the visual system. *Neuron*, 36, 791-804.

Jaśkowski, P., van der Lubbe, R., Schlotterbeck, E., & Verleger, R. (2002). Traces left on visual selective attention by stimuli that are not consciously identified. *Psychological Science*, 13, 48-54.

Kanwisher, N. (2001). Neural events and perceptual awareness. *Cognition*, 79, 89-113.

Kinoshita, S., & Lupker, S. J. (Eds.) (2003). *Masked Priming. The Sate of the Art*. New York. Psychology Press.

Klotz, W., & Wolff, P. (1995). The effect of a masked stimulus on the response to the masking stimulus. *Psychological Research / Psychologische Forschung*, 58, 92-101.

Kosslyn, S.M. (1994). *Image and Brain. The Resolution of the Imagery Debate*. Cambridge, MA. MIT Press.

Kreegipuu, K., & Allik, J. (2003). Perceived onset time and position of a moving stimulus. *Vision Research*, 43, 1625-1635.

Krekelberg, B., & Lappe, M. (2001). Neuronal latencies and the position of moving objects. *Trends in Neurosciences*, 24, 335-339.

LaBerge, D. (1997). Attention, awareness, and the triangular circuit. *Consciousness and Cognition*, 6, 149-181.

Lamme, V.A.F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, 7, 12-18.

Lange, N. N. (1892). The law of perception. I-II. (in Russian). *Voprosy Filosofii I Psikhologii, Spetsial'nyi Otdel*, 13, 18-37.

Lange, N. N. (1892). The law of perception. III. (in Russian). *Voprosy Filosofii I Psikhologii, Spetsial'nyi Otdel,* 14, 44-54.

Lange, N. N. (1892). The law of perception. IV-V. (in Russian). *Voprosy Filosofii I Psikhologii, Spetsial'nyi Otdel*, 15, 55-68.

Lange, N. N. (1892). The law of perception. VI-VII. (in Russian). *Voprosy Filosofii I Psikhologii, Spetsial'nyi Otdel*, 16, 25-38.

Llinás, R., & Ribary, U. (2001). Consciousnes and the brain. The thalamocortical dialogue in health and disease. *Annals of the New York Academy of Sciences*, 929, 166-175.

Llinás, R., Leznik, E., & Urbano, F.J. (2002). Temporal binding via cortical coincidence detection of specific and nonspecific thalamocortical inputs. A voltage-dependent dye-imaging study in mouse brain slices. *Proceedings of the National Academy of Sciences USA*, 99, 449-454.

Llinás, R., Ribary, U., Contreras, D., & Pedroarena, C. (1998). The neuronal basis for consciousness. *Philosophical Transactions of the Royal*

Society of London. B, 353, 1841-1849.
Llinás, R., Urbano, F.J., Leznik, E., Ramirez, R.R., & van Merle, H.J.F. (2005). Rhythmic and dysrhythmic thalamocortical dynamics. GABA systems and the edge effect. *Trends in Neurosciences* (in press; doi 10.1016/j.tins.2005.04.006).
Magoun, H. W. (1958). *The Waking Brain*. Springfield, Illinois. C. C. Thomas.
Manzotti, R. (2006). A process oriented view of conscious perception. *Journal of Consciousness Studies*, 13, 7-41.
Marcel, A. J. (1983). Conscious and unconscious perception. experiments on visual masking and word recognition. *Cognitive Psychology*, 15, 197-237.
Mitroff, S.R., & Scholl, B.J. (2004). Last but not least. Seeing the disappearance of unseen objects. *Perception*, 33, 1267-1273.
Mitroff, S.R., & Scholl, B.J. (2005). Forming and updating object representations without awareness. evidence from motion-induced blindness. *Vision Research*, 45, 961-967.
Moutoussis, K., & Zeki, S. (1997). A direct demonstration of perceptual asynchrony in vision. *Proceedings of the Royal Sosiety of London.* Series B, 264, 393-399.
Moutoussis, K., & Zeki, S. (2002). The relationship between cortical activation and perception investigated with invisible stimuli. *Proceedings of the National Academy of Sciences USA*, 99, 9527-9532.
Naccache, L., & Dehaene, S. (2001). The priming method. imaging unconscious repetition priming reveals an abstract representation of number in parietal lobes. *Cerebral Cortex*, 11, 966-974.
Newman, J. (1995). Thalamic contributions to attention and consciousness. *Consciousness and Cognition*, 4, 172-193.
Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, 370, 256-257.
O'Connor, K.P., & Aardema, F. (2005). The imagination. cognitive, pre-cognitive, and meta-cognitive aspects. *Consciousness and Cognition, 14*, 233-256.
Öğmen, H., Patel, S.S., Bedell, H.E., & Camuz, K. (2004). Differential latencies and the dynamics of the position computation process of moving targets, assessed with the flash-lag effect. *Vision Research*, 44, 2109-2128.
Owen, A.M., Coleman, M.R., Boly, M., Davis, M.H., Laureys, S., & Pickard, J.D. (2006). Detecting awareness in the vegetative state. *Science*, 313, 1402.
Purpura, D. (1970). Operations and processes in thalamic and synaptically related neural subsystems. In F.O. Schmitt (Ed.). *The Neurosciences. Second Study Program.* New York. Rockefeller University Press, 458-470.
Purpura, K. P., & Schiff, N. D. (1997). The thalamic intralaminar nuclei. A role in visual awareness. *The Neuroscientist*, 3, 8-15.
Pylyshyn, Z. (2003). *Seeing and Visualizing. It's Not What You Think*. Cambridge, MA. MIT Press.

Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews. Neuroscience*, 3, 261-270.

Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision Research*, 40, 1469-1487.

Revonsuo, A. (2006). *Inner Presence. Consciousness as a Biological Phenomenon*. Cambridge, MA. MIT Press.

Riffert, F. (2004). Whitehead's theory of perception and the concept of microgenesis. *Concrescence. The Australasian Journal of Process Thought*, 5, online.

Rosenthal, V. (2002). Microgenesis, immediate experience and visual processes in reading. In A. Carsetti (Ed.). *Seeing and Thinking*. Amsterdam. Kluwer.

Saariluoma, P., Karlsson, H., Lyytinen, H., Teräs, M., & Geisler, F. (2004). Visuospatial representations used by chess experts. A preliminary study. *European Journal of Cognitive Psychology*, 16, 753-766.

Sack, A.T., Camprodon, J.A., Pascual-Leone, A., & Goebel, R. (2005). The dynamics of interhemispheric compensatory processes in mental imagery. *Science*, 308, 702-704.

Sander, F. (1962). Experimentelle Ergebnisse der Gestaltpsychologie. In F. Sander, and H. Volkelt, (Eds.). *Ganzgeitspsychologie.* Munich. Beck. (Reprinted from E. Becher, ed., *10 Kongress bericht experimentelle Psychologie,* 1928 (23-87). Jena. Fischer).

Scharlau, I. (2004). The spatial distribution of attention in perceptual latency priming. *Quarterly Journal of Experimental Psychology*, 57, 1411-1436.

Scharlau, I., & Neumann, O. (2003). Temporal parameters and time course of perceptual latency priming. *Acta Psychologica*, 113, 185-203.

Sherman, S.M., & Guillery, R.W. (1998). On the actions that one nerve cell can have on another. Distinguishing "drivers" from "modulators". *Proceedings of the National Academy of Sciences USA,* 95, 7121-7126.

Sheth, B.R., Nijhawan, R. & Shimojo, S. (2000). Changing objects lead briefly flashed ones. *Nature Neuroscience,* 3, 489-495.

Singer, W. (1998). Consciousness and the structure of neuronal representations. *Philosophical Transactions of the Royal Society of London.* B, 353, 1829-1840.

Smith, G. (1957). Visual perception – an event over time. *Psychological Review*, 64, 306-313.

Steriade, M. (1996). Arousal. Revisiting the reticular activating system. *Science*, 272, 225-226.

Steriade, M. (1996). Awakening the brain. *Nature*, 383, 24-25.

Steriade, M. (1997). Synchronized activities of coupled oscillators in the cerebral cortex and thalamus at different levels of vigilance. *Cerebral Cortex*, 7, 583-604.

Steriade, M. (2000). Corticothalamic resonance, states of vigilance and

mentation. *Neuroscience*, 101, 243-276.
Steriade, M., Jones, E.G., & McCormick, D.A. (1997). *Thalamus*. Oxford. Elsevier.
Steriade, M., Jones, E.G., & Llinás, R.R. (1990). *Thalamic Oscillations and Signalling.* New York. Wiley.
Treisman, A. (1998). Feature binding, attention and object perception. *Proceedings of the Royal Society of London.* B, 353, 1295-1306.
Undeutsch, U. (1942). Die Aktualgenese in ihrer Allgemein-Philosophischen und ihrer characterologischen Bedeutung. *Scientia*, 72, 37-42, 95-98.
VanRullen, R., & Koch, C. (2003). Competition and selection during visual processing of natural scenes and objects. *Journal of Vision*, 3, 75-85.
Werner, H. (1940). *Comparative Psychology of Mental Development.* New York. Harper.
Werner, H. (1956). Microgenesis and aphasia. *Journal of Abnormal Social Psychology*, 52, 347-353.

Talis Bachmann
University of Tartu
talis.b@lawinst.ee

Mariano L. Bianca - Lucia Foglia
University of Siena

Non-Perceptive Mental Image Generation: a Non-Linear Dynamic Framework

Abstract Mental imagery is an important topic in classical and modern philosophy, as it is central to the study of knowledge; since subjects can recall features of perceptual experiences in different ways and times, even modifying their structure, in this brief essay we will focus on non-perceptive mental images (NPMI, i.e., mental imagery) and to this purpose we will analyse, on the one hand, the nature of perceptive mental images (PMI, i.e., visual experiences); on the other hand, NPMI generation according to different strategic conditions and retrieval modalities and, so, their structural relationships with PMI. Thus, we will clarify what NPMI amounts to and, to this purpose, we will address the issue of epistemic correlation and semantic reference. As regards the former, we will adopt the notion of 'belief', both to describe which attitude supports the use of an NPMI and to account for its relationship with a PMI. As regards the latter, we will talk about weak ontology between an NPMI and the world, to conceive 'the way in which' mental imagery, despite the absence of a sensorial modification, refers to physical objects. As the subject's mental contents determine, but do not constrain, what NPMI is like, we will deal with the 'involvement' of prior knowledge and the notion of meaningful space to underline that its generation is always the result of a wide process which involves perceptive visual information already processed and stored together with further mental contents. So, though some Authors consider mental imagery to be a 'percept-like' image and, somehow, espouse the theory of reactivation, we will pose a challenge to this theory – i.e. to what we call mental imagery standard framework (MISF) – describing an NPMI in terms of configurational setting and indeterminacy. Consequently, we will propose a non-linear dynamic framework of mental imagery generation (MINLF), according to which an NPMI is not reducible to a mere recalling of figural features of a perceptive image.

Key words Visual perception, mental imagery, retrieval modalities, retainment, configurational setting, indeterminacy, meaningful space, isomorphism, ontology.

1. Introduction

Although mental imagery is a common experience, it is quite difficult to pin down its precise meaning, and different understandings of it have added further material to an already complex debate. Concerning imagery's nature, generation, relationship with visual experience and its function several hypotheses have been formulated and among psychologists, cognitive scientists and philosophers the term *imagery* has been used, at least, in two main, though very conflated, senses: a) percept-like image, that is, a quasi-perceptual experience; b) picture-like mental representation (which gives rise to a percept-like image). The support of experimental findings and the use of special devices like PET and fMRI, while encouraging a better understanding of it, explaining what happens to the brain while processing an imagery, does not completely account for what 'it is actually like'. In any case, we can say that several hypotheses about visual experience and mental imagery (from now on PMI and NPMI[1]) have been put forward, which, though different from one another and not all entirely shared by the majority of authors, outline PMI, NMPI and their relationship in the following ways: a) a PMI is a mental content stored in the LTM (Long-Term Memory); b) stored PMI remain as they are and cannot be completely modified; c) a stored PMI can be retrieved 'as it is' to give rise to an NPMI; d) an NPMI generation is a mere reactivation of perceptive visual information already processed and, so, it is not an *active process*; e) an NPMI always refers to a PMI, at least, to its figural features; f) an NPMI generation can be influenced or not by prior knowledge; g) an NPMI always refers to the original PMI, i.e., the PMI generated at the time of the first visual experience of the physical object or event; h) NPMI, though formulated at different times, are similar to one another and refer to a PMI; i) an NPMI generation retains, mainly, the figural features of the previous PMI and not the cognitive or mental attributes assigned to it.

These hypotheses[2] taken all together give rise to what we call *mental imagery standard framework* (MISF) which considers LTM as a 'store' of visual information which cannot be modified during an NPMI generation process and NPMI, more or less, as 'replica' of such information. Since we intend to pose a challenge to this standard framework, we will present the generation of an NPMI as a dynamic, non-linear process which, each time, gives rise to a *new configurational setting* of perceptive visual data[3] within a specific *cognitive frame* (for this notion see Section 3.1.c). In addition, we will argue that a PMI recalling does not concern, though sometimes it may, only its figural features but its significations, values, and semantic reference, also called *meaningful space* (see Section 3.1.c, where this notion will be explained in detail) and, thus, mental imagery generation, besides perceptive visual information, involves different mental variables[4]. Since these variables can deeply influence, and sometimes determine, both its structure[5] and its relationship with a PMI, it will be possible to claim that '*given a PMI, various NPMI can be derived under different conditions and retrieval modalities*'[6] (see Section 4). Remarking that an NPMI is always a new configurational setting of a PMI_{st}[7], on the one hand, we will refute the notion of percept-like image, and on the other hand, we will argue that the re-enactment of perceptive visual information is not a simple retrieval but the result of a non-linear process in which such information is recalled and enriched by different significations and values. Thus, we will introduce and account for the *principle of irreducibility*, according to which the content of an NPMI is not reducible to that of a PMI_{st}, even in its figural features, and though related to perceptive images already stored and to different cognitive and non-cognitive mental contents, its generation is always an *autonomous neuromental event*. Consequently, we will sketch a mental imagery non-linear framework (MINLF) and we will also introduce new epistemological and epistemic notions which, in many ways, are not included in MISF. Even though in this paper MINLF is just drafted and needs to be fully developed, we believe that neurophysiological findings and psychological results can, somehow, support it and overcome, at least, some of the assumptions included in the MISF.

2. *Facultas Imaginandi* or visual mental activity

The question of mental images has been addressed from classic Greek philosophers up to contemporary psychology and neuroscience. Since their nature and relation with perceptual experience and upper cognitive levels are difficult to pin down, they are still a matter of debate; nevertheless, we can note that, for the majority of Authors, *mental images* refer to what is generated in the absence of the appropriate external stimuli and never are used to mean the result of a visual perception.

> In common usage, a mental image occurs when a representation of the type created during the initial phases of perception is present but the stimulus is not actually being perceived; such representations preserve the perceptible properties of the stimulus and ultimately give rise to the subjective experience of perception. As this characterizing makes clear, we are not limiting mental imagery to the visual modality. Although visual imagery is accompanied by the experience of 'seeing with the mind's eye,' auditory mental imagery is accompanied by the experience of 'hearing with the mind's ear,' and tactile imagery is accompanied by the experience of 'feeling with the mind's skin,' and so forth (Kosslyn, S., Thompson, W., Ganis, G., 2006).

Even if we agree with the common usage of the expression *mental image*, according to which an *image generation* occurs when physical stimuli are not actually perceived, and preserves the perceptible properties of the stimulus, we do not completely accept that it is like a subjective experience of perception, i.e., a "quasi-perceptual experience". We think rather that: a) the term *image* can also be used to mean visual perception; b) its generation, besides the physical properties of an object, can also retain its significances and values; c) imagery is different from "quasi-perceptual experience", though research on brain imaging and case studies in neuropsychology have shown similar neural architectures between visual perception and visual mental images[8] (Lambert, S., Sampaio, E., Scheiber, C., Mauss, Y., 2002). So, not only are visual perceptions to be called proper mental images, but they are also one of the most important processes belonging to visual mental activity, or rather, '*facultas imaginandi*', as medieval philosophers were used to considering it. In the history of Western philosophical thought, indeed, despite substantial differences among authors, for the majority of them this *facultas* was like a psychic activity under which mental image generation falls; shared by all human beings, this activity is what make it possible to give rise to specific visual processes (visual perception, mental imagery, imagination images) which, though different from one another[9], have in common the property of being *figural mental events*.

Thus, the idea that the same *mental activity* underlies the generation of different mental images has ancient roots; to be more precise, it comes from

Greek philosophy. As far as Aristotle is concerned, he seems to have held that imagery is like inner pictures, quite similar to those produced from sensorial impressions; in *De Anima* he actually says that imagination is different from sensation but it does not exist without it (*De Anima*, 427b 15) and, furthermore, that imagination is that by which we say that an image arises (*De Anima*, 428a). Aristotle, therefore, considered that imagination (*phantasia*) allows the passage from sensual experience to reason and, this way, defined *phantasia* as the capacity to have imagery. As regards St. Augustin, images are generated by physical objects through sensations and once formed they can be memorized, linked or distinguished from one another, reduced, amplified, ordered and composed in different ways (*De Vera Religione*, 10, 18). According to Hobbes, sensation is what generates an image when an object is present, whereas, what gives rise to an image when an object is not actually perceived can be called 'fantasy' or imagination. Thus, imagination is nothing but faded or weakened sensation, as the object it refers to is distant, or rather, absent (*De Corpore*, XXV, § 7). As regards Descartes, imagination refers to both visual perception and mental imagery: in such a way images can be derived from sensorial contents or from ideas of mind (*Meditations on First Philosophy*, Meditation VI). According to Locke and Hume, who seem to share the pictorial theory of imagination, mental images are fleeting and ephemeral and so, like Hobbes, they consider imagery a weakened and faded copy of the perceptive image of physical objects; Hume, actually, writes:

> When we remember any past event, the idea of it flows in upon the mind in a forcible manner; whereas in the imagination the perception is faint and languid, and cannot without difficulty be preserv'd by the mind steddy and uniform for any considerable time (*Treatise of the Human Nature*, I, III).

In *Critique of Pure Reason* Kant formulates a significant definition of *Imagination*[10] and saying that:

> Einbildungskraft ist das Vermögen, einen Gegenstand auch ohne dessen Gegenwart in der Anschauung vorzustellung (KrV, B151)

he describes it as 'the possibility to represent an object without its presence to the intuition'. In this way, imagery is not a 'faded picture-like image', rather, what reason deals with. It is worth briefly mentioning also what two other important thinkers have said about imagination and images, as they have contributed greatly to the imagery debate, i.e., Husserl and Sartre. According to Husserl, images are *analogon* of physical objects (*Analogon des Gegenstandes*) and, while visual perception (die *Wahrnemuhung*) can directly (*direkt*) 'present' its contents, imagery can be only their figural 're-presentation'. Husserl, actually, writes:

> Ist nicht in der reinen Wahrnemuhung der darstellende Inhalt identisch mit dem Gegenstande selbst? Das Wesen der reinen Präsentation besthet doch

darin, reine Selbstdarstellung des Gegenstandes zu sein, also den darstellenden Inhalt direct (in der Weise, des "Selbst") als ihren Gegenstand zu meinen. Doch das Wäre ein Trugschluß. Die Wahrnemuhung, als Präsentation, faßt den darstellenden Inhalt so, daß mit und in ihm der Gegenstand als selbst gegeben erscheint. Rien ist die Präsentation dann, wenn jeder Teil des Gegenstandes im Inhalte wirklich präsentiert und keiner bloß imaginiert oder symbolisiert ist. […] Die reine Bildvortellung, die ihren Gegenstand vermöge ihrer Reinheit von allen signitiven Zutaten vollständig verbildlicht, besitzt in ihrem darstellenden Inhalt ein vollständiges Analogon des Gegenstandes (*Logische Untersuchungen*, VI, 83).

In *L'imaginaire* Sartre underlines that imagery is an *act* (l'image est un acte) relating to an object that is absent or inexistent:

> Nous dirons en consequence que l'image est un acte qui vise dans sa corporéité un objet absent ou inexistent, à travers un contenu physique ou psychique qui ne se donne pas en propre, mais à titre de "*représentant* analogique" de l'objet visé (*L'imaginaire*, I, II).

So far, we have noted that philosophers have argued extensively about how we should consider the notion of 'mental image' and, since different accounts have been put forward, and sometimes conflated, we think that the best way to get a clearer idea at what it amounts to is to address the issue of visual mental activity together with its main visual processes 'in detail'. Once we have explained their meanings, it will be possible to pin down what we mean by the terms PMI and NPMI. Therefore, for the purposes of this paper, we will use the expression *visual mental activity* to mean the whole mental process that generates figural contents, precising that, such activity, can also be distinguished into perceptive and non-perceptive processes, which, in turn, give rise to different mental images:

a) *perceptive mental images* (best known as *visual perceptions*), i.e., images of things or events referring to a physical world 'actually' perceived;
b) *non-perceptive mental images* (best known as *mental imagery*) i.e., images of things or events formulated without physical stimuli 'actually' perceived;
c) *imagination images* (also called fantasy) i.e., images of things or events formulated without physical stimuli 'actually' perceived and not referring to the world as it is at the moment of their generation.

Though different from one another, such images have in common the feature of *being figural mental configurations* and, so, they belong to the same visual activity, as illustrated in Figure 1.

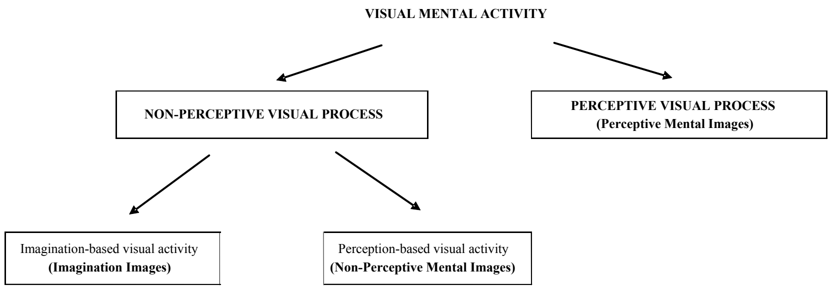

Figure 1. Visual Mental Activity consists of 2 main visual mental processes. Their end products are three different kinds of mental images.

As the Figure shows, there are two different kinds of visual processes and, since what they give rise to is, mainly, either visual perception, if something is present to the senses, or imagery, when something is present to the mind but the physical stimulus is not actually perceived, from now on we will use PMI to indicate visual perception and NPMI to mean mental imagery. Starting from this general outlook, we will deal with their generation process and we will also see that between mental images (PMI and NPMI) and physical objects, in specific conditions, there can be different ontological relationships. It is worth briefly mentioning, as we will later take it up in detail, that an NPMI can be generated *if and only if* the PMI, which it refers to, has been previously processed and stored; only in this way is it possible to establish a relationship between them, and to underline the different semantic references they have with physical objects. In this paper we do not take into account imagination images generation, though they play a remarkable role within visual mental activity.

3. Perceptive mental image (PMI)

> In recent years, the debates surrounding the issue of perceptual experience have become more complex. Lines of influence are often difficult to trace – and a number of independently motivated avenues of investigation have opened up (Gendler, T.Z., Howthorne, J., Edited by, 2006).

In psychology and cognitive science the term *visual perception* usually means the process of detecting, organizing and interpreting sensory data and, since we gather information about the world and act in it through our visual perceptions, these data are critical for knowledge and action[11]. However, the major problem is that what we see is not simply a translation of retinal stimuli, that is, images on the retina; thus, while perceiving and processing visual data, many other neuromental processes seem to occur[12]. Even though it is quite difficult to define what *visual perception* (or rather, a PMI) is about, for the purposes of this paper, it will be considered as a *neuromental*[13] *figural configuration;* more precisely, we can say that:

> *a PMI is the result of a perceptive visual activity i.e., the end product of a neuromental process by which the physical information transduced by the retina photoreceptors*[14] *is carried through different brain structures up to the striate and extrastriate visual cortical areas (such as V1, V2, V3, V4, and V5) to be further processed and spread over different cortical and non-cortical brain areas. The involvement of these last brain areas together with other inner mental contents contribute to the final visual information processing and, so, to the generation of a proper PMI.*

One of the most important features of a PMI is its *awareness*[15]; to have a perceptive image and to be aware of it means that a subject can describe it through a linguistic code, use or modify it both to serve different mental purposes and to act in the world.

Visual information processing, from a neurophysiological point of view, consists of three main steps or levels:

1) ***format level***: the neuronal pathway from the change of state of retina transducers to striate visual cortical area (also known as primary visual cortex or V1);
2) ***pattern integrated level***: the neuronal pathway from primary visual cortex to extrastriate visual cortical areas (also known as associative visual areas);
3) ***frame cognitive level***: the neuronal pathway from associative visual areas to different cortical, subcortical and non-cortical areas which gives rise to a *visual frame*, or rather, to a proper PMI semantically referring to an object of the physical world.

We can sketch such a sequence of steps that visual information takes as it flows from visual sensors to cognitive areas in Figure 2. It is worth briefly mentioning that the use of the term *flow* refers to the neuronal substrates whose main task is to transport visual information from one structure to another (e.g. optic nerve, optic chiasma, optic tract and so forth), whereas the use of the term *processing* means the way in which incoming visual data are processed and organized to generate structured information as output[16]. In the following section we will briefly analyse these three levels to highlight, on the one hand, the specific results they give rise to, on the other hand, the way in which they contribute to generating a PMI as the end product of the whole process.

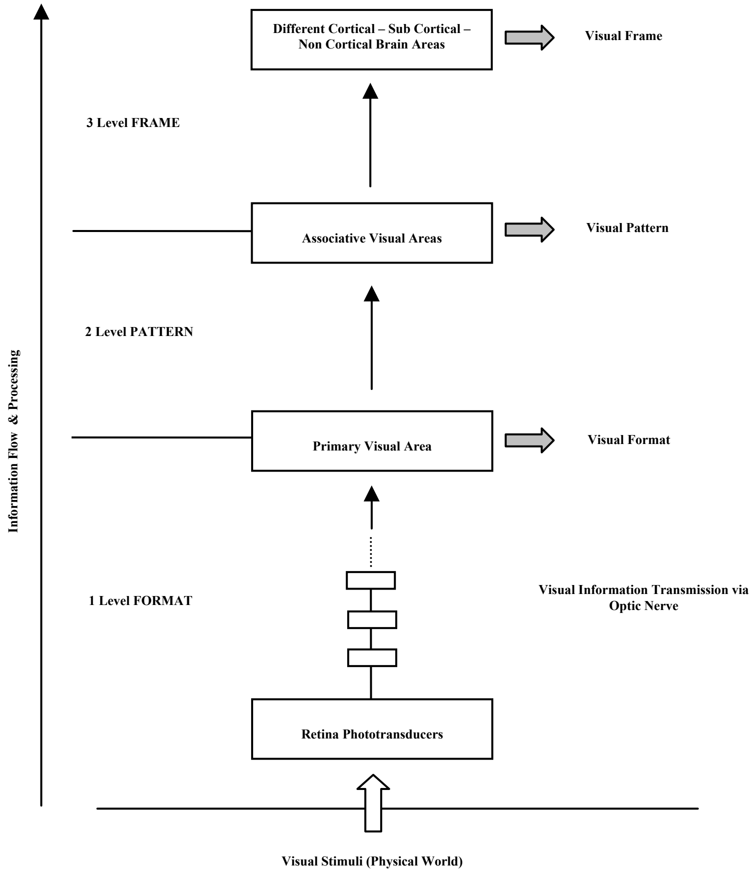

Figure 2. Information Flow and Processing is the sequence of steps that information takes as it flows from visual sensors (retina phototransducers) to cortical areas.

3.1 Information[17] flow and processing

a) Format level

Format level is mainly an *encoding process*, i.e., a process of transforming physical information into a neurochemical code, by which transduced visual information travels from the retina, through different neurophysiological structures, towards the primary visual cortex that is the brain area responsible for processing visual stimuli. The end product of this first level is not yet a PMI but just what we can call *visual format*, since to become a proper perceptive mental image such format needs to reach the upper levels and be further elaborated. In any case, the encoding process gives rise to the following main results:

1) the *translation* of physical information (photons) embodied in a visual stimulus into a neurochemical code[18];
2) the *informational format* (or rather, *visual format*) as output of the incoming information flow and processing. Though it is a primary organization form of the encoded information, a visual format cannot be considered a perceptive image.

b) Pattern level

The primary visual cortex, in turn, transmits the information processed into the visual format to two further brain pathways, called the ventral stream and the dorsal stream (Goodale and Milner,1992), where it is located and, somehow, recognized according to its shape and colour. Anyway, the involvement of these specific brain areas provides neither a complete description of the physical objects in the visual world nor their final identification; rather, it plays a crucial role in focusing their specific features. Hence, at this level, the incoming visual information is further processed and refined and, even if no perceptive mental image appears yet, what comes up is *an integrated visual pattern semantically referred to the world*. Though we will later take it up in detail (see Section 5.1), it is worth briefly mentioning here that *the visual pattern has a semantic reference to the world* and, therefore, that:

1) the physical stimulus brings with it information referred to the structure of the physical object in the world;
2) the neurochemical code of the physical stimulus *maintains* its structural information and, so, the main features of the physical object;
3) such encoded structural features are essentially retained through the pathways up to the associative visual areas.

Therefore, these three statements allow us to consider the *visual pattern* according to two different points of view: the ‘syntactical’ one refers to the pattern as a neuromental configuration whose structure is different from all the other neuromental contents because it consists of visual information processed in an integrated way; the ‘semantic’ one concerns its relationship with the physical world: the visual pattern retains in an encoded way, *via format level*, the structured information of the physical stimulus and, so, the features of the physical object.

c) Frame level

When integrated visual information, semantically referring to the world, leaves the second level and, thanks to a specific *branch-over*[19] process, reaches the upper one, then, visual pattern becomes a proper PMI (visual frame), or rather, a perceptive mental image endowed with specific attributes. Though we cannot investigate the branch-over process in detail, it will be useful to underline, at least, its two main results: a) *pattern recognition* and b) *enrichment* of the visual pattern.

a) Since visual mental activity does more than encoding and conveying figurative and spatial information along the brain structures, *recognition* is the process by which a set of visual information arranged into a certain pattern is *recognized* and *identified* as a specific *figural content*. So, recognition contributes remarkably to a PMI generation and its activation involves preexisting mental contents (*prior knowledge*), more specifically, what we may call *objects' figural type*. A *figural type* is the assembled figural features stored in LTM of an object and it arises from different visual experiences of similar objects. A figural type could be considered as a 'visual generalization'[20]. A visual pattern, for instance, is recognized as a 'black cat' if its figural features 'match' those of the *figural type* of the cat and it is semantically referring to physical cats. Actually, we can say that once a perceptive image is generated it is also cognitively recognized as a 'token' (specific case) of a figural type stored in one's mind. This does not mean that such a *prior knowledge* (and *figural types*) can 'deeply' modify visual information already processed in the previous levels; rather, it suggests that its involvement is useful for generating a complete PMI as an image of a specific object. We do not believe, as conceptualism would say, that what can be perceptually experienced is constrained by a subject's mental contents, because it would be impossible to perceive an object for the first time unless we do know something about it. Even if to experience a physical object, there is no need to possess an associated concept under which it can fall, to identify a visual pattern as a particular PMI (or as an image of an object) it seems to be necessary to draw on such a *prior knowledge*, i.e., information about the world already available, particularly, as figural types.
b) The second result of the branch-over process is the *enrichment* of the visual pattern through the different attributes that can be assigned to it; since the pattern is spread over numerous cortical and non-cortical areas, their involvement allows the generation of a complete perceptive image, i.e., a mental image which has, besides figural features and semantic

reference to physical object, also different kinds of *signification* and *value*. The former can be motivational, emotional, cognitive and so forth; the latter refers to the *worth of something* not only from an ethical point of view, i.e, a moral standard of behaviour, but also in terms of *relevance* in someone's life compared to things or events without any importance. Returning to the above example, we can say that a visual pattern cognitively recognized as a 'black cat' can be further processed and, through the assignment of different significations and values, it can become, ' a nice pet', 'an animal I would like to buy', 'a sign of bad luck' or even 'something I am scared of'. *Thus, a complete perceptive mental image is stored in LTM with its figural features, significations and values* and, in this way, it will be possible to 'recall' it to memory with all, or some of, the attributes originally assigned to it. So, the image of a 'black cat' will not only be an image of an animal with specific features but an image with significations, such as: a 'nice pet', 'something I would like to buy' or ' a sign of bad luck'.

It is, therefore, clear that the end product of the frame level, i.e., a complete perceptive mental image, is due to the contribution of different brain areas whose involvement, together with the influences exercised by other inner mental contents, is what makes a PMI a neuromental visual configuration useful to know and to live in the world[21]. Thus, a *meaningful* PMI can be used for different purposes, for instance, to plan an action or make a decision as well as to be linked to other mental images in order to give rise, in this way, to a wider neuromental configuration; in addition, it can also be available to further mental processes and, so, to be used for concept formation, inferences and so forth. Returning to the conceptualism debate mentioned above, it is worth specifying once more that preexisting knowledge, or rather, *prior knowledge*, contributes to but does not constrain a PMI generation, since its influences concern mainly the recognition process and the assignment of specific significations and values.

On the basis of the branch-over process and its two main results we can tentatively sketch the features of a PMI in the following ways:

1) a PMI retains the visual pattern as it has been processed in the second level (together with its attributes);
2) a PMI retains the same semantic reference to the physical world of the visual pattern, via format level;
3) a PMI is a specific state of awareness, i.e., what a subject is aware of at the moment of its formation, and stored this way, it can be re-enacted through an NPMI as a new state of awareness;
4) a PMI embodies the result of the recognition process and, in this way, a perceptive mental image becomes a 'token' of a *figural type* stored in LTM as an NPMI (this specific NPMI might also be considered as a visual generalization);

5) a PMI is the result of a branch-over process by which different significations and values are assigned to it; thus, a PMI is always embedded in what we call *meaningful space* in which there are its *figural features* with its significations, values and semantic references;
6) a PMI can be codified into a propositional code which describes its contents or assigns to it a proper name; in this way, such a perceptive image gains a propositional format; returning to the previous example, a PMI can be formulated through several propositional expressions, like: 'a black cat', 'this black cat is a lucky charm', 'the black cat belonging to my family' and so forth;
7) a PMI can be modified, consciously or not, and used both for further cognitive mental processes (such as: concept formation, descriptions, decision making, etc.) and to act in the world.
8) From all these features, it follows that:
9) a PMI is a *figural* or *iconic mental representation*[22] of a physical object and, for this reason, it can '*stand fo*r' it (see Section 5.1).

In the present context, it is necessary to distinguish between: 1) PMI processed when physical stimuli are actually perceived, for instance, images referred to the objects I can see, *now*, in my room, and 2) PMI truly stored in memory, i.e., the amount of perceptive information actually retained and available for further mental processes (from now on referred to as PMI_{st}), for instance, images referring to the objects I saw yesterday in my room that I can recall today; such a distinction is remarkable, at least, for two main reasons:

1) a PMI_{st} *can* be different from a PMI and *can* retain less information: all the visual information available, when a physical object is 'actually' present, may not always be completely stored in LTM
2) a PMI_{st} permanence in memory can be altered by different mental contents as well as by brain or psychic pathology and cellular ageing.

4. Non-Perceptive Mental Image (NPMI)

Despite the familiarity of the experience of imaging, the exact meaning of NPMI is remarkably hard to explain and, since the term often seems to swing towards imagination, it is necessary to pin down the way in which it will be used in order to avoid possible misunderstandings. An NPMI is a visual mental content whose generation occurs in the absence of physical stimuli; on the one hand, it recalls and uses structured visual information and, so, its formulation depends on a PMI_{st}, on the other hand, it is influenced by cortical and non-cortical neuronal configurations[23]. Thus, an NPMI is a new configuration and it always refers to *a world already perceived*. On the contrary, imagination images are mental contents whose generation has no semantic reference to the world, as it is neither perceived nor actually present at the moment of their formulation. Since our main aim here is to make clear the relationship between PMI_{st} and NPMI, we will not take into account imagination images in detail, though they are an interesting and fascinating end product of the visual mental activity, and to address the issue of this Section we shall start from the following statements:

1) there is not any NPMI without a previous PMI_{st};
2) an NPMI cannot be reduced to a PMI_{st};
3) NPMI generation, besides PMI_{st}, involves different mental contents;
4) NPMI generation *is always a new neuromental process which gives rise to a new mental configuration*.

The relevance of these statements will be clearer in what follows and we will see that, on the one hand, they are coherent with MINLF, while on the other hand, they refute some of the hypotheses contained in MISF.

To say that an NPMI generation involves a PMI_{st} as well as other mental contents means, above all, to specify, on the one hand, that its formulation is reducible neither to what has been processed and stored in LTM nor to the conceptual or emotional capacities of the subject; on the other hand, that imagery is always the result of a process in which visual information is *retrieved* according to different conditions and modalities. So, we might even say that if there were no PMI_{st}, it would be impossible for a subject to recall past visual experiences and, at the same time, if there were not also other mental contents, which operate as top-down processes according to the global condition of the mind, such recalling would be impossible as well. This global condition includes different variables such as: beliefs, thoughts, emotions, demands, significations, expectations and so forth and, as we can note,

> It has recently been suggested that, during image generation, top-down processes arise in part from prefrontal areas, while during perception bottom-up processes arise from early visual areas. Our results are in agreement with this outcome: the left inferior frontal gyrus during imagery was more affected because top-down processes originated from this region. (Mazard, A., Laou, L., Joliot, M., Mellet, E., 2005).

Although PMI_{st} play a very important role in NPMI generation, we do not believe, as some authors would say[24], that 'imagery generation is the process by which long-term memory knowledge of the visual appearance of objects or scenes is used to create a short-term percept-like image' (Farah, 1995); rather, we think that *there is much more than a mere 'recalling' of a* PMI_{st}, though in some cases it can be this way, i.e., in 'reactivation'. Thus, if an NPMI is PMI_{st}-dependent, but visual information embodied in it is not enough to account for its generation, then, the involvement of different mental contents plays a decisive role, as, for instance, the presence or not of specific goals to reach or expectations to fulfil determine the way in which an NPMI is formulated. As the involvement of different mental contents together with visual information can give rise to several NPMI we can state that:

> *given visual information embodied in a* PMI_{st}*, n NPMI can be derived which can have different relations with such information.*

So, if we accept that NPMI is not a mere recalling of visual information stored, then, we should uphold what we call the *principle of irreducibility* of

NPMI generation, according to which imagery is not reducible to a PMI_{st} and so it cannot be considered *strictu sensu* a percept-like image. On the basis of this principle, if imagery can be a partial recalling of a PMI_{st} and recover some information while neglecting some other, then, only so-called reactivation can be properly considered a percept-like image, i.e., imagery whose generation retains all the information arranged in a PMI_{st}. Returning to the example of the 'black cat', if I recall simply its shape or colour but not what it means to me, I have imagery which retains only a small part of the whole information processed and stored; on the contrary, if I recover all its figural features together with all my memories and significations assigned to the cat, I have imagery which is a complete reactivation of a PMI_{st}, in other words, an NPMI of a black cat with all its significations. Though reactivation is an important retrieval modality, it is not the only one. Thus, although an NPMI is always perceptive-based, it is not 'always' a mere reactivation nor does it entirely depends on what has been processed and stored.

It is necessary, now, to take into account two important questions, though both of them have been previously pointed out: a) imagery is always a kind of retrieval of the information embodied in a PMI_{st} and b) given the same PMI_{st}, *n* NPMI can be generated at different times. Both of them underline that imagery, even as reactivation, can in no manner be reduced to a process by which perceptive contents are recalled from long-term memory; though some experimental findings seem to advocate, somehow, the reactivation process showing specific neuronal activation when subjects are asked to perform a mental imagery task using verbal cues, in everyday life imagery generation involves variables and mental processes that are more far complex than ones managed in experimental conditions. So, if scientific results describe what happens to the mind when imagining 'something' under specific experimental conditions, they cannot put forward, at the same time, a suitable model to understand the ordinary generation of NPMI. Actually, an imagery generation process does not work as informational computer processing and a comparison with the artificial computing systems could be misleading. Thus, what we mean by *retrieval* is not what MISF suggests (i.e., the process by which long-term memory knowledge of the visual appearance of objects or scenes is used to create a short-term percept-like image) but *what contributes to an NPMI generation together with further different mental processes which, in turn, are linked to the global condition of the mind*. As we have already said, we do not believe that every NPMI is a reactivation of visual information embodied in a PMI_{st} or rather, it is not always like this, i.e., not always does NPMI recall the complete information of a PMI_{st}. Hence, we will say that imagery generation is like a process of *reconstructing* and *deconstructing* of visual information and, as there are different kinds of retrieval modalities of a PMI_{st}, besides reactivation (see Section 4.1.2. and 4.1.3), there can be also several NPMI, different from one another referring to the same perceptive image.

In the following sections we will analyze in detail the process by which visual information is retrieved from a PMI_{st} to generate imagery. Indeed, it is necessary to make clear that such a process holds not only between NPMI and PMI_{st} but also between NPMI and a *previous* NPMI; more specifically, this means that imagery can retrieve visual information both from a PMI_{st} and from other imagery which 'stands for' the original perceptive image as its *substitution* and, in such a way, it is being maintained in LTM. Thus, we argue that each NPMI might substitute, totally or partially, a PMI_{st} or a previous NPMI in LTM, being, this way, the *new* mental content from which *new* imagery can be generated. Actually, in everyday life it often happens that what we recover is not the 'very *first'* PMI_{st} but visual information arranged into one of its substitutions which, in turn, though they seem to be vivid, are not at all the very first one. Therefore, *the whole retrieval process is 'iterative' and holds in both cases, i.e., when accounting for NPMI generation from a PMI_{st} and NPMI generation from previously stored NPMI (from now on we will call it $NPMI_{st}$)*[25]. So, each new $NPMI_{st}$ at time t_i retrieves visual information from PMI_{st} or $NPMI_{st}$ at time t_{i-1}, i.e., the last $NPMI_{st}$ in LTM.

Figure 3 describes both processes.

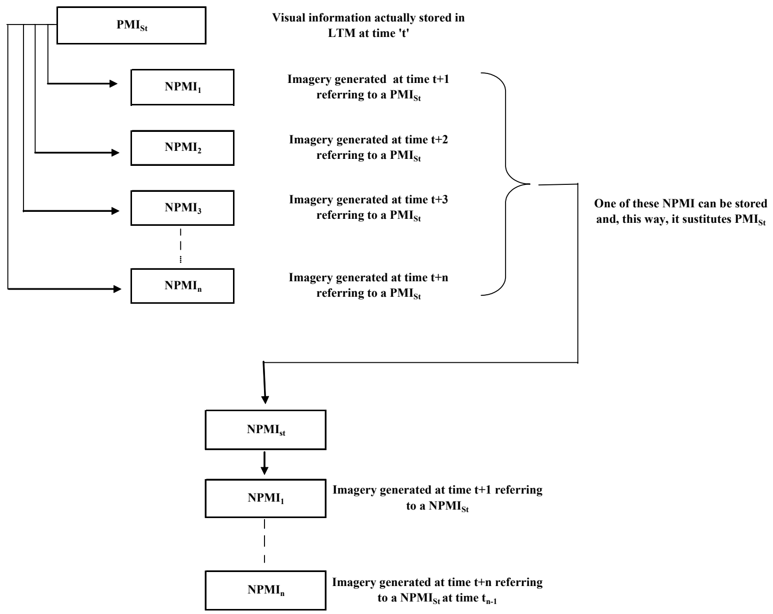

Figure 3. NPMI Generation Process and Substitution Generation Process.

4.1. Retrieval

What is a 'retrieval'? How does it really work? Is it a chance pulling out of perceptive images? As we can notice to address this issue means to run into many crucial questions and one of these properly concerns visual information: does the information structured into a PMI_{st} (or $NPMI_{st}$) and used to formulate *new* NPMI remain the same? In other words, is it retrieved 'as it has been processed and stored' or is it, somehow, modified during the imagery generation process? If so, in what ways does it happen? Indeed, we have already answered, albeit partially, this tricky question, but on this occasion we want to stress, one more, that we believe that visual information embodied in a PMI_{st} can be modified *within the imagery generation process*, and strategic conditions, retainment and retrieval modalities are the ways in which this process happens. Such a process can recall some information rather than other which, this way, can be definitely or not abandoned. Nevertheless, it is necessary to say that *this modification process does not entail a direct action* on a PMI_{st}; rather, it happens within an NPMI generation.

Some authors would say that the imagery process, while working, refers to visual information structured and saved in LTM, conceived as a 'store' where perceptive images are and remain just the way they have been processed (see point b) of MISF). They would also agree that LTM can be subject to modifications caused by brain damage, neuronal decay or specific mental contents like emotions and feelings but, most likely, they would not accept modifications caused by the imagery generation process. On the contrary, we believe that this is possible, in other words, we say that visual information embodied in LTM can be subject to modifications caused by 'retrieval', conceived as an *active process* able to alter visual information of a PMI_{st} (or $NPMI_{st}$) during NPMI generation. Hence, we do not accept point c) of MISF. In this paper we do not take into account questions concerning pathology or neuronal decay, though we believe that they could be useful to grasp the meaning of the relationship between PMI_{st} and NPMI.

In accordance with what has been said so far, we will refer the term 'retrieval' to the whole process of recalling of a PMI_{st} (or $NPMI_{st}$) to generate *new* NPMI. In this way, the retrieval process refutes hypothesis d) of MISF. Such a process can be distinguished in the following ways: a) *strategic conditions*: the general mental conditions by which the retrieval process is triggered; b) *retainment*: '***what***' NPMI generation actually recalls, i.e., the visual information really maintained; c) *retrieval modalities*: '***how***' visual information is recalled; in Section 4.1.3. we will make clear how many modalities there are and how they really operate to give rise to different imagery of the same PMI_{st}.

4.1.1. Strategic conditions

As we have already mentioned, the retrieval process is, in most cases, triggered on the basis of general mental conditions in which there may or may not be motivations, expectations and propositional attitudes. When one, or some of these, are actually involved NPMI generation process is *goal-oriented* (one of this case is the experimental one in which a subject is asked to accomplish a specific visual task); on the contrary, such a process is *non-goal-oriented*, i.e., free, in other words, it is not constrained to recover specific visual information to get a target. To give an example, we might think about what happens while recalling the physical features of somebody, for instance a girl, in two different cases: 1) when we just want to e-mail her or 2) if we would like to buy her a dress. While in the latter case (2), to decide the most suitable clothes or what fits and suits her best, we necessarily have to recall her height, size, eye colour or complexion; in the former (1), information can be retrieved from PMI_{st} in a free way, as we do not have to reach a specific target. So, while *goal-oriented* imagery is aimed at serving a specific purpose, *non-goal-oriented* imagery does not have to pursue anything, as there are no prior intentions or motivations.

It is now necessary to stress that strategic conditions 'drive' the whole retrieval process, orienting the way in which retainment and modalities will operate to generate specific NPMI.

4.1.2. Retainment

Imagery generation does not necessarily have to recall all the visual information arranged in a PMI_{st}, (or $NPMI_{st}$); for instance, if I want to taste again a 'Spaghetti dish' cooked in a seaside restaurant I know very well, I do not need to retain the complete perceptive image of that restaurant but simply the detailed image of the 'Spaghetti dish'. Therefore, we can say that an NPMI can retain, at different times, different contents of the 'same' PMI_{st} (or $NPMI_{st}$), i.e., a) figural features and semantic reference; b) significations and values.

a) *Figural features and semantic reference retainment*

Such retainment happens when the imagery recalls, though not completely, the figural features and semantic reference of a perceptive image, for instance, the shape, colour, contour and size of a 'cat', but does not at all concern what the cat means to me from a cognitive or emotional point of view. Thus, imagery which retains the figural features and semantic reference of a PMI_{st} is an NPMI which recovers only some of its contents.

b) *Significations and values retainment*

Such retainment happens when imagery recovers, partially or completely, the *meaningful space* of a PMI_{st}. When recalling a 'black cat', for instance, I can retrieve only some of the main significations assigned to it, while ignoring some others; therefore, if I retain the attribute of being 'a sign of bad luck' instead of that of being 'a nice pet' or 'an animal I would like to buy', I recall only partially the meaningful space assigned to it. During an NPMI generation not only can some significations be retained but new significations and values can be added to visual information recalled. Thus, the retainment process concerns what is actually retrieved.

4.1.3. Retrieval modalities

As we have already mentioned, the retrieval process concerns '*what*' is actually recalled (retainment) and '*how*' information is recovered (modalities) under specific strategic conditions, which operate as a 'guide' to the whole process. If a strategic condition, for instance a motivation, demands the recalling of specific contents arranged in a PMI_{st}, i.e., detailed visual information, then, such a process has to know '*what*' and '*how*' to recover it. In this way, retainment and retrieval modalities apply, at the same time, to the visual information required. If the motivation is to go back to a nice Italian restaurant to eat a specific Italian dish I have already tasted, then, as far as 'what' is concerned, there is no need to recover all the visual information of that restaurant, but, simply, only one related to the Italian dish; as regards the 'modalities', most likely I will use that of 'focus', i.e., a selective attention which involves focusing on specific information of a scene while ignoring other aspects. Therefore, I will focus on the figural features of 'Spaghetti dish' perceptive image.

We will, now attempt to underline some of these modalities.

a) *Reactivation*

We can define 'reactivation process' as a 'complete retrieval' of visual information arranged into a PMI_{st} (or $NPMI_{st}$) which has been processed at the three levels (visual format, visual pattern and visual frame; see Section 3); thus, only in this case, can an NPMI be said to be an *epistemic isomorphic structure* of a PMI_{st} (for the notion of isomorphism see Section 5) and, so to speak, it can be considered a *ready-made* image. Indeed, this is what the majority of authors mean by the term 'imagery'; on the contrary, we believe that reactivation is only one of the different modalities to retrieve visual information and that it is certainly not the most frequent one; nevertheless, we can say that this process is very common in the ordinary life to recognize what has been already perceived and what is being perceived right now. The isomorphism we have mentioned, in this case, is called *epistemic*, as it is based on a belief: the subject believes that an NPMI of an object has the same content of the perceptive image of this object when it was generated the first time.

b) *Resetting*

The 'resetting process' is *the retrieval of visual information derived from different visual experiences of the same object*, i.e., different PMI_{st} of the same thing, to generate a specific NPMI; to give rise, for instance, to a single image of the face of a person whom we have talked to several times or seen in different conditions we need to recall what he looks like while smiling, crying, sleeping, speaking, eating and so forth. Thus, all visual information we have processed and stored referring to him can be *reset* to generate a *single new* imagery.

c) *Mixing*

The 'mixing process' is *the retrieval of visual information derived from several visual experiences of different objects* to generate a specific NPMI. The result of such a process will be an imagery which *assembles* (puts together) figural contents different from one another, and, so, it is *new* as it does not refer to any of those visual experiences in particular.

d) *Focusing*

The 'focusing process' is *a selective attention to specific visual information while ignoring other*; we usually employ this modality when we are asked to analyse something or pay attention to particular details, for instance, to describe someone's eye. By focusing, a new imagery retrieves only what is relevant to its generation.

e) *Placement*

The placement process is *the retrieval of visual information referred to an object in a specific environmental context.* The face of a best friend, for instance, can be recalled, alone, in 'itself', or in a casino while he is playing black jack. Therefore, this modality gives the chance both to retrieve a perceptive image within its specific environmental context as well as in a different one.

f) *Flowing*

This is *the modality by which PMI_{st} are recalled 'in quick succession'*, i.e., following each other in time or order, to give rise to images linked to one another in a way similar to a movie.

The common features of these retrieval modalities are the following ones:

1) they give rise to imagery which is a *partial image* of a PMI_{st} (or $NPMI_{st}$), or rather, a partial recalling of the whole information arranged in it;

2) all imagery can add *new* information to the visual information retrieved;
3) visual information not retrieved can be definitively abandoned or not and, in the latter case, it becomes *residual*, i.e., it might be recalled in further imagery generation.

4.1.4. Configurational setting

As we have noticed so far, imagery is not always a mere recalling of a PMI_{st} (or $NPMI_{st}$), rather, it is the result of a non-linear retrieval process which gives rise to different new NPMI referred to the same visual information, though formulated at different times and in different conditions. So, an NPMI can be considered a modification of the visual information retained and, in turn, it might be distinguished according to 'time' and 'quality'. As far as 'time' is concerned, a modification and, so, an NPMI, can be temporary or permanent, i.e., it can be maintained for a short or a long period of time; as regards 'quality', a modification can *remove* even definitively some information as well as *replace* it or *add* new information. Thus, visual information not used to generate 'current imagery' can be either abandoned or maintained in memory and, this way, it is available for further strategic conditions and retrievals. We call such information *residual contents of NPMI generation*.

Hence, the whole imagery generation process based on retrieval ends up in what we call *configurational setting* of visual information stored in LTM. Therefore, if every NPMI is always a new neuromental configuration, then, there can be neither an NPMI similar to a PMI_{st} nor two NPMI similar to one another, though referring to the same visual information. We thus have what we can call the *indiscernible principle of neuromental configurations*, according to which there can never be two identical neuromental configurations.

We sum up this process in Figure 4 and 5.

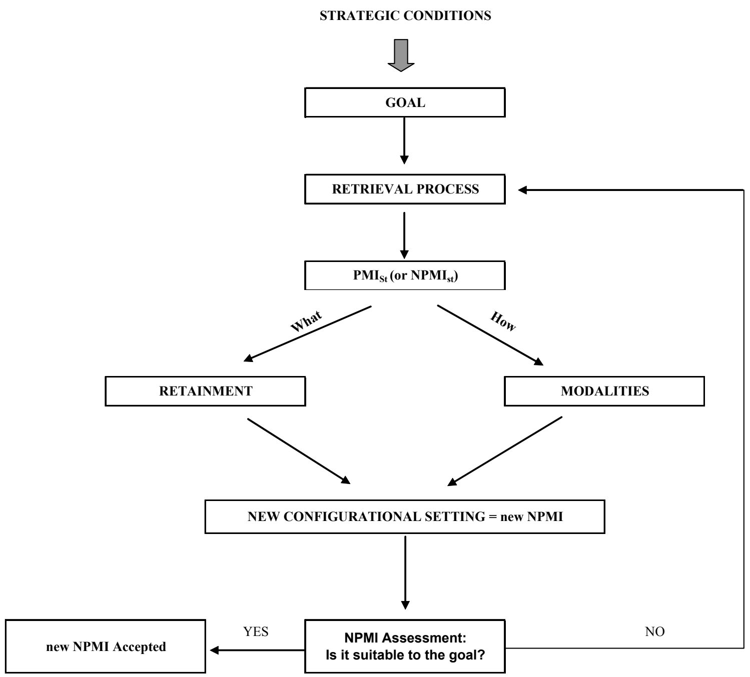

Figure 4. Goal-oriented retrieval process.

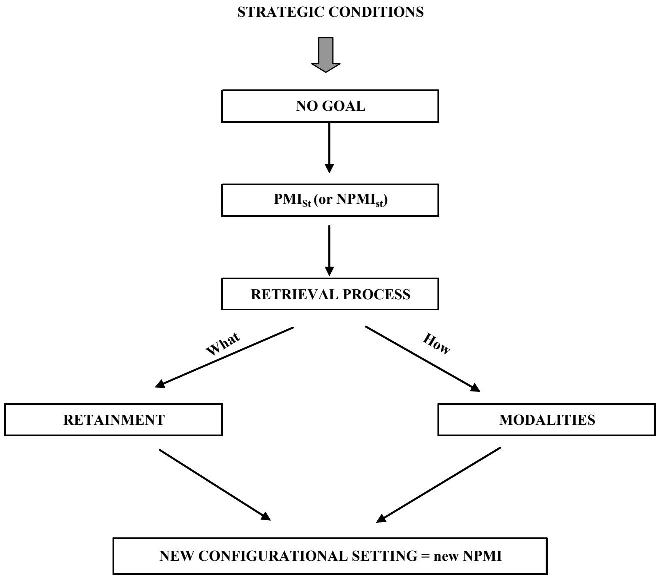

Figure 5. Non-goal-oriented retrieval process.

4.2. NPMI main features

On the basis of what has been said so far, we can underline the following main features of an NPMI :

1) an NPMI is always a new configurational setting of a PMI$_{st}$ or NPMI$_{st}$;
2) an NPMI can be a partial or total retrieval of the visual information embodied in a PMI$_{st}$ or NPMI$_{st}$ and can entail a partial or total modification of information which has been processed in three levels: visual format, visual pattern, visual frame;
3) an NPMI is not reducible to a PMI$_{st}$ or to an NPMI$_{st}$ i.e., it does not recall all visual information embodied in these;
4) an NPMI generation can also retrieve what we have called residual contents;
5) an NPMI brings with it, modified or not, the meaningful space of a PMI$_{st}$ or NPMI$_{st}$;

6) an NPMI is a substitution of a PMI_{st} or $NPMI_{st}$ and *it is believed to semantically refers to a physical object*, and, therefore, it stands for it;
7) an NPMI is a specific state of awareness;
8) an NPMI can be codified into a propositional code which describes its contents or assigns to it a proper name; in this way, such a non-perceptive image gains a propositional format;
9) an NPMI can be modified consciously or not and used both for further cognitive mental processes (such as: concept formation, descriptions, decision making and so forth) and to act in the world;
10) an NPMI might be considered the starting point of imagination-images generation.

5. PMI_{st}, NPMI and physical objects

The distinctions we have drawn and the account we have given of mental image generation urge us to specify in detail the structural relationship between PMI_{st} and NPMI and, consequently, to address the remarkable philosophical issue of their correlation with the physical object they refer to. To deal with these two crucial topics we will adopt the notions of correlation, isomorphism and ontology.

5.1. Isomorphism and correlation

PMI has been defined as 'an iconic (or figural) mental representation of a physical object', i.e., a neuromental content which *stands for* what it refers to (see Section 3.1.c); at this point we aim at clarifying in brief the meaning of the expression *stand for* and, for this purpose, we will use the term *isomorphism* in its algebraic sense: *an isomorphism between two sets A and B is a 1-1 correspondence which 'maps' one element of A in only one element in B so that every relation or operation defined in A is preserved by the correspondence in B*. In our specific case the two sets involved are respectively a physical object (A), or rather, physical stimuli coming from such an object, and PMI_{st} (B). Thus, applying to them the notion of isomorphism, we obtain that in a PMI_{st}[26] what are encoded are both the main features of a physical object and the relations among its components and, this way, isomorphism 'goes' from physical object to PMI and thence to PMI_{st}. In this sense, the structure of a physical object is 'mapped' with a specific code within the PMI_{st}. This isomorphism is based on the change of state of retina transducers together with the spreading of the visual information over the cortical and non-cortical brain areas, where such information is further processed: consequently, it is clear that this kind of isomorphism establishes a semantic relationship between PMI_{st} and the physical object. Thus, we can use the expression *stand for* to mean that a PMI_{st} is an *isomorphic structure of the physical object* and, so to speak, that it semantically refers to it. On the basis of these considerations, we accept a *realistic perspective* on the relationship between PMI_{st} and physical objects, which, as we will see later, does not hold between NPMI and physical objects.

As an NPMI can be generated in different ways at different times, under specific strategic conditions and retainment, through several retrieval modalities and under the influence of various mental contents (so that information embodied in a PMI_{st} can be differently arranged) its formulation cannot always be considered either a mere reactivation of the visual information processed and stored in LTM or a simple recalling of PMI_{st} figural features. Therefore, given a PMI_{st}, not only can we not predict with sufficient and adequate certainty which kind of new NPMI will be generated, but also the way in which it will be produced. In fact, imagery generation is always subject to a *degree of indeterminacy* and, given that it involves different variables and influences, like prior knowledge, residual data, meaningful space, motivational goals, retrieval modalities and other influences from cortical areas, its generation is a probabilistic process. In other words, given a PMI_{st} we can only theoretically compute the conditional probability of the occurrence of a new NPMI; the probability is conditioned by the general process of retrieval (see Section 4.1) and by all the other mental contents involved in its generation. On the basis of what has been said up to now, the whole process is a non-linear one.

This adds up to saying that an NPMI is never reducible to a PMI_{st}, and, though it is based on a visual perception, the involvement of different mental contents and processes determine the *non-linear way* in which the information processed is retrieved. Consequently, if there can be imagery generated through the reactivation modality, then, there can also be NPMI whose structure is different from the perceptive one, i.e., NPMI which is a partial or total modification of a PMI_{st}. For all these reasons, to refer to the relationship between a PMI_{st} (or $NPMI_{st}$) and every *new* NPMI, instead of isomorphism, we will use the term *correlation*, as it allows us to consider and use an imagery as if it were the original PMI_{st}. It is worth briefly mentioning that the only case in which there is an isomorphism between an NPMI and a PMI_{st} seems to be the reactivation one, as imagery fully recalls the contents of a previous perceptive image. Thus, if a subject through an NPMI retrieves what is contained in a PMI_{st}, then, she will also *believe* that such imagery refers to a PMI_{st}, and, consequently, to the physical object as well. As this correlation is based on a 'belief', we will call it *epistemic correlation*; through it, actually, a subject assumes that her new NPMI refers to a mental content already processed and stored in LTM and, in this way, to the physical world as well. Now that we have seen what 'isomorphism' and 'correlation' really mean, it is possible to deal briefly with some ontological aspects. In other words, we will consider the relationship between PMI_{st} and NPMI, on the one hand, and, on the other hand, the physical object they refer to.

5.2. Weak and Strong Ontology

PMI_{st} and NPMI, though different from one another, are useful in enabling us to act and to know the physical environment and, since all human beings use them as *models* of the world they live in, they give rise to *two different ontologi-*

cal relationships between mind and world. Of these relationships, the first, which we will call *strong ontology*, is based on the fundamental isomorphism, mentioned above, between a PMI_{st} and the structure of the physical object (where stimuli come from) it refers to; the second one, which we will call *weak ontology*, concerns the relationship between an NPMI and the structure of the physical object it *can* refer to, as it is generated without actually perceived stimuli.

We consider the first relationship *strong* because it is based on the isomorphism empirically controllable between PMI and physical objects, and *ontological* as there are three kinds of ontological statements referring to a PMI_{st}, a physical object and their relationship. By the first ontological statement, the subject ascertains that there is a PMI_{st} she is aware of; by the second ontological statement, she ascertains, on the one hand, that something has triggered the generation of such PMI_{st}, on the other hand, that there is a relationship between such PMI_{st} and a physical object, assuming, at the same time, that this relationship is a structural correspondence, i.e., an isomorphism. By the third ontological statement, the subject assigns '*existence*' to that physical object.

While a PMI_{st} is directly triggered by physical stimuli, an NPMI refers to the world in a very different way and, since it is activated without something being actually perceived, its *contents* can only refer to visual information that has been really processed and stored in LTM. As an NPMI, though we may be aware of it, is blind to physical stimuli, during its generation we cannot perceive at all what it semantically refers to. Notice that its reference to the world is only *indirect*, as it is perceptive image-dependent, it is clear that the relationship between an NPMI and a physical object is *ontologically weak* and, above all, is based on a *belief*. Such a belief, however, is fundamental, as through it, imagery is believed to refer to a physical object and, consequently, it is used to act and live in the world as if it were a PMI_{st}. In other words, recalling information embodied in a PMI_{st}, imagery is believed to retrieve also its semantic reference to the physical object and, for the same reason, it is believed to be able to stand for it. As the relationship between NPMI and world is not isomorphic, we prefer to use the expression *epistemic ontological correlation* to describe the way in which NPMI refers to physical objects. In any case, though this last perspective seems to be less realistic than one formulated for the relationship between PMI and physical objects (see Section 5.1), it is worth underlining that, although NPMI are based on a belief, they can be used *as if they were* PMI_{st} and, so, referring to the world; of course, it is a sort of *fiction*, but, in a certain way, this is what happens in most cases for human beings very often approach the world they live in *indirectly*, i.e., using the mental contents they have formulated about it as *model*s; indeed, this is the most relevant role an NPMI plays a) to react to and recognize new physical stimuli; b) to know the world through the formulation of different figural models and c) to behave in it in accordance with such models.

Closing remarks

In this paper we have tried to adopt a point of view according to which in order to analyse the nature of non-perceptive mental images and their generation process it is necessary to take into account the structure of perceptive mental images in detail. How can we really understand the generation of imagery if we do not consider the structure of perceptive images? Indeed, as we have shown in this paper, it is precisely the structure of a PMI that allows us to grasp the way in which an NPMI is generated and its relation to visual information previously processed and stored in LTM. According to this perspective, we have analysed in depth the structure of PMI and have considered different levels of visual information processing within the whole process of PMI generation; we have then introduced a detailed analysis of this process and used notions like format, pattern and frame to mean the intermediate levels of visual perception.

We have also noted that only this analysis gives us the opportunity to understand how visual information is recalled to generate NPMI. In our view, such generation is highly complex and is based on the general notion of retrieval, which has been distinguished in three different sub-routines: strategic conditions, retainment and retrieval modalities. As we have seen, this retrieval process is intimately related to the structure of PMI, and, thus, the outline of this complex generation process closely related to the structure of PMI has led us to differentiate our MINLF from MISF and refute some of its assumptions, in particular, the well known notions of reactivation and percept-like image, according to which a PMI is a recall of information arranged in a visual perception. Indeed, we have shown that it is not always so, and that the NPMI generation process does not always refer to the original PMI but also to NPMI which is its substitution in LTM. Furthermore, the main feature of this generation is that of being an active process, not a mere reproduction, and it is non-linear, i.e., from stored visual information different non-perceptive mental images can be derived and their formulation is not completely predictable. To put it simply, to have, now, in our mind visual information does not allow us to say which imagery will be generated, even after some minutes; this concerns not only the figural features of visual information but also its significations and values, which, as we have noticed, are assigned to it. Therefore, to retrieve visual information means not only to recover its figural features but, as we have argued, its significations and values as well. Moreover, figural features and significations are not always recalled 'so as they are' in LTM but, most likely, every time they are retrieved they can be, somehow, modified, at least, according to two variables:

1. goals (motivations) which trigger the retrieval of visual information (that is, what we have called strategic conditions, retainment and modalities);
2. prior knowledge and other mental contents, as we have mentioned.

This amounts to saying that *every new imagery is a configurational setting of stored visual information*. We can also highlight the fact that NPMI and PMI establish different ontological relations with physical objects; thus, we have introduced the notion of isomorphism between PMI and physical objects and the distinction between strong and weak ontology.

From the point of view of the Philosophy of Mind, we could put forward some tricky questions: a) is NPMI a mental content or a mental process which, once activated, gives rise to new different mental processes? b) does an NPMI generation concern only visual information processing or does it involve also other different mental activities?

Indeed, we believe that further investigation is called for, that, so to speak, '*much more water must flow under the bridge of the mind*' before we may understand what mental imagery actually means.

Endnotes

[1] PMI and NPMI stand for perceptive mental image and non-perceptive mental image.
[2] It is useful to note, once more, that some of these hypotheses, usually correlated to the pictorial theory of mental images, are neither fully shared nor all accepted by the majority of authors.
[3] Perceptive visual data means PMI, i.e., visual information processed and stored together with the attributes, meanings and values assigned to it.
[4] "Intons-Peterson and McDaniel (1991) suggested that mental images might be the product of an interaction between a visual representation and an individual's conceptual knowledge of an object, suggesting therefore that mental images might be "knowledge based". In: Gardini, S., De Beni, R., Cornoldi, C., Bromiley, A., Venneri A. (2005). Different neuronal pathways support the generation of general and specific mental images. *NeuroImage*, 27, 544-552.
[5] Structure means the 'way in which visual information is processed and organized'.
[6] "Furthermore, the content of mental images evoked from long-term memory may be less modality specific and produce more types of representation than the corresponding sensory representations directly derived from perception (Cornoldi et al., 1998)". In: Gardini, S., De Beni, R., Cornoldi, C., Bromiley, A., Venneri A. (2005). Different neuronal pathways support the generation of general and specific mental images. *NeuroImage*, 27, 544-552.
[7] PMI_{st} means perceptive visual information already processed and stored in LTM (See Section 3).
[8] Further readings in: Kosslyn, S.M., Thompson, W.L., Alpert, N.M. (1997) Neural System Shared by Visual Imagery and Visual Perception: A Positron Emission Tomography Study. *NeuroImage*, 6, 320-334.
[9] Visual perceptions occur when a physical stimulus is actually perceived; mental imagery are stimulus-independent but semantically referred to the world; imagination images, instead, are visual mental contents without any correspondence to the world 'actually' perceived at the moment of their generation.
[10] Imagination is here also called a 'blind but indispensable function of the soul, without which we should have no knowledge whatsoever, but of which we are scarcely ever conscious'. More precisely Kant distinguishes between reproduc-

tive and productive imagination. Reproductive imagination supplements the fragmentary sense inputs, so that when we see a round green apple, we assume that it has some continuous features of location and quality, and that its identity persists in a continuous temporal series. Whereas the reproductive or associative imagination (empirische Einbildungskraft) is operative in the apprehension of particular empirical phenomena, the productive imagination (die produktive Synthesis der Einbildungskraft) provides a global orientation by placing any perception in the a priori context of one single, unified experience of all possible states of consciousness. Thus, without the productive imagination the reproductive imagination could not work at all (KrV, A 100–103 and A 115–119).

[11] For a biological approach to mental representations and perceptual experience see Bianca, M. (2000). *Fondamenti di etica e bioetica*. Bologna, Patron; Damasio, A. (1995). *L'errore di Cartesio. Emozione, ragione e cervello umano*. Milano, Adelphi; Gazzaniga, M. (2006). *La mente etica*. Torino, Codice Edizioni.

[12] "Some writers (e.g., Churchland, 1988) have argued that the presence of outgoing (centripetal) nerve fibers running from higher cortical centers to the visual cortex constitutes prima facie evidence that vision must be susceptible to cognitive influences" (Pylyshyn, 2003). Further readings in: Han, S., Jiang, G., Zhou, T., Cai, P. (2005). Distinct neural substrates for perception of real and vistual visual world. *NeuroImage*, 24, 928-935; and in: Simmons, A., Miller, D., Feinstein, J., Goldberg, T., Paulus, M.P. (2005). Left inferior prefrontal cortex activation during decision-making task predicts the degree of semantic organization. *NeuroImage*, 28, 30-38, suggest that many brain regions may be involved in discriminating different categories of objects. The inferior pre-frontal cortex, for instance, has been identified as a vital structure for processing competing semantic information when making a decision about the meaning of a stimulus.

[13] A process is neuromental if it involves different cortical brain areas.

[14] Photoreceptors are responsible for transducing light (photons) into nerve signals that can be transmitted to the brain via the optic nerve.

[15] To be aware of something means to be in a neuromental configuration by which a subject a) ascertains to be in this configuration and b) can formulate, albeit not compulsory, propositions which claim that she is in such a configuration; these propositions render explicit a configurational content into a linguistic code. These propositions can also trigger consequent behaviours. So, a neuromental configuration is aware when conditions a) and b) are satisfied.

[16] "In monkeys, over 30 visual areas have been identified and their connectivity and functional properties were studied (…) They exhibit a characteristic anatomical and functional organization, in which the majority of the retinal visual information enters the brain in the occipital lobe, passes through a number of retinotopically organized areas and segregates into two distinguishable but interconnected visual information streams extending along the dorsal and the ventral side of the brain, respectively. Although much less is known about the human visual system, functional imaging studies in the last decade revealed a roughly similar organization in the human brain. (…) Subsequent studies demonstrated the involvement of several cortical areas along the ventral occipital-temporal

brain in various aspects of visual form analysis and colour constancy (e.g., Bartels and Zeki, 2000; Kanwisher et al., 1996; Malech et al., 1995)". In: Stiers, P., Peeters, R., Lagae, L., Hecke, P.V., Sunaert, S. (2006). Mapping multiple visual areas in the human brain with a short fMRI sequence. *NeuroImage*, 29, 74-89.

[17] In this paper we will use the term information to refer to data flowing throughout a channel from one structure to another; these data, according to Floridi, are data + meaning. Nevertheless, we will not consider the 'true' nature of information within the C.N.S. and for this topic we refer to: Floridi, L. (2003). *Outline of a Theory of Strongly Semantic Information*, (http://www.wolfson.ox.ac.uk/-floridi/papers.htm); Maleeh, R., Shafiee, A. (2006). Two Aspects of Information-Driven Interactions: Towards a Theory of Consciousness, submitted to *Anthropology and Philosophy*; Roederer, J.G. (2005). *Information and its Role in Nature*, Berlin Heidelberg, Springer-Verlag.

[18] We can consider the photoreceptors encoding physical information in neurochemical code the starting point of the whole process, thus, the proper kernel of visual information flow and processing.

[19] The branch-over is the process by which integrated visual information (visual pattern) is spread over different brain regions to be further processed.

[20] This term is certainly vague but in this paper we do not take it into account explaining properly what it actually means.

[21] For an analysis of the use of neuromental configurations to know and to act in the world see Bianca, M. (2005). *Rappresentazioni mentali e conoscenza*. Milano, Franco Angeli.

[22] For this notion, *ibid*., p. 47.

[23] The term configuration refers to the contemporary activation in a time t' of n neuronal structures.

[24] "Mental image generation has been defined as the process by which long-term memory knowledge of the visual appearance of objects or scenes is used to create a short-term percept-like image (Farah, 1995). Imagery would therefore consist of the temporary re-activation of representational units retrieved from long-term memory into a specialized processor, or "visual buffer" (Kosslyn, 1994)". In: Gardini, S., De Beni, R., Cornoldi, C., Bromiley, A., Venneri A. (2005). Different neuronal pathways support the generation of general and specific mental images. *NeuroImage*, 27, 544-552.

[25] In the following sections, when we talk about 'visual information retrieval' we will not specify each time that what holds for a PMI_{st}, can be also applied to a $NPMI_{st}$; as we have noted, the imagery generation process can start from both of them.

[26] Strictu sensu 'isomorphism' is established between PMI and the physical object but, since visual information embodied in it can be immediately stored, even not completely, we can state that 'isomorphism' also holds between PMI_{st} and physical object.

References

Essays

Agostino, A. (c1995). *De Vera Religione*. Roma, Città Nuova.

Aristotele (1991). *Opere*. Roma-Bari, Laterza.

Bear, M.F., Connors Barry, W., Paradiso, M.A. (1999). *Neuroscienze. Esplorando il cervello*. Milano, Masson.

Berkeley, G. (1973). *Trattato sui principi della conoscenza umana*. Bari, Laterza.

Bermudez, J., L. (1998). *The Paradox of Self-Consciousness*. Cambridge (MA), MIT Press.

Bianca, M. (2005). *Rappresentazioni mentali e conoscenza*. Milano, Franco Angeli.

Denes, G., Pizzamiglio, L. (1996). (Edited by) *Manuale di Neuropsicologia: normalità e patologia dei processi cognitivi*. Bologna, Zanichelli.

Denis, M., Mellet, E., Kosslyn, S. M. (Eds). (2004). *Neuroimaging of Mental Imagery*. Psychology Press, Hove, UK.

Descartes, R. (1969). *Meditazioni metafisiche*. In Opere Filosofiche, Torino, Utet.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Oxford, Blackwell.

Ferretti, F. (1998). *Pensare vedendo. Le immagini mentali nella scienza cognitiva*. Roma, Carocci.

Floridi, L. (2003). *Outline of a Theory of Strongly Semantic Information*. (http://www.wolfson.ox.ac.uk/floridi/papers.htm.

Formigari, L., Casertano, G., Cubeddu, I. (c1999). *Imago in phantasia depicta: Studi sulla teoria dell'immaginazione*. Roma, Carocci.

Gendler, T., S., Hawthorne, J. (2006). (Edited by), *Perceptual Experience*. New York, Oxford University Press.

Gibson, J.J. (c1999). *Un approccio ecologico alla percezione visiva*. Bologna, il Mulino.

Ghiron, V. (c2001). *La teoria dell'Immaginazione di Edmund Husserl: fantasia e coscienza figurale nella fenomenologia descrittiva*. Venezia, Marsilio.

Hobbes, T. (1972). *Il Corpo* (*De Corpore*). In Elementi di Filosofia. Torino, Utet.

Hume, D. (1888). *Treatise of the Human Nature*. Oxford Press.

Kant, I., (1956). *Kritik der reinen Vernunft*. Frankfurt am Main, Suhrkamp.

Kosslyn, S.M., Thompson, W.L., Ganis, G. (2006) *The case for Mental Imagery*. New York, Oxford University Press.

Langton, R.G. (1998). *Occhio e cervello: la psicologia del vedere*, Milano, Cortina.

Locke, J. (2004). *Saggio sull'Intelletto Umano*. Milano, Bompiani.

Pylyshyn, Z.W. (2003). *Seeing and Visualizing: It's not what you think*. Cambridge, MA: MIT Press.

Roederer, J.G. (2005). *Information and its Role in Nature*. Berlin Heidelberg, Springer-Verlag.

Sartre, J.P. (1948). *Immagine e Coscienza*. Torino, Einaudi.

Sartre, J.P. (1962). *L'immaginazione – Idee per una teoria delle emozioni*. Milano, Bompiani.

Sartre, J.P. (1940). *L'imaginaire. Psychologie phénoménologique de l'imagination*. Éditions Gallimard.

Wittgenstein, L. (1999). *Ricerche filosofiche*. Torino, Einaudi.

Articles

Allen, G., McColl, R., Barnard, H., Ringe, V.K., Fleckenstein, J., Cullum, C.M. (2005). Magnetic a imaging of cerebellar-prefrontal and cerebellar-parietal functional connectivity. *NeuroImage*, 28, 38-48.

Altmann, Erik M., Bonnie, E.J. (1999). Episodic Index: A model of memory for Attention Events. *Cognitive Science*, 23 (2), 117-156.

Boles, B.D. (1984). Global Versus Local Processing: is there a hemispheric dichotomy? *Neuropsychologia*, 22, 445-455.

Calvert, A. Gemma (2001). Crossmodal Processing in the Human Brain: Insights from Functional Neuroimaging Studies. *Cerebral Cortex*, 11 (12), 1110-1123.

Chabris, C.F., Kosslyn, S.M. (2005). Representational correspondence as a basic principle of diagram design. In Tergan, S-O., & Keller, T. (Eds.), *Knowledge and Information Visualization, LNCS 3426*. Berlin: Springer-Verlag: 36 - 57.

Das, P., Kemp, A.H., Liddell, B., Brown, K.J., Olivieri, G., Peduto, A., Gordon, E., Williams, L.M. (2005). Pathways for fear perception: modulation of amygdale activity by thalamo-cortical systems. *NeuroImage*, 26, 141-148.

De C. Hamilton, A., Wolpert, D.M., Frith, U., Grafton, ST. (2006). Where does your own action influence your perception of another person's action in the brain? *NeuroImage*, 29, 524-535.

Denenberg, V.H. (1981). Hemispheric laterality in animals and the effect of early experience. *Behavioural and Brain Sciences*, 4, 1-149.

Denis, M., Kosslyn, S.M. (1999). Does the window really need to be washed? More on the mental scanning paradigm. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 18*, 593-616.

Denis, M., Kosslyn, S. M. (1999). Scanning visual images: A window on the mind. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition,* 18, 409-465.

Denis, M., Mellet, E., Kosslyn, S. M. (2004). Neuroimaging of mental imagery: An introduction. *European Journal of Cognitive Psychology, 16*, 625-630.

Ewbank, M.P., Schluppeck, D., Andrews, T. (2005). fMRI-adaptation reveals a distributed representation of inanimate objects and places in human visual cortex. *NeuroImage*, 28, 268-279.

Farah M.J. (1988). Is Visual Imagery really Visual? Overlooked Evidence from Neuropsychology. *Psychological Review*, 95, 307-17.

Finke, R.A. (1980). Levels of Equivalence in Imagery and Perception. *Psychological Review*, 87, 113-132.

Finke, R.A., Pinker, S., Farah, M. (1989). Reinterpreting Visual Patterns in Mental Imagery. *Cognitive Sciences*, 16, 252-257.

Galaburda, A. M., Kosslyn, S. M., Christen, Y. (2002). Introduction. In A. M. Galaburda, S. M. Kosslyn, & Y. Christen (Eds.), *Languages of the Brain*. Cambridge, MA: Harvard University Press. 1-14.

Ganis, G., Thompson, W. L., Mast, F. W., Kosslyn, S. M. (2003). Visual imagery in cerebral visual dysfunction. *Neurologic Clinics of North America*, 21, 631-646.

Ganis, G., Thompson, W. L., Kosslyn, S. M. (2004). Brain areas underlying visual imagery and visual perception: an fMRI study. *Cognitive Brain Research*, 20, 226-241.

Ganis, G., Thompson, W. L., Kosslyn, S. M. (2005). Understanding the effects of task-specific practice in the brain: Insights from individual-differences analyses. *Cognitive, Affective, & Behavioral Neuroscience*, 5, 235-245.

Gardini, S., De Beni, R., Cornoldi, C., Bromiley, A., Venneri A. (2005). Different neuronal pathways support the generation of general and specific mental images. *NeuroImage*, 27, 544-552.

Goodale & Milner (1992). Separate pathways for perception and action. *Trends in Neuroscience*, 15, 20-25.

Gorno-Tempini, M.L., Price, C.J. (2001). Identification of famous faces and buildings. A functional neuroimaging study of semantically unique items. *Brain*, 124, 2087-2097.

Grèzes, J., Fonlupt, P., Bertenhal, B., Delon-Martin, C., Segebarth, C., Decety, J. (2001). Does Perception of Biological Motion Rgnitive brainely on Specific Brain Regions? *NeuroImage*, 13, 775-785,

Grill-Spector, K. & Malach, R. (2004). The Human Visual Cortex. *Annual Review of Neurosciences*, 27, 649-677.

Han, S., Jiang, G., Zhou, T., Cai, P. (2005). Distinct neural substrates for the perception of real and virtual visual worlds. *NeuroImage*, 24, 928-935.

Hanakawa, T., Honda, M., Sawamoto, N., Okada, T., Yonekura, Y., Fukuyama, H., Shibasaki, H. (2002). The Role of Rostral Broadmann Area 6 in Mental-operation Tasks: an integrative Neuroimaging Approach. *Cerebral Cortex*, 12 (11), 1157-1170.

Klein, I., Dubois, J., Mangin, J., Kherif, F., Flandin, G., Poline, J., Denis, M., Kosslyn, S., Bihan, D. (2004). Retinotopic organization of visual mental images as revealed by functional magnetic resonance imaging. *Cognitive Brain Research*, 22, 26-31.

Klein, I., Paradis, A.-L., Poline, J.-B., Kosslyn, S.M. LeBihan, D. (2000). Transient activity in human calcarine cortex during visual imagery. *Journal of Cognitive Neuroscience*, 12, 15S-23S.

Kosslyn, S.M., Thompson, W. L., Alpert, N. M. (1995). Identifying objects at different levels of hierarchy: A positron emission tomography study. *Human Brain Mapping,* 3, 107-132.

Kosslyn, S.M. (1996). Neural systems and psychiatric disorders. *Cognitive Neuropsychiatry,* 1*,* 89-93.

Kosslyn, S.M., Thompson, W.L., Alpert, N.M. (1997). Neural System Shared by Visual Imagery and Visual Perception: A Positron Emission Tomography Study. *NeuroImage*, 6, 320-334.

Kosslyn, S.M. (1999). If neuroimaging is the answer, what is the question? *Philosphical Transactions of the Royal Society of London B*, 354, 1283-1294.

Kosslyn, S.M., Rabin, C. (1999). Mental imagery. In R.A. Wilson and F.C. Keil (Eds.), *MIT Encyclopedia of Cognitive Science*. Cambridge, MA: MIT Press.

Kosslyn, S.M. (1999). Visual mental images as re-presentations of the world. In J. S. Gero & B. Tversky (Eds.), *Visual and Spatial Reasoning in Design.* Sydney, Australia: University of Sydney Press. 83-92.

Kosslyn, S.M., Sukel, K.E., Bly, B.M. (1999). Squinting with the mind's eye: Effects of stimulus resolution on imaginal and perceptual comparisons. *Memory and Cognition,* 27, 276-287.

Kosslyn, S.M., Thompson, W.L. (2000). Shared mechanisms in visual imagery and visual perception: Insights from cognitive neuroscience. In M. S. Gazzaniga (Ed.). *The New Cognitive Neurosciences, 2nd ed*. Cambridge, MA: MIT Press.

Kosslyn, S.M., Ganis, G., Thompson, W.L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, 2, 635-642.

Kosslyn, S.M., Thompson, W.L., Ganis, G. (2002). Mental imagery doesn't work like that. *Behavioral and Brain Sciences*, 25, 198-200.

Kosslyn, S.M., Ganis, G., Thompson, W.L. (2003). Mental imagery: Against the nihilistic hypothesis. *Trends in Cognitive Science*, 7, 109-111.

Kosslyn, S.M. (2003). Understanding the mind's eye...and nose. (News & Views.) *Nature Neuroscience*, 6, 1124-1125.

Kosslyn, S.M. (2003). Visual mental imagery: A case study in interdisciplinary research. In F. Kessel, P.L. Rosenfeld, & N.B. Anderson (Eds.), *Expanding the Boundaries of Health and Social Science: Case Studies in Interdisciplinary Innovation*. New York: Oxford University Press, 122-146.

Kosslyn, S.M., Thompson, W.L. (2003). When is Early Visual Cortex Activated During Visual Imgery? *Psychological Bulletin*, 129, 723-746.

Kosslyn, S.M. (2004). Preface to B. Libet, *Mind time: The temporal factor in consciousness*. Cambridge, MA: Harvard University Press.

Kosslyn, S.M., Thompson, W.L., Shephard, J.M., Ganis, G., Bell, D., Danovitch, J., Wittenberg, L.A., and Alpert, N.M. (2004). Brain rCBF and performance in visual imagery tasks: Common and distinct processes. *European Journal of Cognitive Psychology*, 16, 696-716.

Kosslyn, S.M. (2005). Mental images and the brain. *Cognitive Neuropsychology*, 22, 333-347.

Kosslyn, S.M. (2005). Reflective thinking and mental imagery: A perspective on the development of Posttraumatic Stress Disorder. *Development and Psychopathology*, 17, 851-863.

Kosslyn, S.M., Rosenberg, R.S. (2005). The brain and our students: How to explain why neuroscience is relevant psychology. In B. Perlman, L. I. McCann, & W. Buskist (Eds.), *Voices of Experience: Memorable Talks from the National Institute on the Teaching of Psychology*. Washington, DC: American Psychological Society. 71-82.

Kosslyn, S.M., Thompson, W.L., Sukel, K.E., Alpert, N. (2005). Two types of image generation: Evidence from PET. *Cognitive, Affective, & Behavioral Neuroscience*, 5, 41-53.

Kourtzi, Z., Huberle, E. (2005). Spatiotemporal characteristics of form analysis in the human visual cortex revealed by rapid event-related fMRI adaptation. *NeuroImage*, 28, 440-452.

Laeng, B., Teodorescu, D-S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. Cognitive Science, 26, 207-231.

Laeng, B., Chabris, C.F., and Kosslyn, S.M. (2003). Asymmetries in encoding spatial relations. In R. Davidson and K. Hugdahl (Eds.), *The Asymmetrical Brain*. Cambridge, MA: The MIT Press.

Levine, D.N., Warach, J., Farah, M. (1985). Two visual system in mental imagery: Dissociation of "what" and "where" in imagery disorders due to bilateral posterior cerebral lesions. *Neurology*, 35, 1010-1018.

Lambert, S. Sampaio, E., Scheiber, C., Mauss, Y. (2002). Neural substrates of animal mental imagery: calcarine sulcus and dorsal pathway involvement – an fMRI study. *Brain Reseach*, 924, 176-183.

Maleeh, R., Shafiee, A. (2006). Two Aspects of Information-Driven Interactions: Towards a Theory of Consciousness, submitted to *Anthropology and Philosophy*.

Mast, F.W., Merfeld, D.M., and Kosslyn. S.M. (2006). Visual mental imagery during caloric vestibular stimulation. *Neuropsychologia*, 44, 101-109.

Mast, F.W., Ganis, G., Christie, S., Kosslyn, S.M. (2003). Four types of visual mental imagery processing in upright and tilted observers. *Cognitive Brain Research*, 17, 238-247.

Mast, F.W., Kosslyn, S.M. (2002). Visual mental images can be ambiguous: Insights from individual differences in spatial transformation abilities. *Cognition*, 86, 57-70.

Mazard, A., Laou, L., Joliot, M., Mellet, E. (2005). Neural impact of the semantic content of visual mental images and visual percepts. *Cognitive Brain Research*, 24, 423-435.

Posner, M.I., Gilbert, C.D. (1999). Attention and Primary Visual Cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 2585-2587.

Pylyshyn, Z.W. (1973). What the Mind's Eye tells Mind's Brain: A Critique of Mental Imagery. *Psychological Bulletin*, 80, I, 1-24.

Pylyshyn, Z.W. (1981). The Imagery debate: Analogue Media versus Tacit Knowledge. *Psychological Review*, 88, 16-45.

Pylyshyn, Z.W. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioural and Brain Sciences*, 22 (3), 341-423.

Pylyshyn, Z.W. Is the Imagery Debate over? If so, what was it about? http://ruccs.rutgers.edu/faculty/pylyshyn-mehler.htm.

Pylyshyn, Z.W. (2002). Mental imagery: in search of a theory. *Behavioural and Brain Sciences*, 25 (2), 157-237.

Pylyshyn, Z.W. (2003a). Return of the Mental Image: are they really pictures in the brain? *Trends in Cognitive Sciences*, 7 (3), 113-118.

Reisberg, D., Pearson, D.G., Kosslyn, S.M. (2003). Intuitions and introspections about imagery: The role of imagery experience in shaping an investigator's theoretical views. *Applied Cognitive Psychology*, 17, 147-160.

Renier, L. Collignon, O., Poirier, C., Tranduy, D., Vanlierde A., Bol, A., Veraart, C., De Volder, A.G. (2005). Cross-modal activation of visual cortex durino depth perception using auditory substitution of vision. *NeuroImage*, 26, 573-580.

Roland, P.E., Gulyas, B. (1994). Visual imagery and representation. *Trends in Neurosciences*, 17, 281-287.

Rouw, R., Kosslyn, S.M., Hamel, R. (1998). Aspects of mental images: Is it possible to get the picture? *Cognition*, 66, 103-107.

Simmons, A., Miller, D., Feinstein, J., Goldberg, T., Paulus, M.P. (2005). Left inferior prefrontal cortex activation during decision-making task predicts the degree of semantic organization. *NeuroImage*, 28, 30-38.

Slotnick, S.D., Thompson, W.L., Kosslyn, S.M. (2005). Visual Mental Imagery Induces Retinotopically Organized Activation of Early Visual Areas. *Cerebral Cortex,* 15, 1570-1583.

Sparing, R., Mottaghy, F.M., Ganis, G., Thompson, W.L., Toepper, R., Kosslyn, S., Pascual-Leone, A. (2002). Visual cortex excitability during visual mental imagery – a TMS study in healthy human subjects. *Brain Research*, 938, 92-97.

Stiers, P., Peeters, R., Lagae, L., Hecke, P.V., Sunaert, S. (2006). Mapping multiple visual areas in the human brain with a short fMRI sequence. *NeuroImage*, 29, 74-89.

Stose, S.J., Kosslyn, S.M. (2001). Imagery versus propositional representation. In P. B. Baltes and N. J. Smelser (Eds.), *International Encyclopedia of the Social and Behavioral Sciences,* 11, 7193-7197.

Thompson, W.L., Kosslyn, S.M. (2000). Neural systems activated during visual mental imagery: A review and meta-analyses. In: A. W. Toga and J. C. Mazziotta (Eds.), *Brain Mapping: The Systems.* San Diego: Academic Press. 535-560.

Thompson, W.L., Kosslyn, S.M., Sukel, K.E., Alpert, N.M. (2001). Mental imagery of high- and low-resolution gratings activates Area 17. *NeuroImage,* 14, 454-464.

Toepper, R., Kosslyn, S.M., Pascual- Leone, A. (2002). Visual Cortex Excitability Increases during Visual Imagery – A TMS Study in Healthy Human Subjects. *Brain Research*, 983, 92-97.

Ungerleider, L.G., Mishkin, M. (1982). Two cortical visual system. In D.J. Ingle, M.A. Goodale, & R.J.W. Mansfield (Eds)., *Analysis of visual Behavior* (pp. 549-586) Cambridge, M.A: Mit Press.

Voellm, B., Taylor, A.N.W., Richardson, P., Corcoran, R., Stirling, J., McKie, S., Deakin, J.F.W., Elliott, R. (2005). Neuronal correlates of theory of mind and empathy: A functional magnetic resonance imaging study in a nonverbal task, *NeuroImage*, 29 (1), 90-98.

Sources on web

http://www.neuroingegneria.com/illusioniottiche/index.php
http://plato.stanford.edu/
http://www.imagery-imagination.com/index.htm
http://cogprints.org/3750/

Mariano L. Bianca - Lucia Foglia
Università di Siena
bianca@unisi.it
foglia3@unisi.it

*Le sezioni 2., 3., 4., sono a cura di Lucia Foglia. Le sezioni 1., 5., sono a cura di Mariano L. Bianca.

Anna M. Borghi - Claudia Scorolli
University of Bologna

Object Concepts and Mental Images

Abstract The paper focuses on mental imagery and concepts. First we discuss the possible reasons why the propositional view of representation was so successful among cognitive scientists interested in concepts. Then a novel perspective, the embodied view, is presented. Differently from the classic cognitivist view, this perspective acknowledges the importance of perceptual and motor imagery for concepts. According to the embodied perspective concepts are not given by propositional, abstract and amodal symbols but are grounded in sensorimotor processes. Neural and behavioral evidence favouring this perspective is presented. The paper discusses the continuity, but also the differences, between the imagery view and the embodied view of conceptual representation.

Key words Imagery, motor imagery, concepts, representation, embodiment.

1. Introduction

This paper will focus on the relationship between mental imagery and concepts. We will start by indicating possible reasons why the propositional view of representation has gained popularity in cognitive science, and trying to highlight some of its shortcomings. Then we will focus on a recent view of conceptual representation that acknowledges the importance of perceptual and motor imagery. In this perspective concepts are not given by propositional, abstract and amodal symbols but are grounded in sensorimotor processes. We will describe neural and behavioral evidence favouring this perspective. At a more general level, in this paper we aim at showing the deep continuity, but also the differences, between the imagery view of representation and one of the most successful recent theories of concepts, the "embodied" theory.

2. The propositional view of concepts and meaning: some reasons of its success

The propositional view has been and probably still is the dominant view among cognitive scientists who study concepts. Image-based theories have been quite popular among philosophers since the antiquity (Barsalou & Prinz, 1997; Prinz, 2002). Aristotle claimed "The soul never thinks without an image". Among others, Epicurus, Lucretius, Locke, and Berkeley underlined the importance of perceptual aspects for knowledge. The belief that images were crucial for thinking remained popular into the first part of the last century. In more recent times image-based views seemed to lose a lot of their appeal. Cognitive scientists have always demonstrated a certain resistance to imagery and preferred to endorse propositional-based theories. Let us address the question of why this happened and of why the propositional view had such a great success in the study of concepts.

A first reason is that the complexity of cognitive computations calls for the symbolic nature of representations, and language is the most powerful symbolic system in our possession. The advantage of propositional symbols is that, similarly to words, they are productive and can be differently combined in order to generate new meanings. Thus they can provide a powerful explanation of the way cognition works.

A second reason of why the propositional view was endorsed by the majority of cognitive scientists is that the most part of research on concepts was conducted using experimental paradigms that involved language. Consider for example a typical task used in order to access conceptual organization, the property generation task. In this task participants are required to produce all the properties that come to their mind for a given concept (for example, dog: tail, fur, head, barks, etc.). The output of such a task is a list of features. It is quite an easy step to identify the production output with the way in which objects are represented. Given that in standard feature production tasks participants generated features in order to describe concepts, this suggested that the way in which concepts were represented was based on propositional features.

A further reason of the resistance of cognitive scientists to an image-based view of concepts might be due to the theoretical and methodological weaknesses of the mental imagery perspective within the broad context of cognitive science. The literature on mental imagery was often misunderstood, and this misunderstanding led to simplifications that induced scientists to endorse a more easily defendable propositional view.

A final reason might lie in the fact that, in the years, different successful computational models of propositional view have been proposed (Collins & Loftus, 1975; Newell & Simon, 1976; Turing, 1950). Among the most recent models, Latent Semantic Analysis (LSA; Landauer & Dumais, 1997) and Hyperspace Analogue to Language (HAL; Burgess & Lund, 1997) have shown using disembodied symbols that statistical linguistics indeed approximates behavioral results (Foltz, Kintsch & Landauer, 1998; Landauer, McNamara, Kintsch & Dennis, 2006). Consider for example LSA (Landauer and Dumais, 1997) that uses a high-dimensional space to measure the association between words found in texts. LSA provides an index of a type of co-occurrence of words in similar texts, and this index has been demonstrated to be significantly correlated with psychological effects depending on associations, such as semantic priming.

Even though some propositional models of representation, as for example LSA, can be powerful instruments useful to investigate associations between words, they have the limits of using ungrounded and disembodied symbols. Arguments from Harnad, Searle and others have shown that meaning cannot rise without symbols grounding, so it follows that models like LSA cannot account for the emergence of meaning, or at least they cannot fully explain it. Namely, in accordance with claims of propositionalists, they equate meaning with relationships between abstract and amodal symbols that are arbitrarily linked to their referents in the world.

3. The embodied view of concepts and its relationship with image-based views of cognition

In the last ten years much has changed. The resistance of cognitive scientists to imagery-based representations has been overcome by the emergence of a new perspective, the embodied view, that is rapidly gaining success. According to this perspective knowledge is grounded in sensorimotor systems; knowledge acquisition and use are influenced by the characteristics of our body and its peculiar way to interact with the environment. Studies adopting an embodied perspective have produced an impressive amount of behavioral results that cannot be accounted for by propositional models based on the idea that meaning emerges from associations in a semantic network. Importantly for the aim of this paper, the endorsement of an embodied perspective has lead scholars to re-evaluate an image-based view of knowledge.

According to the embodied view, conceptual information is distributed over modality specific domains (Barsalou, Simmons, Barbey, & Wilson, 2003; Boronat, Buxbaum, Coslett, Tang, Saffran, Kimberg, & Detre, 2005; Gallese & Lakoff, 2005; Martin, Wiggs, Ungerleider, & Haxby, 1996). Thus, thinking of an object or of an entity leads to a re-experiencing (simulation) of the interaction with that object or entity. For example, thinking of a "dog" leads to the activation of multimodal information – the sound of the dog barking, its colour and shape, the smoothness of its fur while we caress it, etc. In this view, concepts are thus conceived of as "simulators", as they make it possible to run simulations (Barsalou, 1999; Barsalou et al., 2003). Simulations consist of re-enactments of our sensorimotor experiences with objects and entities. Namely, the neural areas recruited when we think about an object or an entity and prepare to act are the same that are recruited when we perceive and interact with its referent.

Adopting an embodied view of concepts implies the rejection of the idea of a transduction process from perceptual and motor states to propositional symbols. Here we advance three arguments useful to reject the idea of a transduction.

The first argument concerns the way evolution works. Evolution typically proceeds in a conservative way. There is no clear reason why in human beings a further process, the process of transduction, should be introduced between the object experience and the way we retain information on object.

The second argument regards the way science and scientific explanation typically work. In science the most parsimonious explanation of a process is typically the best one. If a simple explanation of a process can be provided, there is no reason why one should advance a more complicated explanation of the same process. The propositional explanation is clearly more complex than the embodied explanation. Namely, the first implies a transduction of experience into propositional symbols, that are arbitrary linked to their referents (objects and entities). On the contrary, the embodied explanation simply claims that while thinking and speaking we reactivate our previous experiences with objects and entities.

The third argument in favour of the embodied view is a provocative argument. In principle, there should be no need to defend an imagery based view of concepts. Objects have visual, acoustic, tactile features, and there is a rich body of evidence demonstrating that they are represented through the activation of visual, acoustic, tactile etc. brain areas. As we will show below, there is much recent evidence showing that object processing, as well as processing of words referring to objects, activate sensori-motor areas of the brain. So, in principle we would need counter-evidence, that is evidence showing that concepts are NOT grounded in sensorimotor processes. As far as we know, there is no convincing demonstration of the existence of a transduction process. Therefore, scholars who claim that object concepts are represented through propositional, quasi-linguistic symbols rather than modality specific features should provide convincing demonstrations of their claim.

As it should be clear from our presentation so far, the image-based and the embodied vision of concepts share the idea that concepts are not amodal, abstract, arbitrary and propositional symbols. However, this does not mean that imagery can be conceived of in a simplistic way as the process of creating “images in the head”. Also, it does not mean that images can be equated to “visual images”.

Concepts cannot be equated to “images” in the head because they are not holistic representations, but they are componential in nature. They are given by the activation of modality specific features across different domains. In line with connectionist claims, concepts are represented in a sub-symbolic rather than in a symbolic way (Elman, Bates, Johnson, Karmiloff-Smith, Parisi, & Plunkett, 1997): the single components are not symbolic per se, but meaning arises and symbols emerge from the activation of neurons in different modality specific areas. It is the compositional character of concepts that guarantees for their productivity.

The imagery components of concepts shouldn’t be simply equated with visual images. Rather, concepts can be conceived of as activation of neural patterns in different modality specific domains. There is compelling evidence showing that information is distributed in the brain across modality specific areas, and that information related to different features is stored in different brain regions (Martin et al. 1996; Boronat et al., 2005). In other terms, concepts imply the simultaneous activation of visual as well as tactile, motor, acoustic, taste features. In order to be productive, the semantic system must be compositional, and its components are different kinds of features distributed in different brain areas. These domains and these features are more or less activated depending on their relevance during knowledge acquisition. Importantly, according to this view, which is different from the classic information processing view, perception and action are seen as intimately related processes. In line with this account, various evidence on cortical object representation has shown that tools and manipulable objects, unlike non-manipulable artefacts, activate motor-related areas, whereas animals and natural objects activate vision-related areas (Gerlach, Law, & Paulson, 2002).

A potential difference between the traditional image-based view and the embodied view concerns the role played by "mental representation". Even if they defended an analogical rather than propositional form of representation, still the very idea of representation was crucial for defenders of image-based views of concepts. The same was true for traditional theories of concepts. Consider for example a claim by Keil (1995): "When I think of the category of dogs, a specific mental representation is responsible for that category and roughly the same representation for a later categorization of dogs by myself or by another". In the traditional view, concepts are representation; more importantly, they are rather stable and constant representations, and this explain the stability of our behavior. In antithesis with this position, within the embodied framework the dynamicity of behaviour has been underlined. In this framework, the notion of representation has been either dropped or deeply revisited. In fact, due to its theoretical ambiguity some scholars have proposed to dismiss it. As Smith (2005) claims, the dynamical character of cognition does not require stable forms of representation. For example, depending on the context we think of dogs in very different ways: "in the context of frisbies, we think of playful puppies, in the context of race tracks, we think of streamlined (and not at all playful) greyhounds, and in the contexts of muzzles, the main thought is fear" (Smith, 2005; p. 280). Other authors, though they still make use of the notion of representation, deeply revisit it as they equate representation with the activation of multiple neuronal areas (e.g., Barsalou, 1999; Barsalou et al., 2003; Gallese & Lakoff, 2005). Our impression is that the debate within defenders of the embodied perspective focuses more on a terminological rather than on a theoretical matter. Once representation is equated with the activation pattern of multiple modality specific areas, once the dynamicity rather than the stability of concepts is stressed, once cognition is intended as "a complex set of processes bound to each other and to the world through perception and action with no fixed and segregated representations of anything" (Smith, 2005; p. 279), then the very fact of using or not the notions of concepts and of representation is not crucial. Even if one chooses to stick to the notion of representation, this notion has completely lost its original meaning.

4. Evidence of the imagery character of concepts

Immediately after the most important proposals that concepts were grounded in sensorimotor processes were formulated (Barsalou & Prinz, 1997; Barsalou, 1999), a rich body of evidence in favour of the imagery character of concepts was collected. For example, Zwaan, Stanfield and Yaxley (2002) used a recognition task in order to address whether the linguistics input was converted into a propositional representation or into a 'mental image', that is a representation having an analogical relationship with its referent. They showed that, for example, the sentence "The ranger saw the eagle in the sky" lead to a faster recognition of a picture of a bird with outstretched wings than of a bird with folded wings, whereas the opposite was true with the sentence “The ranger saw the eagle in its nest”. This result, that is not predicted by a propositional view, suggested that subjects simulated in an analogical way the object shape implied by the sentence. Further evidence shows that the spatial and functional perspective suggested by a sentence activate mental imagery. In a study by Borghi, Glenberg and Kaschak (2004), participants read a sentence describing an object or a location from an inside (e.g., "You are driving a car ") or an outside (e.g., "You are fueling a car ") perspective. Then they were presented with a word and had to verify if the word referred or not to a part of the object / location. There could be inside or outside parts (e.g., seat vs. trunk) and parts located either near or far from the place where the action expressed by the sentence typically occurs (e.g., the inside part “back seat” is far from the place where the action of “driving” occurs, whereas the inside part “steering wheel” is near to the driver-place). Subjects performed the verification task more quickly when the perspective implied by the sentence and the part perspective matched. Then, within this perspective, they were quicker to verify near parts than far parts. These results clearly showed that the different perspectives implied by the sentences influenced the accessibility of information and made different conceptual knowledge available. Importantly, using latent semantic analysis (LSA), the authors ruled out the possibility that the results were due to semantic associations between words. Thus the results cannot be accounted for by a propositional theory and strongly suggest that during sentence comprehension we use imagery to mentally scan the situation described. Interestingly, the results found in these experiments are comparable with those obtained by Kosslyn, Ball and Reiser (1978). They also found quicker responses in scanning a smaller region of the image. Crucially, whereas Kosslyn et al. asked participants to generate a mental image, in the study by Borghi et al. the task did not require any explicit engagement of perceptual information: no imagery instructions were given and participants were presented only with linguistic stimuli. Clear evidence in favour of an image-based view comes from eye tracking studies that show that participants listening to stories describing objects orient their eye movements in the direction of the imagined object. For example they orient their eyes upward while listening to someone talk about skyscrapers, downward while listening to someone talk about canyons (Spivey & Geng, 2001).

The activated ‘mental images’ are not only visual, but multimodal, as an experiment by Pecher, Zeelenberg and Barsalou (2003) clearly shows. The authors selected concept nouns and properties regarding vision, motor action, audition, taste, touch and smell. Participants were presented with a sentence like ‘A *lemon* can be *sour*’ and they had to respond if the sentence was true or false. Crucially, the task did not require to create a mental image. Response times showed that switching modality, for example from a property related to taste (eg: *lemon – sour*) to an auditory property (*leaves – rustling*), required a cost. A control experiment ruled out the possible alternative explanation that *amodal* symbols for the same modality are more associated than *amodal* symbols for different modalities. These findings clearly favour the idea that concepts imply the simultaneous activation of different modality specific domains.

Evidence discussed so far suggests that concepts are made of 'perceptual symbols' (Barsalou, 1999). However, in the last years many studies focused on the relevance for concepts of motor information (and of motor imagery) (Glenberg, 1997). Consider for example a study performed the Positron Emission Tomography (PET), a scanning technique for monitoring regional blood flow (Grafton, Fadiga, Arbib & Rizzolatti, 1997). Authors found a specific activation of motor areas of the brain for manipulable man-made tools, like scissors or hammer. Importantly, the left premotor cortex was activated not only when subjects had to say the use of the object or when they had to name the object, but also when the task consisted only of looking at the object. This suggests that both naming and observing tools activate motor imagery and that perception and action are strictly interwoven. The role of motor imagery and action for concepts has been shown also in behavioural studies. For example, Borghi et al. (2004) used the same task used in the previous experiment but chose sentences that did not imply any action and selected upper and lower object parts (e.g, roof vs. wheel of a car). Participants performed the task using as responding device a vertically oriented button box. They had to move the hand upward or downward to press the response key. Responding in a direction incompatible with the part location (e.g., responding downward to verify that a car has a roof) was slow relative to responding in a direction compatible with the part location. A recent debate concerns the degree of specificity of the simulations activated during language comprehension. Different behavioural and neuro-physiological experiments demonstrate that these simulations are quite specific (see for example Buccino, Riggio, Melli, Binkofski, Gallese, Rizzolatti, 2005). Pulvermüller, Härle and Hummel (2001) investigated brain activity elicited by visually presented verbs that could be referred to movements of the arms (eg. *to write*), of the legs (e.g. *to walk*) or of the face muscles (e.g *to talk*). The behavioural part of the study consisted in a lexical decision task. In the physiological part they recorded Event Related Potentials (ERPs), that is a measure of the electrical activity produced by the brain in response to a sensory stimulus or associated with the execution of a motor, cognitive, or psycho-physiologic task. Behavioural results showed faster response times for face related verbs followed by arm related verbs and leg related verbs, supporting the idea that words semantic properties are reflected in the brain response they induce. Recorded ERPs revealed significant topographical differences: different kinds of verbs, referring to actions performed using different effectors, are processed in different ways in the brain.

These conclusions are confirmed in a behavioural study by Scorolli & Borghi (submitted). Participants were presented with pairs of nouns and verbs that could be referred to hand and mouth actions (e.g., to unwrap vs. to suck the sweet), or to hand and foot actions (e.g., to throw vs. kick the ball). Their task consisted of deciding whether the combination made sense or not: a group of participants responded by saying *yes* loudly into a microphone, another group by pressing a pedal. Results suggest that sentence processing activates an action simulation. This simulation is quite detailed, as it is sensitive to the effector involved. Namely, a facilitation emerged in responses to 'mouth sentences' and 'foot sentences' compared to 'hand sentences' in case of congruency between the effectors – mouth and foot – involved in the motor response and in the sentence. A propositional theory can hardly account for these results. Namely, participants were simply asked to decide if a sentence like *to kick the ball* was sensible or not sensible. If concepts mediated by words were abstract, amodal and arbitrarily related to their referents, why should the sentence '*to kick the ball*' produce faster responses with the pedal than "*to throw the ball*"?

5. Conclusion

As we have seen, the embodied view incorporates and extends many ideas initially proposed in a different cultural and historical contexts by authors working on mental imagery. However, these ideas have been deeply modified and revisited. The critiques of authors in favour of a propositionalist account have been taken into consideration and many of the weaknesses of the traditional mental imagery approach have been eliminated. For example, proponents of the embodied cognition perspective got rid of some reductive and misleading interpretations of mental imagery, as for example the idea that we have "images in the head" or the equation of "mental imagery" with "visual mental imagery". In addition, the notion of "mental representation", that was extensively used by defenders of the image-based view of cognition, was either rejected or deeply transformed in the context of embodied views of cognition.

A final remark should be made regarding the importance of imagery for concepts according to the embodied perspective. Immediately after the most important proposals that concepts were grounded in sensorimotor processes were formulated (Barsalou & Prinz, 1997; Barsalou, 1999), the role played by imagery (and particularly by visual imagery) for conceptual representation was stressed. As our brief review has shown, initial evidence favouring an embodied approach to concepts showed that thinking of an object or comprehending a word implied referring to the perceptual aspects of its referent (MacWhinney, 1999; Zwaan, 2004). In the last years the panorama has slightly changed. More and more authors who adopt an embodied perspective underline the importance for concepts not just of perceptual but of sensorimotor and particularly motor processes. More importantly, in the last years it has become more clear that perception and action are not separate and sequential processes but that they are deeply interwoven. In sum, the centrality of action for concepts has been fully recognized (for a review see Borghi, 2005). The recognition of this centrality has been favoured by the diffusion of studies on motor imagery and by the broad literature on the motor resonance processes triggered by mirror neurons in both monkeys and humans. Consider for example the way in which the term "motor image" has changed: whereas this term classically referred to explicit or conscious representation of an action (e.g., imagine yourself running), recent research focuses also on implicit or inconscious aspects, and with "motor imagery" we refer to a subliminal activation of the motor system (Jeannerod & Frak, 1999). At the same time, studies on mirror neurons have shown the relevance of motor resonance processes (see discussions on the role played by mirror neurons for concepts, see Borghi and Scorolli, submitted).

Once established that concepts strongly rely on sensorimotor processes, a further problem that needs be addressed concerns the ornamental character of modal representations. Proponents of an amodal, propositional view might claim that is possible that objects are represented through modality specific features, but that these features are not NECESSARY for concept representation. This objection reminds of the critiques advanced by Pylyshyn (1973) to imagine-based theories. Phylyshyn (1973; but see also 2003) argued that images were just epiphenomenal to the context of thinking. Some recent studies on sentence comprehension provide some useful cues in order to solve the question. For example, data from Borreggine and Kaschak (in press) suggest that the ACE effect, at the very least, was due to the simultaneous occurrence of a motor preparation phase and sentence comprehension. Namely, the motor resonance effect could occur either during sentence comprehension or after the sentence has been understood in order to prepare for action. If it occurred after sentence comprehension, motor imagery would not be necessary for comprehension. In order to solve this complex matter, more detailed analyses of timing during conceptual and language comprehension tasks are needed. Providing an answer to this objection represents one of the most exciting challenges of current research in the field of embodied cognition.

References

Barsalou L. W. (1999). Perceptual Symbol Systems. *Behavioral and Brain Sciences*, 22, 577-609.

Barsalou, L.W. & Prinz, J.J. (1997). Mundane creativity in perceptual symbol system. In T.B. Ward, S. M. Smith, & J. Vaid (eds). *Conceptual structures and processes: Emergence, discovery, and change*. Washington, DC: American Psychological Association.

Barsalou, L.W., Simmons, W.K., Barbey, A.K., & Wilson, C.D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Science*, 7, 84-91.

Borreggine, K.L., & Kaschak, M. (in press). The Action-sentence Compatibility Effect: Its all in the timing. *Cognitive Science.*

Borghi, A.M. (2005). Object concepts and action. In D. Pecher & R.A. Zwaan (Eds). *Grounding Cognition: The role of perception and action in memory, language, and thinking.* Cambrige University Press.

Borghi, A.M., Glenberg, A.M., Kashak, M.P. (2004). Putting words in perspective. *Memory and cognition*, 32, 863-873.

Borghi, A.M., Scorolli, C. (submitted) Sentence comprehension and hand motion simulation.

Boronat, C.B., Buxbaum, L.J., Coslett, H.B., Tang, K., Saffran, E.M., Kimberg, D.Y., Detre, J.A. (2005). Distinction between manipulation and function knowledge of objects: Evidence from functional magnetic resonance imaging. *Cognitive Brain Research*, 23, 361-73.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., Rizzolatti, G. (2005). Listening to action related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research*, 24, 355-63.

Burgess, C., & Lund, K. (1997). Modeling parsing constraints with high-dimensional context space. *Language and Cognitive Processes*, 12, 177-210.

Collins, A.M. & Loftus, E.F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, 82, 407-428.

Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D., e Plunkett, K. (1997), *Rethinking Innateness. A Connectionism Perspective on Development,* Cambridge, MA, MIT Press.

Foltz, P.W., Kintsch, W. & Landauer, T.K. (1998). The measurement of textual Coherence with Latent Semantic Analysis. *Discourse Processes*, *25,* 285-307.

Gallese, V., Lakoff, G. (2005). The brain's concepts: The role of the sensorimotor system in conceptual knowledge. *Cognitive Neuropsychology*, 21, 455-479.

Gerlach, C., Law, I., & Paulson, O. B. (2002). When action turns into words. Activation of motor-based knowledge during categorization of manipulable objects. *Journal of Cognitive Neuroscience*, 14, 1230-1239.

Glenberg, A.M. (1997). What memory is for. *Behavioral and Brain Sciences*, 20, 1-55.

Grafton, S.T., Fadiga, L., Arbib, M.A., & Rizzolatti, G. (1997). Premotor cortex activation during observation and naming of familiar tools. *Neuroimage*, 6, 231-236.

Keil, F. (1995). The growth of causal understandings of natural kinds. In: *Causal cognition: A multidisciplinary debate*. Sperber, D., Premack, D., Premack, A. (Eds.) Oxford, England, Oxford University Press.

Jeannerod, M. and Frak, V. G. 1999. Mental simulation of action in human subjects. Current Opinion in Neurobiology, 9, 735-39.

Kosslyn, S.M., Ball, T.M., & Reiser, B.J. (1978) Visual images preserve metric spatial information: Evidence from studies of image scanning, *Journal of Experimental Psychology: Human Perception & Performance*, 4, 47-60.

Landauer, T.K., Dumais, S.T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.

Landauer, T., McNamara, D.S., Dennis, S. & Kintsch, W. (Eds.). *LSA: A road to meaning*. Mahwah, NJ: Erlbaum.

MacWhinney, B. (1999). The emergence of language from embodiment. In B. MacWhinney (ed.). The emergence of language. London: Erlbaum (pp. 213-256).

Martin, A., Wiggs, C.L, Ungerleider, L.G., & Haxby, G.V. (1996). Neural correlates of category specific knowledge. *Nature*, 379, 649-652.

Newell, A., & Simon, H.A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the Association for Computing Machinery*, 19, 113-126.

Pecher, D., Zeelenberg, R., & Barsalou, L.W. (2003). Verifying properties from different modalities for concepts produces switching costs. *Psychological Science*, 14, 119-124.

Pylyshyn, Z.W. (1973). What the Mind's Eye Tells the Mind's Brain: A Critique of Mental Imagery. *Psychological Bulletin*, 80, 1-24.

Pylyshyn, Z.W. (2003). Mental imagery: in search of a theory. *Behavioral and Brain Sciences*, 25, 157.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology,* 9(2), 129-154

Prinz, J. (2002). *Furnishing the mind. Concepts and their perceptual basis*. Cambridge, MA: MIT.

Pulvermüller, F., Härle, M. & Hummel, F. (2001). Walking or talking? Behavioral and neurophysiological correlates of action verb processing. *Brain and Language*, 78, 143-168.

Rizzolatti, G., Craighero, L. (2004). The mirror neuron system. *Annual Review of Neuroscience,* 27*,* 169-192.

Scorolli, C. & Borghi, A.M. Sentence comprehension and action: Effector specific modulation of the motor system.

Smith, L. B. (2005). Cognition as a dynamic system: Principles from embodiment. *Developmental Review,* 25, 278-298.

Spivey, M.J., Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychological Research*, 65, 235-241.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59, 433-460.

Zwaan, R. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B.H. Ross (ed.), *Psychology of learning and motivation*, (vol.44, pp.35-62). New York: Academic.

Zwaan, R., Stanfield, R.A., Yaxley, R.H., 2002. Language comprehenders mentally represent the shape of objects. *Psychological Science*, 13, 168-171.

Anna M. Borghi - Claudia Scorolli
Università di Bologna
annamaria.borghi@unibo.it
claudiascorolli@libero.it

Francesco Ferretti
University of Roma Tre

Imagery, Perception and Creativity

Abstract The aim of this paper is to justify the role of mental imagery in creativity. In more specific terms the central idea of this paper is that the justification for the role of mental images in the creative process lies in the analysis of the relationship between vision and imagery. Mental images are present in thought just in those situations in which the ideal way to solve a problem would be the perception of those same things before our own eyes. The production of a mental image is, from this point of view, an analogous process to the perception of an object in its absence. In these cases the images become a good substitute for reality: this is possible because they share common structural elements with percepts. These shared common characteristics provide a validation to the idea that the role of imagery in creativity depends on the structural properties which mental images share with the representation of the perceived information. Such a result opens the way to the idea that there is a close link between imagery and vision and, thus, provide justification for the theory that imagery is closer to perceiving than to thinking.

Key words Creativity, imagery, visual perception, blindness.

«Early in my career I got into fights with other engineers at meat-packing plants.
I couldn't imagine that they could be so stupid as not to see
the mistakes on the drawing before the equipment was installed.
Now I realize it was not stupidity but a lack of visualization skills.
They literally could not see» (Grandin, 1996: 26).

The harsh opinion expressed by Temple Grandin[1] towards the engineers with whom she was working on the design and planning of ranches and cattle pen layouts serves to highlight the distinction between logical-linguistic and imaginative-spacial abilities in thought processes. Grandin's idea is that really innovative design and planning are linked to mental images. It is true that this is an intuitive and entirely introspective judgement, but similar anecdotes, recounted by scientists, artists and writers, are so numerous and detailed as to lead us to believe that there is a connection between imagery and creativity. The aim of this work is to justify the intuitive judgement underlying such experiences from both theoretical and empirical perspectives.

In more specific terms the idea central to this paper is that the justification for the role of mental images in the creative process lies in the analysis of the link between imagery and perception. That perception – those cognitive devices with which an organism relates to the surrounding world – is an inexhaustible source of information constitutes a patently obvious fact. Less obvious is the fact that imagery is a similar source of information. In this paper we aim to examine what at first sight appears to be an obvious paradox: if mental images are the product of memorised information (and therefore *already available* to the subject), how can they be used in situations where *new* information is discovered? It is our theory that in order to explain the role of mental images in creative processes (thus solving the paradox) it is necessary to examine the relationship between imagery and visual perception.

1. Perception and interpretation

Although it may seem superfluous to draw attention to the fact, perception and imagery are cognitive capacities with different characteristics. Thus, imagery comes into play when there is insufficient perceptive information and it is in these cases that mental images become a good substitute for percepts. In order to understand what is behind the relationship between imagery and vision it is first of all necessary to start from visual perception.

Initially cognitive science used the metaphor of the computer as its basic reference point. This metaphor has given rise to a conception of mental processes in which inferential processes play a significant role. Models of visual perception also come into this picture. According to Gregory (1966), who considers perception as analogous with scientific research procedures, seeing is the result of a series of hypotheses which the subject projects onto the world. Perception is akin to a problem-solving activity. Taken thus, perception can be considered as a process of «interpretation» of the external world. Rock (1982) puts forward a conception of visual perception which is radically conceptual (or epistemic). In this case each perceptive act is an interpretative process which involves the beliefs and the background knowledge of subjects. Ambiguous figures (such as the celebrated duck-rabbit of Jastrow) highlight the fact that perception cannot be determined by sensory stimulus alone. The same stimulus can give rise to different interpretations of the same retinal images, that is, it can be seen now *as* a duck, now *as* a rabbit.

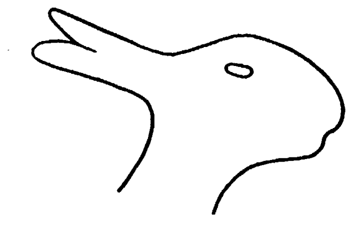

Figure 1. The famous Jastrow duck-rabbit.

Pylyshyn (1981) defines as «cognitive penetrability» the theory by which perceptive content is largely determined by the background knowledge of the subject. An author who has very much insisted upon this point is Bruner (1957), a true precursor of the *New Look* in the psychology of vision. This idea also has deep repercussions in epistemology; Kuhn (1962), Hanson (1958) and Goodman (1968), for example, claim that perception of the world depends on how the perceiver *conceives* the world. In fact, two individuals (two scientists) who observe the same phenomenon starting from two different theories see two different things. Thus it is how we interpret the world which determines what we see.

There is no doubt that there are aspects of perception which may be interpreted in terms of cognitive penetrability. But is it really possible to conceive *all* perception in this way? The answer must be no. Some aspects of visual perception show a very strong resistance to conceptual penetration. Fodor (1984) attempts to limit extreme forms of the interpretativistic viewpoint by making a distinction in visual perception between «belief fixing» and «appearance fixing». With the aim of re-establishing the right balance between interpretation and observation, Fodor shows how the background knowledge of subjects is impermeable to visual perception (or at least a part of it). The typical case is that of optical illusions: *knowing* that the two horizontal lines of the Muller-Lyer illusion are of equal length does not prevent us *seeing* them as different.

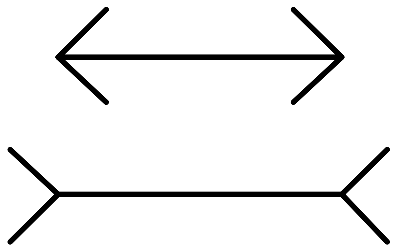

Figure 2. The Müller-Lyer illusion.

Criticism of the interpretative hypothesis takes into account the common sense intuition (underlying *fisica ingenua*) of the independence of reality from conceptual categories in the epistemic subject. If modular processing is independent of background theory, then it is possible to claim that two scientists sustaining two different theories are, in fact, observing the same world (even though they interpret it differently). According to Fodor, the point is distinguishing between fixing appearances and fixing the perception of beliefs. While it is clear that the latter calls into question the holistic relationship between beliefs, as regards the former (the fixing of appearances), it is possible to speak of a fixing process which is largely independent of the background beliefs of the individual. The theory of Fodor is put forward as a valid reconciliation of two aspects (both essential) of perception theory.

With regard to experimental phenomenology, Kanizsa (1980, 1991) provides proof, in a series of experiments which have a significant visual impact on the observer, that it is not possible to reduce perception to inferential processes. Amodal completion (Fig. 3), one of the essential elements of visual perception (those processes which transform the fragmentary input we find before our eyes into a single reality) lends itself perfectly to demonstrating the existence of a specific perceptual character (primary vision), which cannot be reduced to general laws of thought. There is no question that some aspects of perception can be assimilated with thought. What is called into question is the theory that thought and perception are one and the same. According to Kanizsa (1991) the organisation of primary visual perception comes about before and independent of signification processes which attribute an interpretation to the world. Indeed, at this level, although there is no meaning,

> there is organisation, that is segmentation, structure, precise spacial, chromatic, dimensional and topological relationships. Everthing but the sense data so dear to philosophers in the English-speaking world, raw material, disorganised as well as without sense, "blotches of colour" waiting to be organised. (…) But if there can be organisation without meaning, (…), the meaning is attributed to a visual reality which is already segmented into distinct objects and provided with a shape. Even if the process of incorporation of the meaning is not generally observable, the constitution of the visual object must necessarily precede its recognition. It can only be recognised insofar as it already exists (ivi: 21).

Even though it is not possible to consider primary perception as passive, we should also recognise that it is a form of perception more similar to «receiving» than «giving». The perceived object imposes itself on the observer (the white rectangle imposes itself on the observer as if it were a real object).

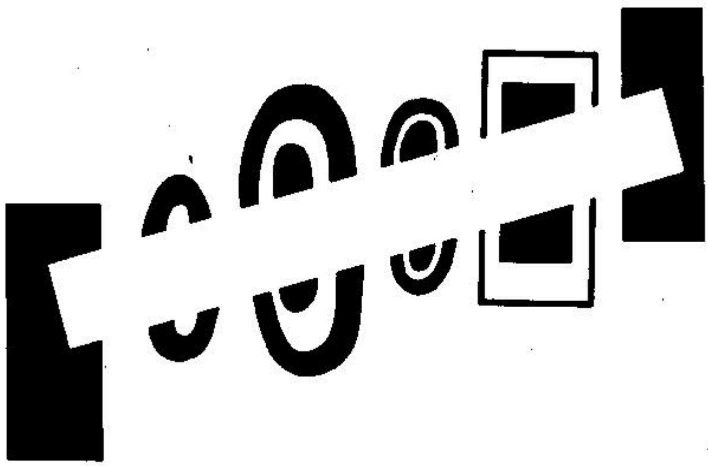

Figure 3. Amodal completion. From Kanizsa (1980: 283).

Amodal completion shows how far this level of organisation of experience is independent of the beliefs or wishes of the perceiver. What Kanizsa demonstrates is that at least some aspects of the organisation and structuring of experience (primary perception) are different from thought. A similar argument may also be put forward for imagery. The central point in the debate on the relationship between imagery and perception is to understand if and to what extent imagery may be considered similar to primary perception.

2. Imagery and perception

Our starting point concerns functional analysis; images are not always necessary in our thought, but only in specific circumstances, unusual or new situations in which what we have in mind should ideally be before our eyes. It is in these circumstances that images become a good substitute for sensory perceptions. Denis (1991) provides a definition to support this claim:

> A considerable body of data converges on the idea that imagery provides individual with cognitive products which can be used in the same way as products of perception, and have behavioural effects that are often very similar to those of perception (ivi: 38).

Here, the basic idea is that the peculiar role of imagery in the creative process depends on certain common characteristics shared between mental images and percepts. How is it possible to justify this hypothesis? How may the production of an image be considered as the perception of an object which is not physically present?

It must first be said that the issue here is not a presumed identity or match between the two cognitive models (Kosslyn et al., 2001). Since imagery, just like perception, is a complex system of processing devices, a possible way of dealing with the problem is by «levels of equivalence» (Finke, 1989). How far is it possible to extend the analogy between imagery and vision? An enormous series of studies have been dedicated to this theme, to the point that we may say that current debate on the nature of imagery is based almost entirely on this theme (Bértolo, 2005; Kaski, 2002).

The first step in dealing with the problem is that of understanding if the similarities which are found at a functional level may also be found in equivalent processing devices. One way of dealing with the question is that of understanding which parts of the perception processing system are also employed in imagery. The sophisticated techniques available today for mapping brain activity provide detailed data about this aspect of the question. The fact that *some* components of visual perception are active in visual perception tasks is not controversial. There is no calling into question, for example, that the parieto-occipital and temporo-occipital associative areas are involved in carrying out imagery tasks (Roland and Gulyàs, 1994). What is controversial is whether or not the primary visual areas are involved. This point is controversial because around the question as to whether images are more similar to thinking (conceptualising) or seeing rekindles the debate on the figural nature of mental images. Finke et al. (1989) thus summarise the issue:

> The process of perception begins with the geometry of the retinal images, and ends with a description of objects in the world. The controversy over imagery has largely concerned whether images are like early perceptual representations containing information about the geometric properties of visual inputs, or like later cognitive representations containing information about the conceptual categories of interpreted objects (ivi: 51-2).

Since, as we have mentioned in the introduction, our idea is that the role of imagery in the creative process depends on the fact that mental images share common functional and structural features with percepts, demonstrating the involvement of the primary visual areas in imagery would provide reasonable backing for our theory.

Blindness and imagery

A good way to understand how far it is possible to extend the analogy is that of examining an extreme case: by studying the relationships between imagery and visual perception in congenitally blind subjects. Kennedy (1980, 1983, 1993) studied the representational capacities of the blind analysing their drawings. Kennedy and Jurecevic (???) describe the case of the congenitally blind Turkish painter Esraf Armagan who was asked to draw a cube taking into consideration variations in the observer's viewpoint. His drawings are quite remarkable, since they respect perspective, a property which depends on the perception of a view point at a distance from the object (Fig. 4).

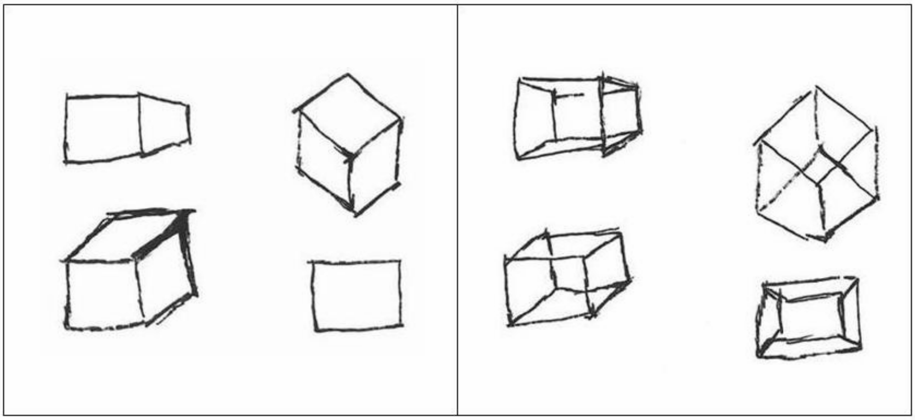

Figure 4. The congenitally blind Esraf Armagan who was asked to draw a cube taking into account variations in the observer's viewpoint. Figure from Kennedy e Juricevic (???): http://www.utsc.utoronto.ca/~kennedy/one-point.pdf

So how is it possible for the blind to represent perspective in images of the world which are typical of the viewpoint (at a distance) of the sighted?

Masini and Antonietti (1992) claim that «pictorial» perception through touch confirms that representation «at a distance» in space is not a specific characteristic of vision. According to these two authors, the fact that the blind can carry out tasks which seem to involve visual representations is not dependent upon visual perception. Kerr (1983) holds a similar view. A direct consequence of these considerations is that the properties to which we have referred as visual properties (viewpoint, perspective, occlusion) can no longer be considered such.

Nevertheless, the data coming from studies on the blind could be interpreted in the light of an alternative theory, a theory which also recognises visual-spacial properties in the mental representations of the blind. The hypothesis supporting the previous argument maintains that since the blind have not had visual experiences, the figural character of their representations cannot depend on visual perception. Against this interpretation we shall claim that it is not necessary to have had a visual experience in order to have a representation with visual characteristics. Firstly, in fact, claiming that a person can simply see or not see is to take the question from the wrong viewpoint: visual perception is a complex system (with a vertical as well as a horizontal dimension). Many blind people have only damage to peripheral centres of their visual system, while they conserve to a great extent whole the processing systems at a higher level. And if it were possible to activate these higher level visual centres independently from light stimulation of the retina? Could we then say that the representations produced in this way would be wholly independent from visual perception?

Cornoldi et al. (1992) make an important step towards recognising visual-spatial properties in the mental images of the blind. They call into question the *visual buffer* of Kosslyn (the functional space in which visual images are generated):

> Since, on the other hand, [the blind] cannot have, in long-term memory, traces derived from visual experience, we must think that such images are the result of the construction in an active system of the memory (called active memory, buffer, visuo-spacial notebook or – a term referred to in the present work – visuo-spacial working memory) of a represention with visuo-spacial properties processed on the basis of information acquired through non-visual means (ivi: 183-84).

Nevertheless, in spite of this recognition, their judgement is that: «La ricerca sulle immagini dei ciechi congeniti totali (...) ridimensiona ulteriormente l’assunzione di una analogia fra percezione visiva e immaginazione visiva: se l’immagine non è la filiazione diretta della percezione appare infatti altamente improbabile che una serie di processi, per quanto supportati da sistemi in parte condivisi con la percezione visiva, porti a risultati perfettamente sovrapponibili» (ivi: 185). If we are seeking a «perfect match» between visual perception and imagery, Cornoldi et al. are quite right. But is it really this level of relationship which is at the centre of the debate on mental images?

By saying that the blind have visual-spacial representations we do not intend to claim that they have representations which *perfectly match* those of the sighted. What we intend to claim is that the difference between the two representational capacities is not sufficient to consider the mental representations of the blind as completely void of figural characteristics. This point is of great significance and deserves further attention.

Lambert et al. (2004) in a fMRI study provide data which support the involvement of the primary visual areas (PVA) in visual imagination tasks carried out by congenitally blind subjects. They arrive at the conclusion that: «without going quite as far as to say that blind subjects did generate visual mental images, we support the idea that, as in sighted subjects, mental imagery mechanisms undergo perceptual influence in blind people» (ivi: 10).

The most interesting data on whether visual experience is necessary to produce visual images come from studies on the dream activity of congenitally blind subjects. Since dream activity is associated with visual perception for the sighted, the prevailing idea is that the blind are unable to dream. Against this idea Bértolo (2005) suggests examining an alternative path:

> Dreams with visual content are expressions of visual imagery. Therefore if dreams with visual content could be demonstrated in congenitally blind persons, this would imply that visual imagery is possible in subjects who have been prevented from having visual experiences. Furthermore, this would allow one to infer that visual imagery does not depend on specific visual perception, but can emerge from activation of visual cortex by non visual inputs (ivi: 180).

Firstly, a number of studies have proved that the blind have dream experiences showing that the dreams of those who cannot see «contain mostly sounds, touch sensations or emotional experiences» (Bértolo et al., 2003: 277). However, there is still strong support for the idea that the dreams of the congenitally blind or those becoming blind before the age of 5-7 have no visual content (Kerr, 2000). The question of visual content is extremely controversial because it concerns subjective experience and is, therefore, a factor which is difficult to demonstrate experimentally. Nevertheless, today we have available objective tests for the presence of visual processing in dream activity. The test used by Bèrtolo et al. is the EEG and the measurement of alpha power as an index of visualisation activity. Effectively, alpha rhythm is an index of the visual activity in which the brain is engaged. Studies carried out on sighted subjects have shown that alpha activity attenuation or blocking is a good indicator of visual imagery capacity.

Lopes da Silva (2003) was the first to provide empirical evidence (in a study based on EEG) for the decrease of alpha strength recorded from the central and occipital regions of blind subjects carrying out visual imagery tasks. In so doing he provides proof in support of the idea that «subjects who have never had visual experiences can have dreams with virtual images that are probably mediated by the activation of the cortical areas responsible for visual representations» (ivi: 328). Starting from this basic correlation, Bértolo et al. (2003) studied the dream activity of blind subjects concentrating, in particular, on the visual content of their dreams. The most significant result of this study regards «a significant negative correlation between the visual content of the dreams and the alpha power» (Bértolo et al., 2003: 282).

Figure 5. Graphical representation of an oneiric scene of a blind subject. From Bértolo et al., 2003: 181.

Figure 5 shows the drawing of one of the blind subjects asked to reproduce graphically the dream previously described verbally. No statistically significant difference was found between blind and sighted subjects. As well as being able to verbally describe the visual content of their dreams, the blind subjects were also able to draw a graphic representation of the same visual content. Bértolo arrives at the conclusion that:

> According to these results, the congenitally blind, who have never experienced sight, are able to visualise (…) The observation of alpha attenuation/visual content correlation along with the no differences in the graphical representations leads us to hypothesize that blind subjects can produce virtual images, that is, that their dreams correspond to the activation of visual cortical regions (Bértolo, 2005: 183).

With regard to the debate on the relationship between visual perception and imagery, the most significant data emerging from studies on congenitally blind subjects concerns the fact that the visual character of the representations in play in the tasks assigned does not effectively depend on a visual experience, but on the sharing of a specific processing system. As Farah (1988) in fact underlines:

> Imagery is not visual in the sense of necessarily representing information acquired through visual sensory channels. Rather, it is visual in the sense of using some of the same neural representational machinery as vision. That representational machinery places certain constraints on what can be represented in images and on the relative ease of accessing different kinds of information in images (Farah, 1988: 315).

The question of the activation of the primary visual areas in visual imagination tasks is among the most controversial in the current debate. There is ample experimental proof to back up such an hypothesis (Ishai and Sagi, 1995; Klein et al., 2004; Kosslyn and Ochsner, 1994; Miyashita, 1995; Sakay and Miyashita, 1995). But there are also experimental results which would indicate quite different conclusions. In support of the theory for the independence of imagery from vision Pylyshyn (2003) provides a number of studies relative to subjects with brain damage; Kaski (2002) outlines a series of experiments which exclude the activation of the primary visual areas in imagery tasks. After explaining how to interpret these contradictions, Kosslyn et al. (2001) define the question in quantitative terms: the overwhelming majority of experimental data available today provide convincing evidence for the theory that the primary visual areas are involved in imagery tasks. It would therefore seem that the idea of a processing system common to both imagery and vision is amply backed up by empirical evidence.

Shape and content

In the debate on mental images, why is it so important to recognise the involvement of the primary areas in imagery tasks? For two reasons: the first is that confirming such an involvement means guaranteeing a close link with the processes which give rise to the basic characteristics of the perceptive stimulus; the second is that the retinotopically-organised nature of the primary visual cortex can be used to settle an old issue: the dispute between advocates, on the one hand of the propositional nature and, on the other, the pictorial nature of mental images, which in the late twentieth century characterised the «imagery debate» from the beginning of the seventies up until the nineties. In fact, demonstration of the involvement of the primary visual areas in imagery tasks provides, according to Kosslyn (1994; and Kosslyn et al., 2001), decisive proof in favour of the figural nature of mental images. For two reasons:

> First, these areas are topografically-organized: they preserve (roughly) the local spatial geometry of the retina, so patterns of activation in them depict shape. If these areas are activated during imagery, and such activation has a functional role, this would be evidence that imagery relies on representations that depict information, not describe it. (In other words, this would be evidence that mental imagery relies on actual images). Second,

> such findings could not be explained solely by "tacit knowledge" stored as descriptions, which Pylyshyn used to explain away the findings from earlier behavioural experiments that attempted to show that imagery relies on depictive representations (Kosslyn et al., 2001: 639).

Pylyshyn (2003), however, claims that the data in favour of the relationship between imagery and vision provide no evidence regarding the question of representational format since «imagery and vision might involve the very same form of representation without it being pictorial in either case (indeed the case against pictorial representations in vision is as strong as it is for imagery)» (ivi: 115). Pylyshyn is right to highlight the fact that in vision what is perceived is the object and not its image. This objection is justifiable, but his argument does not go far enough to exclude the presence of figural characteristics in visual perception: the best theory of vision to date, put forward by Marr (1982), is, in fact, based on the processing of the «shape» of objects. Considering visual perception on the basis of this type of processing has important repercussions on the idea of the relationship between shape and content and, thus, on the idea of representational format.

The role of shape in the recognition of objects has an influence on categorisation processes. The concept of dog which adults have is different from the perceptive categorisation which infants employ in order to recognise a dog. In the concept of dog there are high level properties largely dependent upon language (that the dog is an excellent pet or that it gives birth to puppies, for instance) which are not perceptible in the form of the dog. As well as these properties, however, there are others at a more basic level which are dependent upon categorisation based on perceptive data. Quinn and Eimas (1986) demonstrated, for example, that the recognition of objects is mediated by the ability to perceive «closed forms», proving the fact that categorisation, at least in the initial phases, employs bottom-up and not top-down process. From this point of view perceptive categorisation works as a «primitive base» on which the concepts of adults can be constructed. A similar position is that of Mandler (2004a, 2004b). That perceptive categorisation must provide the basis of concepts (which make use of linguistically codified knowledge) is, moreover, shown by the fact that the attribution of a name to an object *presupposes* the recognition of that object as the occurrence of a type. According to Landau et al. (1998) shape similarity constitutes a basic element in the acquisition of lexis because it guarantees the mechanism for generalising new names to new objects:

> however the meanings underlying object names are ultimately characterized shape similarity constitutes a critical bootstrapping mechanism operating to initiate learning of object names in young children by allowing them to identify category members in the absence of dense knowledge about category (Landau et al., 1998: 21).

Theoretical interpretations of this type provide evidence in favour of the role of shape recognition in categorisation processes: one part of the cognitive content which the subject employs to recognise objects in the world is preverbal by nature and depends on the perceptive analysis of the shape of objects.

At this point, the moral of the story. Examples of preverbal categorisation show that at least a part of the information content (necessary for the recognition of objects), acting as a trigger for high level interpretation mechanisms, depends essentially on the perception of the shape of the object. Claiming that visual perception is linked to the shape of objects has important repercussions for the central theme of this paper: it means highlighting the constitutive and not simply expressive role of shape in the representation of the content. The point is that a degree of similarity links the structure of percepts and the structure of mental images. For this reason guaranteeing a high level of similarity between imagery and vision provides a means for confirming the figural character of mental images. But there is more. Our hypothesis is that the activation of the primary visual areas in tasks using the imagination provides evidence for recognising a degree of «receptivity» towards images and, therefore, of a similarity between images and percepts which is not only structural but also functional. The figural character of images (the structural resemblance with percepts) and the degree of receptivity which characterises them are two sides of the same coin. It is with respect to this inherently dual nature that the role of images in the discovery of new information and, consequently, creativity come clearly to light.

3. Imagery and creativity

Grandin's introspective reflection with which we began this paper is not an isolated case. Writers, artists and scientists claim that visual imagination is involved in creative tasks and the process of discovery (Finke, 1990; Intons-Peterson, 1993; Miller, 1984, 2000). These reflections provide evidence of the role of mental visualisation processes in the discovery of new information. Even though it is true that the anecdotes themselves are not sufficient to confirm the functional role of images in cognitive processes, they do, however, certainly represent a starting point. It would be a grave error, in fact, not to recognise the importance of a basic question which these anecdotes leave unsolved, namely their justification (Shepard, 1978).

The debate on the role of the creativity of mental images has reopened the question of their figural nature. The problem to be explained is that whether the content transmitted by a mental image coincides with its interpretation (or rather, with the information necessary to generate that image), or whether, instead, the mental images (once generated in the short-term memory) carry a content which goes *beyond* that necessary for their generation.

Gilbert Ryle (1949) is a precursor of the idea of the coincidence of the content of a mental image with the interpreted information necessary to generate it. His idea is that the content of an image is in itself «transparent» to the subject, or rather that it is fully interpreted:

> It would be absurd for someone to say: "I vividly see something in my

> mind's eye, but I cannot make out even what sort of a thing is". True, I can see a face in my mind's eye and fail to put a name to its owner, just as I can have a tune in my head, the name of which I have forgotten. But I know how the tune goes and I know what sort of face I am picturing. Seeing the face in my mind's eye is one of the thing which my knowledge of face enables me to do (Ryle, 1949: 265-6).

In this way, the content of an image coincides with its interpretation, excluding any functional role in the way (the format) in which it is expressed. Pylyshyn (2003) considers the content of an image interpretable in terms of «tacit knowledge», or rather in reference to what people know «about how things tend to happen in the world» (ivi: 113). Along with this hypothesis, Chambers and Reisberg (1985) also claim that images are interpreted symbols. In this respect the «transparency» of the content to the subject who imagines it is particularly illuminating. It is not possible to have an image in mind without knowing what image it is. Coinciding with their interpretation, images cannot be reinterpreted:

> Images, unlike the stimuli that give rise to perception, cannot be reinterpreted. Images can certainly willed to be different, or can be replaced, but without a construal process, there is no possibility for reconstrual. (...) One regard in which images are not picture-like is that there is no such thing as an ambiguous image. "To imagine something differently is to imagine something different" (Casey, 1976: 159) (ivi).

In order to demonstrate the difference between images and percepts, Chambers and Reisberg invented a series of experiments which were able to prove the impossibility of reinterpreting a mental image. To understand the issue at stake, it is important to draw a clear distinction between the two cases described in the quote from Casey. The first is when we imagine something as different simply because we have produced another image; instead, the second is when we interpret the same image differently. It is the second case, of course, which needs to be demonstrated in order to speak of an analogy between imagery and vision.

In order to test the impossibility of reinterpreting the same image, the two researchers carried out the following experiment. The subjects were required to briefly observe an ambiguous figure like the famous duck-rabbit of Jastrow, reproduce it in a mental image and then «inspect» it with the mind's eye trying to produce a different interpretation. Only in the case where they were able to see the alternative interpretation would they have experienced the typical «reversal» which characterises the perception of ambiguous figures, thus showing the close relationship between vision and imagination. The results of the experimental reports were quite clear: none of the 35 subjects was able to reinterpret the mental image in any of the experimental tasks. In contrast with these results, all the subjects were able to draw the image which they had in mind and, starting from the visual perception of the drawing, discover the alternative interpretation, therefore demonstrating the radically heterogeneous nature of vision and imagination.

The conclusion to be drawn from these experiments is that the production of an image is entirely determined by the content that interprets it. Mental images cannot be «examined» both because they are not observable objects and, above all, because it is useless to examine a content which is already known. Seeing the situation is completely different: the sensorial stimulus (and therefore the real object) plays a double constitutive role in visual perception. In the first place, it acts as a limit to the interpretation: contrary to the creative freedom which characterises the imaginative act, the real object places strong limitations on possible interpretations. But it would be wrong to consider the stimulus in this way only. It is, in fact, also what feeds the interpretive process unceasingly. It is here that we find the real difference between perception and imagination: in vision, the stimulus situation is in a state of constant change, and this means that interpretive processes are continuously active. From a cognitive point of view what characterises perception (as receptivity) is the fact that it is possible to constantly add new information to contents which have already been interpreted. Only the perceived objects, differently from those which are imagined, can therefore be «reinterpreted». And it is on the crucial question of the reinterpretation of mental images that we find the crux of the problem, not only in the relationship between vision and imagination, but in the debate on the very nature of mental images.

The results produced by Chambers and Reisberg are not without unproblematical: their theory still has to explain the significant problem of the *extra* information which images produce in the process of discovery and creativity. Images are particularly useful in unusual or new situations, situations in which the information (verbal) in the memory is not enough to deal with a given task. If all the contents transmitted by an image were interpreted, the role of images would be inexplicable. For what purpose is it necessary to duplicate a content already available to the subject in a different format? After the numerous criticisms of their article in 1985, Reisberg and Chambers (1991) themselves outline the problem:

> We claimed that mental images are inherently unambiguous. Yet this poses a puzzle: It seems obvious that images can surprise us of instruct us. If images are unambiguous embodiments of ideas we already have, how can new information or insights be gleaned from images? (ivi: 336-7).

Although in the successive debate Reisberg and Chambers (1991, 1992; Chambers, 1993) attempted to solve the problem by attributing a dual nature (depictive and descriptive) to mental images, their solution appears heavily influenced by «interpretativistic prejudice». To deal with the problem, there appears to be only one choice: to explain the role of images in discovery and creativity it is necessary to consider mental images as being able to transmit more information than that available in the content necessary for their generation. In order to do this it is necessary to sustain the hypothesis of the strict continuity between vision and imagery. If imagining an object were a way of «seeing» the object in its absence, in a way which is subject to limitations similar to those imposed on visual perception by the real object, then mental images could also transmit additional information, thus justifying their role in processes of creativity and discovery. In order to do this, however, images, like percepts (although not exactly in the same way), should be reinterpretable. But is it really possible to characterise mental images in this way?

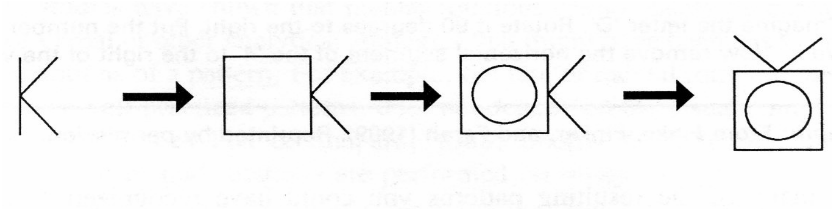

Figure 6. Reinterpreting mental images (Finke, 1989).

To demonstrate the fact that the reinterpretation of images is possible, Finke (1990, 1993) attempts to test if subjects were able to recognise the image of a familiar object without providing any interpretation in advance. The experiment, carried out by subjects with their eyes closed, consisted of the generation and transformation of images following a sequence described by the researcher. A typical sequence was the following: «Imagine the letter 'K'. Place a square next to it on the left side. Put a circle inside of the square. Now rotate the figure 90 degrees to the left» (Finke et al., 1989: 61). After the transformations, the subjects were required to say if they succeeded in recognising the object in the final image (in this case a television). The results proved that many subjects were able to represent the model described by the researcher in images, to carry out transformations on it in accordance with the instructions given and to reinterpret the emerging image from the transformations of the preceding model. As it was impossible to know the content of the final image in advance (since it depended on the sequence of transformations), the result is that the subjects succeeded in recognising the objects only by reinterpreting their mental images (attributing a new content to them). What does this additional information depend on? The most interesting answer from the research of Finke et al. is that the reinterpretation depends on the structural properties of the images (those that provide them with a figural character); their experiments are enlightening in this respect: there is nothing in the interpretation of the letter "J" and the letter "D" which leads one to think of an umbrella; the fact that by composing the images of the two letters in a certain way it is possible to reinterpret the configuration in the mind as an umbrella shows that «it is the geometry of the pattern that allows these inferences to be made. Since these inferences *can* be made, subjects must have more than the pure symbolic or conceptual residue of these visual patterns available to them» (Finke et al., 1989: 70).

Experiments of this type have been confirmed by numerous studies on the role of mental images in creativity and discovery (Finke, 1990, 1993; Finke et al., 1992; Roskos-Ewoldsen, 1993; Helstrup and Anderson, 1991, 1993). To explain this role it is essential to draw a distinction between «generative processes» and «explorative processes» (Finke, 1993). And it is this very distinction which can provide a solution to the «paradox of the functional role of mental images» from which we started. In order to generate an image it is always necessary to know what to generate: it is not, therefore, to the generation processes (which employ interpreted content) that we must look to explain the role of imagery in creativity. The additional information may be ascribed to processes of image exploration (the same activated in the interpretation of the visual stimulus). But why should the exploration of an image generated from an interpreted content (fully «transparent» to the subject) open the way to the discovery of new information? The only plausible answer to this question is that the added information which an image transmits depends on the representational format. Constructing an image means codifying the mental content in a certain shape; this operation is never neutral with respect to the represented content. Because information is dependent on structure (the format of the representation), the information transmitted by an image never coincides with that of the interpreted content which is necessary for the generation of that image. To speak of representation format in this case means to speak of the shared structure of images and percepts. From this point of view, exploration processes represent a way of perceiving an object in its absence.

> According to this view, mental images can stimulate visual processing mechanisms directly. Thus, when mental images are formed, these mechanisms would respond in much the same way as they do when objects and events are observed, resulting in the sensation that an image can be "seen" as if it were an actual object or event (…). The types of effects that occur when a particular image is formed would be determined by which visual mechanisms the image activate (Finke, 1980: 130).

Up to this point we have outlined the theory of Finke as an hypothesis in favour of the reinterpretability of mental images. At this point, however, it is necessary to draw a distinction which goes deeper. According to some authors (Peterson et al., 1992; Verstijnen, 1997) Finke's experiments are convincing, but they fail to reach the same level of processing as in the ambiguous figures of Chambers and Reisberg's experiments. Peterson et al. (1992), for instance, claim that the ambiguous figures lead to both «reconstrual» and «reference-frame realignment»: in cases of this type the reinterpretation goes through a true reorganisation of the structure of the stimulus to be interpreted. Finke's figures, on the other hand, are only «reconstrual». Verstijnen (1997) also claims that it is necessary to distinguish two processes. The first concerns the capacity to break up the starting structure in order to identify new components in a structure. This is a case of «restructuring», a process considered «analytical» by Verstijnen, easy in visual perception but very difficult to reproduce in imagery, and the reason for which mental images are not reinterpretable. Finke's experiments fail to reach this level of processing: they stop at what Verstijnen defines «figural combination» a synthetic, not analytic process. Verstijnen's conclusion is that:

> the opposing conclusions from the imagery literature can be reconciled on the assumption that two forms of processing in imagery have to be distinguished, viz. ANALYSIS and SYNTHESIS. The two different experimental paradigms, familiar from literature, involve analytic and synthetic processes to different degrees (ivi: 16).

The issue is still open and requires further investigation. Having discarded the false problem of seeking a «perfect match» between imagery and perception (which remain distinct faculties), it is up to future research to explain more fully the degree and type of mechanisms involved in these types of process. Whatever the appropriate type of analysis may be, however, one thing is clear: since the content transmitted by mental images does not coincide with that necessary for their generation, an interpretativistic position seems untenable. This result opens the way to the idea that the role of mental images in creativity may be ascribed to the relative similarity between imagery and vision.

Conclusion

In this paper the nature of imagery has been explored from the point of view of the relationship between vision and imagery. This relationship forms the basis of the specific cognitive role of mental images: they are not always necessary in thought but only in specific situations – in those situations in which the ideal way to solve a problem would be the perception of those same things before our own eyes. In these cases the images become a good substitute for reality: this is possible because they share common structural elements with percepts. These shared common characteristics provide a solution to the paradox of the functional role of mental images in cognition: the role of imagery in creativity depends on the structural properties which mental images share with the representation of the perceived information. The production of a mental image is, from this point of view, an analogous process to the perception of an object in its absence. Such a result opens the way to the idea that there is a close link between imagery and vision and, thus, provide justification for the theory that imagery is closer to perceiving than to thinking.

Endnotes

[1] Expert in animal behaviour and famous for her splendid autobiographical books describing first hand experience of autism.

References

Bértolo, H. (2005). Visual Imagery without Visual Perception? *Psicològica*, 26, 173-88.

Bértolo, H., Paiva, T., Pessoa L., Mestre, T., Marques, R. & Santos, R. (2003). Visual Dream Content, Graphical Representation and EEG Alpha Ac-

tivity in Congenitally Blind Subjects. *Cognitive Brain Research*, 15, 277-84.
Casey, E. (1976). *Imagining: A Phenomenological Study*. Bloomington, Indiana, Indiana University Press.
Chambers, D. & Reisberg, D. (1985). Can Mental Images Be Ambiguous? *Journal of Experimental Psychology: Human Perception and Performance*, 11, 317-328.
Chambers, D. (1993). *Images Are Both Depictive and Descriptive*. In B. Roskos-Ewoldsen et al. (Eds.) (1993), 77-97.
Cornoldi, C., Rocchi, P. & De Beni, R. (1992). *Memoria e uso di immagini mentali in ciechi congeniti totali*. In Dario Galati (a cura di), *Vedere con la mente*, Milano, Franco Angeli, 1993, 63-189.
Denis, M. (1991). *Image & Cognition*. New York, Harvester-Wheatsheaf.
Farah M.J. (1988). Is Visual Imagery really Visual? Overlooked Evidence from Neuropsychology. *Psychological Review*, 95, 307-17.
Finke, R.A. (1989). *Principles of Mental Imagery*. Cambridge (Mass), The MIT Press.
Finke, R.A. (1990). *Creative Imagery. Discoveries and Inventions in Visualization*. Hillsdale, New Jersey, Lawrence Erlbaum Associates Pubblishers.
Finke, R.A. (1993). *Mental Imagery and Creative Discovery*. In B. Roskos-Ewoldsen et al. (Eds.) (1993), 255-285.
Finke, R.A., Pinker S. & Farah M.J. (1989). Reinterpreting Visual Patterns in Mental Imagery. *Cognitive Science*, 13, 51-78.
Finke, R.A., Ward T.B. & Smith S.M. (1992). *Creative Cognition. Theory, Research, and Application*. A Bradford Book, Cambridge (Mass.), The MIT Press.
Fodor, J.A. (1984), Perception Reconsidered. *Philosophy of Science*, 51, 1984, 23-43.
Goodman, N. (1968). *Languages of Art*. Indianapolis, Bobbs-Merrill.
Grandin, T. (1996). *Thinking in Pictures*. New York, Vintage Books.
Gregory, R. L. (1966). *Eye and Brain. The Psychology of Seeing*. London, Weidenfeld & Nicholson.
Helstrup, T. & Anderson, R.E. (1991). *Imagery in Mental Construction and Decomposition Tasks*. In R.H. Logie & M. Denis (Eds.) (1991), 229-240.
Helstrup, T. & Anderson, R.E. (1993) *Multiple Perspectives on Discovery and Creativity in Mind and on Paper*. In B. Roskos-Ewoldsen et al. (Eds.) (1993), 223-253.
Intons-Peterson, M.J. (1993) *Imagery's Role in Creativity and Discovery*. In B. Roskos-Ewoldsen et al. (Eds.) (1993), 1-37.
Ishai, A. & Sagi, D. (1995). *Common Mechanisms of Visul Imagery and Perception. Science*, 268, 1772-74.
Kanizsa, G. (1980). *La grammatica del vedere*. Bologna, Il Mulino.
Kanizsa, G. (1991). *Vedere e pensare*. Bologna, Il Mulino.

Kaski, D. (2002). Revision: Is Visual Perception a Requisite for Visual Imagery? *Perception*, 31, 717-731.

Kennedy, J.M. (1980). Blind People Recognizing and Making Haptic Pictures. In Hagen M.A. (ed), *The Perception of Pictures*. Academic Press, New York, 1980.

Kennedy, J.M. (1983). What Can We Learn About Pictures from the Blind? *American Scientist*, 71, 19-26.

Kennedy, J.M. (1993). *Drawing and the Blind*. Yale University Press, New Haven.

Kerr, N.H. (1983). The role of Vision in "Visual Imagery" experiments: Evidence from the Congenitally Blind. *Journal of Experimental Psychology*: General, 112, 265-277.

Kerr, N.H. (2000). Dreaming, Imagery and Perception. In Kryger M.H. et al. (eds.), *Principles and Practice of Sleep Medicine*, W.B. Saunders Company, 482-90.

Klein, I. et al. (2004). Retinotopic Organization of Visual Mental Images as Revealed by Functional Magnetic Resonance Imaging. *Cognitive Brain Research*, 22, 26-31.

Kosslyn, S.M. (1994). *Image and Brain. The Resolution of the Imagery Debate*. Cambridge (Mass.), A Bradford Book, The MIT Press.

Kosslyn, S.M. & Ochsner K.N. (1994). In Search of Occipital Activation During Visual Mental Imagery. *TINS*, 17, 290-92.

Kosslyn, S.M., Ganis, G. & Thompson, W.L. (2001). Neural Foundations of Imagery. *Nature Reviews*, 2, 635-642.

Kuhn, T.S. (1962). *The Structure of Scientific Revolutions*. The University of Chicago, Chicago.

Lambert, S., Sampaio, E., Mauss, Y. & Scheiber, C. (2004). Blindness and Brain Plasticity: Contribution of Mental Imagery? An fMRI Study. *Cognitive Brain Research*, 20, 1-11.

Landau, B., Smith, L. & Jones, S. (1998). Object Perception and Object Naming in Early Development. *Trends in Cognitive Sciences*, 2, 19-24.

Lopes da Silva, F.H. (2003). Visual Dreams in the Congenitally Blind? *Trends in Cognitive Sciences*, 7, 328-330.

Mandler, J.M. (2004a). Thought before Language. *Trends in Cognitive Sciences*, 8, 508-513.

Mandler, J.M. (2004b). *The Foundations of Mind*. Oxford University Press, Oxford.

Marr D. (1982). *Vision*. San Francisco, Freeman.

Masini, R. & Antonietti, A. (1992). Processi percettivi e rappresentativi nei non vedenti. In Dario Galati (a cura di) *Vedere con la mente*. Milano, Franco Angeli, 1992, 113-162.

Miller, A.I. (1984). *Imagery in Scientific Thought: Creating 20th Century Physics.* Boston, Birkhäuser.

Miller, A.I. (2000). *Insights of Genius: Imagery and Creativity in Science and Art*. Cambridge, MA, The MIT Press.
Miyashita, Y. (1995). How the Brain Creates Imagery: Projection to Primary Visual Cortex. *Science*, 268, 1719-20.
Peterson, M.A., Kihlstrom, J.F., Rose, P.M., & Glisky, M.L. (1992). Mental Images Can Be Ambiguous: Reconstruals and Reference-Frame Reversals. *Memory and Cognition*, 20, 107-123.
Pylyshyn, Z.W. (1981). The Imagery Debate. Analogue Media versus Tacit Knowledge. *Psychological Review*, 88, 16-45.
Pylyshyn, Z.W. (2003). Return of the Mental Image: Are There Really Pictures in the Brain? *Trends in Cognitive Sciences*, 7, 113-18.
Quinn, P. & Eimas, P.D. (1986). Pattern-line Effects and Units of Visual Processing in Infants. *Infant Behavior and Development*, 9, 57-70.
Reisberg, D. & Chambers, D. (1991). Neither Pictures nor Propositions: What We Can Learn from a Mental Image? *Canadian Journal of Psychology*, 45, 336-352.
Reisberg, D. & Chambers, D. (1992). What an Image Depicts Depends on What an Image Means. *Cognitive Psychology*, 24, 145-174.
Rock, I. (1982). *The Logic of Perception*, The MIT Press, Cambridge, MA.
Roland, P.E. & Gulyàs B. (1994). Visual Representations of Scenes and Objects: Retinotopical or Non-Retinotopical. *TINS*, 17, 294-97.
Roskos-Ewoldsen, B. (1993). *Discovering Emergent Properties of Images*. In B. Roskos-Ewoldsen et al. (Eds.) (1993), 187-221.
Roskos-Ewoldsen, B., Intons-Peterson, M.J. & Anderson, R.E. (Eds.) (1993). *Imagery, Creativity, and Discovery. A Cognitive Perspective*. North-Holland, Elsevier Science Publishers B.V.
Ryle, G. (1949). *The concept of Mind*. Chicago. The University of Chicago Press, 1984.
Sakay, K. & Miyashita, Y. (1994). Visual Imagery: An Interaction Between Memory Retrieval and Focal Attention. *TINS*, 17, 287-89.
Shepard, R.N. (1978a). The Mental Image. *American Psychologist*, 33, 125-37.
Verstijnen, I. (1997). *Sketches of Creative Discovery*. Thesis Technische Universiteit Delft.

Francesco Ferretti
Università di Roma Tre
ferretti.f@tin.it

James Genone
University of California, Berkeley

Concepts and Imagery in Episodic Memory

Abstract The relationship between perceptual experience and memory can seem to pose a challenge for conceptualism, the thesis that perceptual experiences require the actualization of conceptual capacities. Since subjects can recall features of past experiences for which they lacked corresponding concepts at the time of the original experience, it would seem that a subject's conceptual capacities do not impose a limit on what he or she can experience perceptually. But this conclusion ignores the fact that concepts can be composed of other simpler concepts that a subject possessed earlier, and that demonstrative capacities can explain how a subject can experience a particular feature of her environment, even when she lacks a fully general concept for that feature. Using these resources, conceptualism can explain the relation between perceptual experience and memory. Nevertheless, a puzzle remains for the defender of conceptualism. A certain view about the relation between perceptual experience and mental imagery in episodic memory – that imagery in recall matches the experience retained in it – can make it difficult to understand how conceptualism could be true. For if a subject's conceptual capacities determine what the phenomenology of an experience (or memory of it) is like, then one would expect a perceptual experience and its recall in memory to differ in phenomenology if they involve different concepts. In this essay, I solve this puzzle for conceptualism by undermining the assumption that there is a match between imagery in episodic memory and the phenomenal character of experience.

Keywords Concepts, experience, imagery, memory, perception, phenomenology.

1. Concepts, Memory and Imagery

Conceptualism about perceptual experience is the view that what can be experienced perceptually is constrained by a subject's conceptual capacities. While a variety of formulations of this view exist, the key claim for the view is that a subject cannot experience her environment as containing a particular object or property unless she possesses an associated concept under which the object or property falls, and which is actualized in her having the experience. While arguments both for and against conceptualism have received a great deal of attention in recent literature on perceptual experience, few have paid attention to how the relation between perceptual experience and memory bears on the debate about the role of concepts in experience. An important exception is Michael Martin, who in 'Perception, Concepts, and Memory' (Martin 1992) argues that reflecting on how newly learned concepts can be applied to perceptual experiences that have been preserved in memory shows that conceptualism is false. I have elsewhere argued that a conceptualist can acknowledge Martin's account of the links between perceptual experience and memory without accepting his conclusion (Genone ms.). There I claim that by appealing to the fact that some concepts are composed of simpler ones that the subject possessed at an earlier time, and that others have their basis in more primitive demonstrative abilities, a conceptualist view can be sustained in the face of Martin's challenge.

After reviewing these issues in order to introduce a conceptualist account of the relation between perceptual experience and memory, I consider a separate but related problem facing conceptualism. Some philosophers have held that in order for a memory to involve experiential recall of a past perceptual experience the imagery accompanying the memory must match the remembered experience (Locke 1690/1975, Russell 1921).[1] But given that the concepts involved in having a particular experience in part determine the phenomenology of the experience, a conceptualist view can seem to undermine the possibility of a match. For if different concepts are involved in the initial experience and its later recall, this would seem to suggest that the two states will differ in phenomenology. If the matching view is correct, then conceptualism would seem to be falsified.

The nature of imagery in memory is a topic that has also been explored by Martin (Martin 2001). In solving the puzzle of how a conceptualist can account for differences in phenomenology between a perceptual experience and the imagery involved in its being recalled, I will explore Martin's views about memory imagery, as well as those of John Campbell put forward in his book *Reference and Consciousness* (Campbell 2002). Drawing on their insights, I will suggest that by rejecting the view that there must be a match in phenomenology between a perceptual experience and its recall, conceptualism can give the correct account of the connection between perception and memory.

2. Memory and Conceptualism

What precisely does it mean to claim that perceptual experiences are conceptual? Many theorists have supposed that what matters here is what sort of *content* we attribute to perceptual experiences. Some philosophers have claimed that perceptual experiences are propositional attitudes, and hence have propositional contents (Brewer 1999, McDowell 1994, Searle 1983). Since propositions are thought by these philosophers to have concepts as their constituents, such contents will be by definition conceptual. The question of whether or not perceptual experiences are propositional attitudes, or even representational at all, however, is not one that needs to be taken for granted in formulating conceptualism. We can instead say that a perceptual experience is conceptual if and only if having that experience requires the possession and actualization of the same conceptual capacities required for having an associated belief. So if perceptual experiences are conceptual, then perceiving a cup on a table will require a subject to possess and actualize the same conceptual capacities required for having a belief about a cup on a table.[2]

If this is right, then we can ask whether we have reason to think perceptual experiences *are* conceptual. If we can find evidence that one can have a perceptual experience of a particular object or property (under a certain aspect or description) while lacking the conceptual resources to have a belief about that object or property (under the same aspect or description), then we will have reason to reject the view that perceptual experiences are conceptual. According to Martin (1992), reflecting on the relationship between perceptual experience and memory provides us with such evidence. Martin claims that it is possible to have experiences of objects or properties even though one lacks an associated concept. To see this, we can imagine a case in which a subject experiences an object or property without noticing it. If we suppose that the subject in the meantime gains a concept that she lacked at the time of the original experience, and employs it in recalling the experience, then this would suggest that she did not need to possess the concept to have had an experience of the object or property in the first place. For example, when I was young, I didn't pay much attention to people's appearances, and although I lacked the concept of a dimple, it is highly likely that I had many experiences of people with dimpled cheeks. It is not hard to imagine that in recalling the appearance of someone from my childhood, I remember the person as having dimpled cheeks, a concept I have since learned. This would seem to suggest that my experience was not limited by the concepts I possessed at the time. Although I lacked the concept of a dimple, this person must have appeared to me as having dimples if I am later able to remember him or her that way. If this example and others like it are legitimate, they would seem to constitute counterexamples to the thesis of conceptualism.

One strategy for combating this challenge would be to deny the legitimacy of the phenomenon Martin calls attention to. One might think that if a subject did not initially notice the object or property that is supposedly later recalled, then it was not really experienced at all. We could then explain the purported memory as a case of false memory – we might think in the above example that I seem to remember people from my past as having dimples only because I now know them to have them. I think this is the wrong strategy for the conceptualist. First of all, while one might try to undermine purported counterexamples on a case by case basis, Martin only needs to convincingly describe one instance of the phenomenon in order to advance his argument against conceptualism. Moreover, it is often the case that one can legitimately recall having seen or heard something that one failed to notice at the time of the experience. For example, while searching the living room for my keys I might realize that I just saw them in the bedroom without in my haste having noticed them. There is no *prima facie* reason to suppose that a case like this could not involve a subject having gained a concept that she could employ in recalling the experience at a later time.

The correct strategy for responding to the phenomenon that Martin draws our attention to is to grant its possibility, but argue nevertheless that it does not threaten conceptualism. In order to do this, it is necessary to affirm that the subject did not experience anything for which she lacked conceptual resources. There are two kinds of cases here. To return to the above example, a conceptualist could argue that if the person's dimples really were part of my experience, then I must have conceptualized them on the basis of geometric concepts I had available to me at the time. Assuming for the sake of argument that the concept of a dimple has the concept of an indentation as one of its components, then we can see how the feature later conceptualized in memory with the concept of a dimple could have been part of my earlier experience. Supposing I have since gained the more complex concept, we can imagine that by recognizing its constituent in the earlier experience when I recall it, I can apply the concept of dimpled cheeks to the person I remember.[3]

A second sort of case concerns a memory involving a non-composite concept that was lacked by the subject at the time of the original experience. Color concepts provide a convenient example. If we suppose that I lacked the concept of hazel when I was young, and we suppose that we cannot understand the concept of hazel simply as a composite of brown and green, then it will not be possible to explain my remembering people from my youth as having hazel eyes in terms of simpler concepts I possessed at the time. Another strategy is available to the conceptualist, however, and that is to claim that I possessed a demonstrative capacity to attend to hazel colored objects, which I could have employed by thinking of them or referring to them as being *that* shade. If I had not possessed such a capacity, the conceptualist can insist, then it is not correct to say that I could have had an experience of the property in the first place.[4]

So according to the line of defense suggested here, a conceptualist can maintain that subjects can recall experiences on the basis of concepts they initially lacked, but only if other appropriate conceptual capacities were available to them. There are, of course, further considerations on behalf of both the conceptualist and the nonconceptualist that can be raised in this context. I have examined these issues in detail elsewhere (Genone ms.). Assuming for now that this version of conceptualism can withstand Martin's challenge, I want to consider how it leads to a related difficulty concerning the relation between perceptual experience and memory – one that a conceptualist needs to be able to explain in order for the view to be sustained.

3. Concepts and Phenomenal Character

According to the defense of conceptualism I suggested in the previous section, a conceptualist should maintain that a subject can employ different concepts in recalling a perceptual experience than were employed at the original time of the experience. While this might make sense of the challenge presented by Martin, it leads to further difficulties for conceptualism. To see what these difficulties are, it is first necessary to explore the relation between concepts and the phenomenology of perceptual experience and memory. A natural assumption for a conceptualist view about experience is that there is a close connection between the concepts that are involved in a subject having a particular experience and what that experience is like for her.[5] It is because a subject has the concept of an automobile that she can see that the large mass of metal and plastic before her is an automobile, rather than just a mass of metal and plastic. Likewise, it is because her experience presents her with a particular mass of metal and plastic that the concept of an automobile is relevant to her experience. According to the view presented in the previous section, a perceptual experience and its recall in memory can involve different concepts if the subject has in the interim gained a new concept. Although I argued that there must be a connection between the concepts applied to an experience that is recalled in memory and the concepts involved in the original experience, it nevertheless follows that the fact that a new concept is applied to the experience in memory will involve a phenomenal difference between the experience and its recall.

To see this, we can, returning to the above example, suppose that on first being confronted with a person who has dimpled cheeks, I can experience the dimples on the basis of my concept of an indentation. Later, having gained the concept of dimples, I recall the person as having dimpled cheeks. It is natural to think that there is a phenomenal difference between these two experiences, for if they were exactly the same, there would be no reason to suppose that the concept of dimples would be involved in my memory of the person. The difference consists in whatever else is involved in seeing someone as having dimples rather than just seeing the person as having indentations in his or her cheeks. Although the concept of an indentation is a component of the concept of a dimple, the two cannot be equivalent, otherwise there would be no reason to suppose that I didn't have the concept of a dimple in the first place. What this difference amounts to, phenomenologically speaking, is at the very least a difference in the salience or relevance of various aspects of what one experiences. For example, a person with dimples might look jolly to me, but someone who I see as having indented cheeks would probably not look that way.

It might be objected that what is really going on is that there is a common experiential element present in both the perceptual experience and the memory of it, but that this common element has a different significance for me given the change in my conceptual resources. This would ignore, however, the ways in which a dimple *looks* different than a mere indentation. To suppose that there is no *visual* difference between the two would undermine the basis for the application of a different concept in each case. So the conceptualist is committed to the idea that insofar as different concepts are involved in a perceptual experience and its recall, the memory and the experience will differ phenomenologically.

To see how this creates a problem for conceptualism, it is necessary to make some distinctions with respect to the kinds of memories that might be relevant here. Among species of conscious memory, philosophers and psychologists have distinguished between procedural, factual, and episodic memory (Sutton 2004).[6] Procedural memory – remembering *how* to do something – is not relevant for present purposes. Factual memory, sometimes called semantic or propositional memory, involves remembering *that* something happened. Episodic memory, sometimes called personal memory, involves remembering a particular past episode that one experienced. One can have a factual memory without having experienced what is thereby remembered. For example, I can remember that yesterday was my sister's birthday, without having had a perceptual experience of her birthday that I am remembering.[7] On the other hand, episodic memory *does* require that I experienced what I now remember. I cannot remember the way my sister looked on her birthday if I did not see her on that day.

From the point of view of thinking about links between concepts and the phenomenology of memory, we can note that both factual and episodic memory can involve imagery, and that as suggested above, it seems natural to suppose there is a tight link between the imagery and a subject's conceptual capacities. For example, it would not seem to make sense to suppose that a subject could remember that aardvarks have teeth while imagining what one looked like with its mouth open if she lacked the associated concepts. Likewise, it seems highly plausible that the imagery involved in remembering seeing an automobile will differ depending on whether or not one possesses the concept of an automobile – for if one lacks the concept, then one will just remember it as, say, a large mass of metal and plastic. As suggested above, we should think of this as a phenomenological difference in order to make sense of why the subject would apply a different concept in the two cases.

In the present context, what needs to be examined is whether or not conceptualism, at least as I have proposed it should be developed, creates a problem for the relationship between the phenomenology of perceptual experience and that of memory by claiming that a perceptual experience and its later recall can involve different concepts. Some philosophers have held that there is a constitutive link between memory and perceptual experience that is preserved by imagery (Locke 1690/1975, Russell 1921).[8] On this view, part of what constitutes a particular memory as being an episodic one is that it matches the experience it recalls. I will call this view "the matching view". It is easy to see that the matching view poses no problem for factual memory. Although factual memory can involve imagery, as for example when I imagine my sister while remembering that her birthday was yesterday, there is no link between any particular perceptual experience I may have had and my memory, so there is no reason to think that the imagery involved in the memory should be true to any particular experience of my sister.[9] According to the matching view, however, the same is not the case for episodic memory. If I remember what my sister looked like when she blew out the candles on her birthday, my experience of seeing her would seem to be constitutive of my memory being a memory of that particular episode as opposed to some other one.

To see why this is a problem for conceptualism, imagine that I see my sister on her birthday with her hair done up in a certain style, one that I have never encountered before. Later, being bored at the hair salon while waiting for my wife, I might study the pages of a hairstyle magazine and gain lots of new concepts for hairstyles. Then upon being queried by my sister as to whether I noticed her hairstyle on her birthday, we can imagine that I recall what she looked like and remember her as having done her hair in a twisted bun, the concept of which I acquired through my reading at the hair salon. The question for the conceptualist is whether my gaining and actualizing this new concept influences the imagery involved in my memory of seeing my sister's hairstyle. Given what has been said above about the relation between concepts and imagery, it would seem that we must conclude that it does. Although my original experience, according to conceptualism, will have involved various geometric and spatial concepts, perhaps some of which were demonstrative, these won't be equivalent to the concept of a twisted bun, the concept I have since acquired. Hence, it follows that what it was like for me to see her hair on her birthday, and what it is like for me to recall how it looked are phenomenologically different. If this is correct, however, it violates the relationship the matching view maintains as holding between a perceptual experience and the imagery involved in episodic memory. If the imagery doesn't match the original experience, then the constitutive link that, according to the matching view, makes the memory a memory of a particular episode seems to have been broken. How can conceptualism deal with this difficulty?

4. Imagery and Episodic Memory

We have seen that the matching view, if it were correct, would pose a significant problem for conceptualism. In order for conceptualism to maintain its coherence, it is necessary to find a way of responding to this challenge. What this requires is explaining how an episodic memory can be a memory of a particular past experience without assuming there to be matching phenomenology. Although the matching view can seem plausible, it has not gone unquestioned in contemporary philosophy.[10] In 'Out of the Past: Episodic Recall as Retained Acquaintance', Mike Martin aims to undermine precisely this view (Martin 2001).[11] According to Martin, the view that a perceptual experience and its recall in episodic memory must be phenomenologically alike poses a significant challenge for the idea that episodic memory involves recall of a past experience. This is because part of what it is to remember something one has previously experienced is to relate to it as something that happened *in the past*. If the memory and the recalled experience are phenomenally alike, however, it is difficult to see what the *pastness* of the memory will consist in. After all, recalling a past experience is not the same as experiencing something that happened in the past again (as seems to happen when we experience *déjà vu*). As Martin writes, "If episodic memory is to be the experience of the past … we need to have the experience of past events as being past" (Martin 2001: 268).

An obvious response to this difficulty might seem to be just to reject any connection between memory imagery and the phenomenology of the experience recalled. If this connection is broken, however, what seems to distinguish episodic memory from factual memory, namely that it is partly constituted by the experience it is a memory of, seems to be lost. Martin formulates this problem as a dilemma:

> We seem to be faced with a choice: either we insist on the idea of episodic memory as retained apprehension or experience, in which case we can have no distinctive experience of the past as past; or we insist on the idea that the episodic memory has a distinctive phenomenology associated with the past, but thereby give up the idea that this has anything to do with retaining something form earlier experience (Martin, 2001: 269).

According to Martin, the solution to this problem is to realize that the connection between episodic memory and the remembered experience is representational rather than phenomenological. According to this line of thought, while a perceptual experience and an episodic memory of it will have the same object (i.e., they will both be *of* the same thing), they relate a subject to the object in different ways:

> The idea here is that although perceptual experience and imagery may coincide with respect to the objects of experience, the events or qualities which are present to the mind, they will still differ in the manner by which these objects are given or presented to the mind. In general, perceptual experience allows for the presentation of objects and qualities, where imagery allows only for the re-presentation of such things (Martin, 2001: 271).

Just as the imagery involved in imagination in general involves representing the way something might be experienced, the imagery in episodic memory is here taken to be constituted by representing a past experience. For example, recalling the way my sister looked while blowing out the candles on her birthday involves representing that experience to myself. Doing this in no way requires that the imagery that constitutes my memory match the phenomenology of the original experience. The moral of these considerations, according to Martin, is that "imagination and memory relate to perception not through replicating the sensational or imagistic component of perception, but through being a form of representing such experiential encounter with the world" (Martin 2001: 273-74).

Although Martin's suggestion allows us to reject the view that the imagery involved in episodic memory and the phenomenology of the experience recalled must match, the connection is not severed entirely. By posing a representational, or intentional connection between what was experienced and its later recall, Martin allows us to see how an episodic memory can still be tied to a particular perceptual episode.[12] This can account both for the pastness of a memory as opposed to the original experience, and for the discrepancy that we would expect in phenomenology given the conceptualist account of differences in conceptual capacities that can be exercised in a perceptual experience and its recall in episodic memory.

John Campbell offers further support for the idea that the phenomenology of perceptual experience and the imagery involved in memory need not be alike (Campbell 2002). In characterizing how demonstrative reference is possible via memory, Campbell points out that matching phenomenology is neither necessary nor sufficient for connecting an experience with its recall. Against sufficiency, he describes a situation in which someone tries to jog someone else's memory by describing an object from her childhood. In having the object described to her, the subject may come to visualize the object in a way that perfectly matches an experience she had of it in the past. Until she actually recognizes what she is imagining as the object of her past experience (Campbell describes this as the "Aha" moment), her imagining will not count as a case of memory.

Campbell also argues that it cannot be essential that the imagery that figures in memory match a particular past experience. If it were, it would be impossible for a subject to use the same demonstrative expression to refer to an object remembered on the basis of several previous experiences. Instead, Campbell suggests that all that is required is that past experience provides the subject with a referent for her use of the demonstrative. What enables a subject to use a demonstrative expression to refer to something she experienced in the past is what Campbell calls "deep decentering" – imagining oneself in the position of the past experience, and then treating it as if it were the present: "To understand the memory demonstrative is to simulate the time at which a past perceptual demonstrative could have been used" (Campbell 2002: 187). This is akin to Martin's suggestion that the relation between an episodic memory and the experience it recalls is a representational one.[13] Considering a case in which a demonstrative expression is used in memory to refer to an object experienced on different occasions, Campbell writes:

> […] your grasp of the memory demonstrative will depend on its being true that there is just one object from which your current memory derives. And that in general will not be something that can be guaranteed by the contents of your memory images or perceptual images alone – they could be exactly the same whether they derived from one object, a number of objects or no objects at all (Campbell, 2002: 191).

The upshot of these considerations, along with those offered by Martin, is that we have reason to reject the view that imagery in memory must match the phenomenology of the recalled experience. Reflecting both on the need to account for the pastness of memory, as well as our ability to refer demonstratively to remembered objects provides evidence that the relation between the imagery in memory and the recalled object of experience is an intentional one. With these lessons in mind, we can return to the conceptualist view offered above to see how it avoids the difficulties that seemed to face it in accommodating the relationship between perceptual experience and memory.

5. Conceptualism and Memory

The version of conceptualism defended here claims that a subject's perceptual experience and its recall in memory have the phenomenology they do thanks in part to the conceptual capacities the subject possesses. This connection implies that if a perceptual experience and its recall involve the actualization of different conceptual capacities, the memory will differ phenomenologically from the original experience. While this initially seems to threaten the constitutive connection between an episodic memory and the experience it is a memory of, our examination of the views of Campbell and Martin gives us reason to doubt that matching phenomenology could have secured this connection in the first place. As they point out, a match between memory imagery and experience is not only inessential to the relation between perception and memory, but supposing that there is such a match actually generates puzzles about the possibility of demonstrative reference via memory and about the pastness of memory.

By rejecting the matching view of the relation between the phenomenology of perceptual experience and memory, conceptualism can maintain that it is possible to remember an experience on the basis of concepts one lacked at the time of the original experience, provided that the newly gained concepts are appropriately linked to concepts that were originally available to the subject. By requiring this link between the concepts involved in a perceptual experience and a memory of it, conceptualism preserves the idea, suggested by Campbell and Martin, that there must be an intentional relation between perception and episodic memory. Since, according to conceptualism, the involvement of a newly learned concept in memory is grounded in the involvement of a more primitive but nevertheless related conceptual capacity in the original experience, a conceptualist view can insure that the memory is of the same feature of the world that was initially experienced by the subject. In light of the defense of conceptualism provided here, we see that respecting the relationship between perceptual experience and memory presents no difficulties that a conceptualist view cannot overcome.[14]

Endnotes

[1] For purposes of this paper, I treat the term "imagery" as a way of talking about the experiential, or phenomenological, element involved in non-perceptual experiences.

[2] These formulations are deliberately vague. I have done this so as to accommodate the wide variety of views that exist on the nature of perceptual experiences and intentional mental states in general. Adopting one or another of these views is inessential to the present line of inquiry, and the claims throughout this paper can be reformulated in terms of whatever view about the nature of intentionality one prefers. For example, if one is skeptical of propositions and thinks that mental states have some other kind of representational contents, then a mental state will be conceptual when being in a state with a certain content will require possessing and actualizing the concepts required for having a belief with the same content. The possibility of framing the debate about the role of concepts in perceptual experience in terms of conceptual states rather than contents was first pointed out by Heck (2000), and has since been further explored by Byrne (2004), and Speaks (2005), who both suggest that the primary interest of the debate concerns issues which can be separated from disagreements over the nature of intentionality.

[3] To respond in this way is not to allow that there is some common experiential element between the experience and the memory of it that is conceptualized in two different ways on separate occasions. Rather, the conceptualist claims that the concepts applied in the original experience make it possible both for the su-

bject to have the experience, and for related concepts to be applied in remembering it. I am grateful to John Campbell for pressing me to clarify this point.

[4] This way of envisioning the relationship between demonstrative capacities and conscious attention is not uncontroversial. I say a bit more about this elsewhere (Genone ms.). For an alternative account of the relation between demonstrative capacities and attention, see Campbell (2002).

[5] This assumption in no way entails a commitment to the view held by Dretske, Harman, and Tye that the qualitative aspects of an experience can be reduced to representational properties (Dretske 1995, Harman 1990, Tye 1997).

[6] Locutions such as "I remember John" or "I remember London" might seem to present a fourth kind of conscious memory – memory of a person or place. On reflection, however, it would seem that examples like this are shorthand for either factual or episodic memories, which might be glossed, for example as meaning that the subject remembers that there is someone she knows named "John" or remembers the experience of meeting him.

[7] Although it is no problem if I have had one – it will just be causally rather than constitutively relevant to my factual memory of her birthday.

[8] This view has its origin in the empiricist "theory of ideas" which supposes memories to be copies of perceptual impressions. Russell's view, which rejects the suggestion that the objects of memory are ideas, nevertheless assumes that the matching of a memory image to an original experience plays a key role in explaining how the memory can be of a particular experience.

[9] Arguably, we should suppose that the imagery must match some experience I have had of my sister, or perhaps a range of such experiences. If not, it might be unclear whether I have succeeded in remembering my sister or am instead remembering some other person. I will not pursue this issue further here.

[10] Psychologists have challenged it as well. See, for example, Nigro and Neisser (1983).

[11] It might seem odd to recruit Martin's ideas in order to defend a conceptualist view, given that, as we have seen, he elsewhere advances substantial criticisms of conceptualism. Nevertheless, there is no internal inconsistency in Martin's views – his conception of the role of imagery in episodic memory is logically independent of his arguments about the role of concepts in perceptual experience and memory.

[12] The notion of intentionality or representation invoked here suggests one might be able to have an episodic memory where what is represented in the memory fails to refer to anything. This is misleading however, because memory is a factive attitude, meaning that one cannot succeed in remembering something that did not happen. So in the case of an episodic memory, where a subject remembers something she experienced in the past, the object of experience represented in the memory must have a unique reference in order to secure the state she is in as an actual case of memory.

[13] It is worth pointing out that Campbell's notion of "deep decentering" involves the idea that in episodic memory a subject simulates the original experience rather than representing it. According to Campbell (2002: 191), this requires conceiving the object of the experience as a constituent of the memory, rather than

the memory as bearing a representational relation to the object. Nevertheless, both Campbell and Martin envision the constitutive link between episodic memory and the recalled experience as involving some sort of relation between the memory and the experience that in no way depends on memory imagery matching experience. It might do justice to their views to say that they both posit an intentional relation between memory and the recalled object, where representing is one sort of intentional relation and simulating is another. Interestingly, both Campbell and Martin conceive their views as developing Russell's (1912) notion of memory as acquaintance with past.

[14] I am extremely grateful to John Campbell and Alva Noë for suggesting invaluable improvements to an earlier version of this essay.

References

Brewer, B. (1999). *Perception and Reason*. Oxford: Oxford University Press.

Byrne, A. (2004). Perception and Conceptual Content. in *Contemporary Debates in Epistemology*. ed. E. Sosa and M. Steup. London: Blackwell Publishers.

Campbell, J. (2002). *Reference and Consciousness*. Oxford: Oxford University Press.

Dretske, F. (1995). *Naturalizing the mind*, *The Jean Nicod Lectures, 1994*. Cambridge, Mass.: The MIT Press.

Genone, J. Manuscript. 'Memory and the Content of Experience'.

Harman, Gilbert. (1990). The Intrinsic Quality of Experience. In J. Tomberlin, ed., *Philosophical Perspectives* vol. 4, Ridgeview Publishing Co.

Heck, R. (2000). Nonconceptual Content and the "Space of Reasons". *Philosophical Review*, 109 (4).

Locke, John (1690/1975) *An Essay Concerning Human Understanding.* P.H. Nidditch (ed.). Oxford: Clarendon Press.

Martin, M. (1992). Perception, Concepts, and Memory. *Philosophical Review* 101 (4).

------.2001. 'Out of the Past: Episodic Recall as Retained Acquaintance'. In C. Hoerl and T. McCormack (eds) *Time and Memory.* Oxford: Oxford University Press.

Nigro, G. and Neisser, U. (1983). Point of View in Personal Memories. *Cognitive Psychology*, 15.

Russell, B. (1921). *The Analysis of Mind*. London: Allen & Unwin Limited.

------.1912. *The Problems of Philosopshy*. Oxford: Oxford University Press.

Searle, J. (1983). *Intentionality*. Cambridge: Cambridge University Press.

Speaks, J. (2005). Is there a problem about nonconceptual content? *Philosophical Review*, 114 (3).

Sutton, J. (2004). ‘Memory’. *Stanford Encyclopedia of Philosophy*.
Tye, M. (1997) A Representational Theory of Pains and Their Phenomenal Character. in Gunther 2003.

James Genone
University of California, Berkeley
jgenone@berkeley.edu

Janet Levin
University of Southern California

Can Mental Images Provide Evidence for What is Possible?

Abstract Recently, a number of philosophers have argued that sensory images – "mental pictures" or other sense-based images of various situations – provide the best evidence for what is possible. In this paper I identify the best argument for this conclusion, but contend that it shows that certain non-sensory representations provide good evidence for possibility as well. That is, though I endorse the claim that the sensory imagination can be a source of evidence for what is possible, I deny that it is the only source. I also sketch some consequences of this view for the thesis that sensations and perceptual experiences are identical to physical states.

Key words Metaphysically possible, metaphysically impossible, non-sensory representations, sensory representations.

Recently, philosophers working in the analytic tradition have shown interest in the role of the sensory imagination – the capacity to "picture" or form some other sense-based image of various situations – in modal epistemology, the study of whether and how we can have knowledge about what is not just physically, but metaphysically, possible and impossible.[1] Informally, when trying to justify such modal claims, we make reference to images we can see or hear or feel "in our minds": we take ourselves to be justified in believing that the dog *could have* barked, even though it in fact remained silent, or that oyster shells *could have been* smooth even though they all are rough, or that there *could have been* humans who can fly, even though this is *physically* impossible, because we can form a clear "mental picture", or some other kind of sense-based representation, of such things. Conversely, we take ourselves to be justified in believing that there *could not be* round squares, or surfaces that are simultaneously red and green all over, since we cannot picture any such things. But this is often regarded as merely a manner of speaking, since it also seems that our purely imaginative capacities are too limited to provide evidence for what is possible, given that there are known possibilities that the sensory imagination cannot depict.

These limitations were made clear (perhaps most famously) by Descartes, in his *Sixth Meditation* discussion of the distinction between imagining and conceiving (or understanding), that "purely intellectual" mental activity that yields representations with pure conceptual content. It is conceiving, and not imagining, Descartes argues, that should be taken to provide evidence for what is possible, since imagining would miss too much; for example, we can conceive (or understand) that there is a difference between a 1000-sided figure and a 1001-sided figure, even though we are unable to form a "mental picture" that can distinguish between the two.[2] And, as Berkeley pointed out in his argument for immaterialism, we cannot form a mental image of an object too small to be seen or touched, or a sound too high to be heard, or a surface that is extended but neither colored nor textured; thus, if the sensory imagination provides the best evidence for possibility, we are forced to conclude, along with Berkeley, that such things are not possible. (Descartes, or course, used these observations about the (sensory) imagination to argue, conversely, that since such things are clearly possible, there must be another, better, means of representing and acquiring knowledge about what is possible and impossible, namely, pure understanding or intellection.)

However, contemporary theorists who argue that the sensory imagination has a special role in establishing claims about possibility have a more complicated picture of the relationship between the two. One of the most articulate proponents of this view is W.D. Hart, who argues (1988, 2003) that "perception is to knowledge of actuality as [sensory] imagination is to knowledge of possibility", a view about our justification for claims about modality that he likens to empiricism about the external world.[3] The roles of perception and sensory imagination, as Hart sees it, are analogous in a number of ways. Just as sense-perceptions cause, and are considered to be the "basic" source of justification for, the beliefs they produce about *what there actually is*, so sensory imaginings cause, and are considered to be the basic source of justification for, the beliefs they produce about *what there could be*.

However, as Hart acknowledges, there are also "non-basic" sources of justification for claims about possibility, just as there are "non-basic" sources of justification for claims about actuality, such as induction, inference to the best explanation of the data from imagination, and coherence among the theories that are invoked to explain these data. In this way we can use induction to argue for the possibility of a situation for which we cannot have a clear mental image, such as the existence of a 1000-sided figure: we can form an image of an octagon, and then a nonagon – and then add an inductive “and so on”.[4] Similarly, we could argue for the possibility of an object too small for humans to see or feel by forming a mental image of a pebble, dividing it, dividing it again, and again – and again add an inductive “and so on”.[5] Or (contrary to Hume), even if we could *not* form a mental image of it, we could use interpolative procedures on the shades of color we can picture to affirm the possibility of a “missing shade of blue”. But, because we cannot use these (or other “abductive”) methods on our mental images of a circle and a square to argue for the possibility of a round square, or on our mental images of a red surface and of a green surface to argue for the existence of a surface that is simultaneously red and green all over, we will not generally draw illegitimate conclusions by use of these "non-basic" sources of justification. Nonetheless, Hart argues, deliverances of the sensory imagination are not infallible; there can be imaginative illusions, even hallucinations – just as there are perceptual illusions and hallucinations – and, in both cases, the rejection or retention of a piece of basic data will proceed by appeal to further "basic" data, and a variety of inductive and explanatory considerations of the sort described above.

The thesis, then, is not that the sensory imagination must be able to depict every situation that is genuinely possible, nor that the sensory imagination, either by itself or enhanced by abductive methods, cannot sometimes lead to mistake. It is rather that the sensory imagination provides basic (if not infallible) justification for our modal beliefs, and also that such justification cannot be provided by the more intellectual faculty of conceiving or understanding. I endorse the view that the sensory imagination can provide evidence for what is possible. However, the view that this role is played *only* by the sensory imagination, and not the more intellectual operations of conceiving or understanding, is a controversial thesis, and it is the one that I will examine, and criticize, here.

Before proceeding, however, there are two prior questions that must be addressed. The first is whether the perceptual experiences and “mental pictures” taken to justify beliefs about (respectively) actuality and possibility are purely imagistic, or unconceptualized, or rather have a propositional structure as well. Ever since Sellars argued, in his well-known discussion of the “Myth of the Given” (1956), that relations of justification or evidential support can hold only between representations with propositional content, the dominant view has been that beliefs can be justified only by further beliefs (or other mental states with propositional content), since only they can display the logical, inductive, or, more generally, explanatory, relations that make one proposition evidence for another. Indeed, as many theorists have argued, it is unclear whether pure images can be representations at all; for example, unless a propositional “caption”, or interpretation, is added (at least implicitly) to a picture of three bees, it cannot be determined whether the image represents three individual bees, the bee species, the virtue of industriousness, or the Barberini family.[6] Thus it was often assumed that even if sensory experiences have an imagistic component that cannot be reduced to their propositional contents, it will be epistemically inert, since it is of the wrong form to figure in justification.

More recently, however, a number of epistemologists have challenged these conclusions, and have begun to develop theories in which unconceptualized sense-experiences can be representations, and can by themselves serve to justify at least some of our beliefs about the world.[7] These views are controversial, and cannot be evaluated without addressing the debate about the existence of “non-conceptual content” and its place in the “space of reasons”.[8] We can sidestep this debate here, however, since, as will become evident, the primary question I will address is whether our sensory imaginings can provide *greater* justification for our beliefs about possibility than their paler, purely conceptual, counterparts – whether, for example, imagining, rather than merely conceiving, of some proposition P provides more compelling evidence that P is possible.

The second question is why we should trust either our ability to imagine *or* conceive (understand) some proposition P to provide evidence that P is possible. Ever since Kripke reminded us of the distinction between propositions that are (metaphysically) necessary and propositions that are knowable *a priori*, there have been doubts about whether either imagining or conceiving can have any evidential force whatsoever in determining what is metaphysically possible. If one agrees with Kripke that proper names pick out the same objects in all possible worlds in which those objects exist, it follows that 'Carlo Collodi is Carlo Lorenzini' is a necessary truth. But this truth is not knowable *a priori*, and so – or so it seems – we can conceive or imagine a world in which Collodi is distinct from Lorenzini. Similarly, if one agrees with Kripke that natural kind terms, such as 'water' or 'heat', pick out substances with the internal structure of the paradigm examples we denote by such terms in every possible world (in which it exists), and also that scientific terms work like proper names, it follows that 'water is H2O' and 'heat is molecular motion' are necessary truths as well. But it seems as if we can conceive or imagine that water exists in the absence of H2O, and heat in the absence of molecular motion (and, in each case, vice versa).

There are important and controversial issues about the relation of imagining (or conceiving) to possibilities that are metaphysical but cannot be known *a priori*,[9] but again we can sidestep this debate by restricting the discussion to cases of metaphysical possibility which we *can* know *a priori*, which include (among other things) questions about whether objects endure or perdure, whether causation is nothing but constant conjunction, whether knowledge is justified true belief, whether natural kinds are essentially constituted by their microstructural properties, and whether mind is identical with body.[10] All the important controversies about the role of images in providing evidence for possibility can be addressed even if we restrict ourselves to modal claims of this sort.

However, as Hart acknowledges, there is an important epistemic *asymmetry* between perception and sensory imagination that threatens to undercut the argument for a unique evidential link between (sensory) imaginability and possibility. When perception is veridical, there is a straightforward causal connection between our sense-experiences and the objects and properties perceived, and though there is debate about why beliefs about the world based on sense-perceptions are (often) justified, there is agreement that this causal connection must be at least part of the story. Possible worlds (possibilities), on the other hand, are supposed to be abstract objects, so they cannot be causally responsible for our episodes of sensory imagination – or, for that matter, our conceivings or understandings. As Hart puts it (1988, p. 11), "it is a metaphysical axiom that causal signals cannot pass from unactualized *possibilia* to *actualia*. Independently existing possible worlds seem to be required for objective modal truth, but also to make modal knowledge impossible." So even if our modal beliefs are based upon our sensory imaginings in the way that our empirical beliefs are based upon our sense-experiences, we still must ask: what – if not causal contact with its subject matter – ensures that our sensory imaginings provide even *prima facie* evidence about the existence and nature of the possibilities that we affirm?

It is true, of course, that we are not in causal contact with *all* the entities of which we take ourselves to have knowledge. Numbers and other mathematical entities are also abstract objects incapable of having effects on our mental states, but surely, it seems, we can have knowledge or justified beliefs about mathematics. As Quine has argued, however, our mathematical beliefs can be justified, even though they are causally isolated from their subject-matter, because the existence of numbers (or at least sets) is needed to explain and systematize the data of natural science. One might thus think that (some coherent class of) our beliefs about possibilities can be justified on this model as well, but, as Hart points out, we don't need anything beyond causal necessity and subjunctive conditionals to do natural science, and thus the analogy between mathematics and modality fails. It may therefore seem that the analogy between perception and sensory imagination is not sufficiently strong to justify the modal beliefs that derive from our sensory imaginings, and that there are no other considerations that can be used to argue that our beliefs about modality can be justified, or count as knowledge.

Nonetheless, Hart presents an argument that the sensory imagination provides evidence for possibility which has certain affinities with Quine's argument for the possibility of mathematical knowledge. He does not argue that we need modal truths (that go beyond claims about *physical* possibility and necessity) for science, or for anything important for survival, or for the acquisition of other sorts of knowledge about the world. Indeed, he acknowledges, we don't *need* them for anything except to satisfy the intellectual curiosity that demands an answer to questions such as the mind-body problem, and thus requires the ability to ascertain the truth or falsity of modal claims.[11] Hart also argues (1988, p. 15), in contrast to Descartes, that the sensory imagination can be the *only* source of evidence for such claims, since the more intellectual faculty of conceiving or understanding leads to too many mistakes, given that it permits us to entertain *impossibilities* in constructing *reductio* proofs. Hart's "argument from need", however, may seem weak – given that the "need" is noticed by so few – and can easily prompt endorsement of what he describes (p. 13) as the Quinean suggestion that "modality is all a delusion and to be outgrown when one gives up the metaphysics of childhood and the stone age." The force of this argument can be enhanced, however, when combined with another consideration that Hart mentions, but does not sufficiently emphasize.

This argument begins by noting another analogy between sense-perception and the sensory imagination, namely, that there is a special *phenomenology* common to both which makes it seem, overwhelmingly, as if we are being presented with the truth.[12] Seeing is believing, it is often said, and so, often, is the formation of particularly vivid memory images – and this persists even in situations in which one has been warned about the unreliability of eyewitness testimony, or the ease with which false memories can be induced. And the same can be said, or so it seems, for the relation between sensory imaginings and beliefs about possibility. Hart notes this special phenomenology of our sensory imaginings in passing (1988, p. 13), and it has been noted by others interested in the epistemology of modality as well.[13] As it stands, however, this commonality, though perhaps interesting, seems epistemologically irrelevant, since this "truth reflecting" phenomenology is merely a *psychological* feature of all of these mental states.[14]

One could attempt to argue that the best explanation of how *sense-experiences* came to have this phenomenology shows that beliefs based upon them are justified. This special phenomenology of our sense-experiences, one might argue, tends to compel beliefs about the existence and nature of the objects in one's environment, which we act upon in our daily lives. And since this phenomenology is an enduring feature of our sense-experiences, one might continue, it is clearly the product of natural selection. But if nature has selected for this belief-compelling feature of our sense experiences, then it is reasonable to think that these experiences indeed provide accurate information about the world, since having, and acting upon, accurate beliefs is most conducive to survival.

There are many problems with this argument.[15] But even if it were correct, one could not use a similar argument to explain the analogous phenomenology of our sensory imaginings, since the possession of accurate information about metaphysical possibility and necessity would have no effect on survival in the actual world. Once again the lack of any causal link between our sensory imaginings and what is possible undermines the analogy between imagination and perception, since unactualized possibilities could not have had an effect on our ancestors at any point in our evolutionary history.

But there is a more promising way to use the analogy. Some philosophers have suggested that the special phenomenology of perceptual experiences gives them *prima facie* justification; these experiences seem, intuitively, to reflect the presence (and nature) of objects in the external world, and thus we are justified, at least until there is conflict between these experiences and other perceptual experiences or well-justified beliefs, to take this at face value. This, at least arguably, was G.E. Moore's strategy in his (1939) attempt to challenge the skeptic's claim that, since we cannot know that we are not dreaming, we cannot have knowledge of the external world, and it has recently been defended forcefully by James Pryor (2000, 2004). As Pryor puts it (2000, note 37)

> In my view…[what] explains why our experiences give us the immediate justification they do…is the peculiar "phenomenal force" or way our experiences have of presenting propositions to us. Our experiences represent propositions in such a way that it "feels as if" we could tell that those propositions are true – and that we're perceiving them to be true – just by virtue of having them so represented (…) It is difficult to explain what this "phenomenal force" amounts to, but I think it is an important notion, and that it needs to be part of the story about why our experiences give us the justification they do.[16]

This view, to be sure, is controversial. But one could argue, analogously, that the peculiar phenomenology of our sensory imaginings, and their tendency to compel beliefs about possibility, gives them *prima facie* justification, and the modal skeptic can be challenged on these grounds as well.[17]

This strategy appears to have more promise. One might wonder, however, whether sensory imaginings have anywhere near the phenomenological force, or the belief-compelling efficacy, of sense-perceptions. It may seem, for example, that Pryor himself is dubious. He writes, in a remark embedded in the note quoted above, "I think this "feeling" is part of what distinguishes the attitude of experiencing that p from other propositional attitudes, like belief and *visual imagination*. Beliefs and *visual images* might come to us irresistibly, without having that kind of "phenomenal force" [my emphasis]." But the strategy in fact is compatible with Pryor's remark, since, while visual (and other sensory) images may not "represent propositions in such a way that…"feels as if" we could tell that those propositions are true", they do, arguably, represent propositions in a way that feels as if they are *possible*; that is, when we focus upon them in our modal deliberations, they possess a distinctive phenomenology much like that of sense-perceptions, and prompt belief that the situations they represent are possible. This, I suggest, is the distinctive phenomenological feature of sensory images noticed by modal epistemologists such as Hart.

It may, however, seem dubious that sensory imaginings can be compared to sense-perceptions in their ability to compel belief (which at least seems to be *part* of the phenomenological force of sense-perceptions). Many theorists, including Descartes and Hume, have pointed out that it is nearly impossible to sustain any sort of skepticism about the external world. Skeptical worries invariably vanish when one goes about one's daily life, or enjoys oneself at the backgammon table, thus demonstrating the force of our sense-experiences in compelling belief about the existence and nature of objects in the world. In contrast, it seems that one can perfectly well go about one's daily life when in the grip of *modal* skepticism, and even be tempted to agree, with Quine, that all this requires is the willingness to give up what may be called the "metaphysics of childhood".

Daily life, however, is not the place to test the belief-compelling efficacy of our sensory imaginings, since none of our ordinary activities would be affected by the truth or falsity of modal claims. A better place to test the force of these imaginings is in intellectual inquiry, in trying to determine the nature of knowledge, the conditions for personal identity, the nature of right action, and the relation between mind and body. Here, if our sensory imaginings prompt us to take a position about such metaphysical possibilities and necessities – or to worry if we do not possess other evidence that can defeat the claims which we feel thereby compelled to believe then this should be enough to sustain the analogy between sense-perceptions and sensory imaginings, and justify our taking sensory imaginings, at least provisionally, as evidence for modal claims.[18]

This, therefore, does indeed appear to be the most promising strategy for arguing that our sensory imaginings provide justification for our modal beliefs. But it will not support the thesis that it is *only* our sensuous imaginings, and not our conceivings or understandings, that can play this justificatory role if these non-sensory conceivings or understandings share the distinctive phenomenology of our sensory imaginings that serves to compel belief. And it seems clear that at least some of them do. Don't we feel just as compelled, when we consider the question, to think that simple arithmetic falsehoods just *couldn't* be true as to think that there couldn't be a surface that is both red and green all over? As Descartes puts it in his *First Meditation*, “For whether I be awake or asleep, two plus three makes five…nor does it seem possible that such obvious truths can fall under the suspicion of falsity.” And aren’t we just as convinced that each natural number must have a successor, when we apply simple induction to what we can explicitly conceive, as that there can be a 1000-sided figure? It seems plausible, that is, that what we may call "simple conceivings" deserve to be part of our epistemic base as much as sensory imaginings. And, of course, what differentiates these “simple conceivings” from the assumptions we entertain in *reductio* proofs is that the *reductio* assumptions *lack* the phenomenology and compellingness in question.

Thus I am happy to endorse the view that both our simple conceivings and our sensuous imaginings can be evidence for claims about metaphysical possibility and necessity. However, I am a materialist – and the best-known and most compelling arguments against the identity of minds and bodies, or mental states and physical states, rely on the premise that we can conceive (Descartes) or sensorily imagine (Hart) disembodied minds, or creatures physically identical to ourselves without conscious experience. One might wonder, then, whether it is possible to have it both ways.[19] My answer is yes; that though there is a general link between imaginability (or conceivability) and possibility, it breaks down in just this case.

The imaginability (or conceivability) of disembodied minds, or physical duplicates of ourselves without conscious mental states, I suggest (along with many other theorists),[20] has a different explanation. Crucial to the imaginative depiction of such creatures is the use of subjective, first-person experiential concepts—those that purport to denote the "what it's like" to feel pain and see red, or maybe even "be me". These seem importantly different from physical, functional, mathematical, or other objective concepts – in the way they are acquired, the role they play in thought, and in the (demonstrative-like) way that they refer. Thus, many theorists have argued, there may be principled grounds for refusing to take imaginings involving such concepts to give epistemic support to beliefs about what is possibile – even if we do, in general, think that imagination provides (at least prima facie) evidence of modal fact in cases involving third-person, objective, concepts. In arguing that subjective, first-person concepts are peculiarly unsuited to provide evidence for claims about what it possible, one must argue that the features that make them so are in fact unique to first-person concepts, and provide a satisfying explanation of why first-person concepts might behave, in our sensory imaginings, in such non-standard ways. This is not the place to evaluate this suggestion, or go into the details of what it entails.[21] But if such a principled distinction between imaginings that involve first- and third-person concepts is possible, and if it turns out that the imagined scenarios that seem most to undermine materialism involve these first-person concepts,[22] then perhaps the materialist's attempt to discount or explain away these episodes of sensory imaginings will look to be a matter of principle, rather than a case of special pleading.[23]

Endnotes

[1] By “metaphysically possible” I mean simply *true in some possible world*, and by “metaphysically impossible” I mean *true in mo possible world*.
[2] He writes (1641/1979, pp. 45-6): “And to make this very clear, I first examine the difference between imagination and pure intellection. So, for example, when I imagine a triangle, I not only understand that it is a figure bound by three lines, but at the same time I also intuit by my powers of discernment these three lines as present – this is what I call “imagining”. But if I want to think of a chiliagon, I certainly understand just as well that it is a figure consisting of a thousand sides...but I do not imagine those thousand sides in the same way…[even if] whenever I think about something corporeal, I always, out of force of habit, imagine something – nevertheless it is evident that it is not a chiliagon. This is so because it is not really different from the figure I would represent to myself if I were to think of a myriagon or any other figure with a large number of sides.”
[3] For a similar view, see Kung, 2005. Hart also argues that only sensuous imaginings provide evidence for possibility, and so, as a mind-body dualist, argues at length in his (1988) that one can *imagine* being disembodied.
[4] See Hart (2003). In his (1988, p. 18), he makes a similar point about imagining that the world contains an infinite number of objects.
[5] This example should be familiar from Locke’s discussion of the distinction between primary and secondary qualities in *Essay* II. viii. For another discussion of what may be called “imaginative induction”, see Peter Kung (2005), who also argues for a unique evidential link between (sensuous) imaginability and possibility.
[6] Another example is Wittgenstein’s (1953) well-known figure of a person who could be interpreted either as climbing up or sliding down a hill. In addition, it is only quite recently that philosophers of mind working in the analytic tradition have ceased to be embarrassed by the existence of mental images. It was hard to see how mental images, whether components of sense-perceptions or thoughts, could be accommodated by a computational theory of mind, or realized in the brain, but recent psychological theories have suggested that both can be done.
[7] For example, Bonjour (2002, p. 214), suggests that the mere awareness of these features of sensory experiences can justify beliefs about the nature of these experiences, since there is a special relationship between this sort of awareness and belief which, though not propositional, involves something more than mere causation. Peacocke (2001, sec. IV) argues that if the perceiver can judge, of her sensory experience, that it has certain features, and if the cause of that experience (when veridical) is sufficient both for the experience to have such features and for the perceiver’s judgment about it to be true, then the sensory experience can justify (or, as Peacocke puts it “make rational”) the judgment.
[8] See Evans (1982) and Peacocke (1992) for influential explanations and defenses of the notion of non-conceptual content; McDowell (1994) for criticisms. For an overview of this debate, see Bermudez (2004).
[9] Kripke’s own view about these cases is that, when apprised of these identities (and internal structures), we will retract our claim to have imagined a world in

which Collodi exists without Lorenzini, or water without H2O, and reconstrue what we took ourselves to have imagined as a world in which a mere qualitative (or epistemic) *counterpart* of Collodi (for example, the author of *Pinocchio* – which in other worlds may be someone other than Collodi) can exist without Lorenzini, or a world in which a mere qualitative (or epistemic) *counterpart* of water (for example, the stuff that fills the lakes and comes out of the faucets – which, in other worlds may be distinct from water) can exist without H2O. But others suggest, instead, that it is better to say that we *can* imagine these scenarios, and that imaginability is not always evidence for possibility.

[10] Alternatively, we could ask whether what we initially interpret as a sensory imagining that P would provide better evidence for P than what we initially interpret as a conceiving that P, if these initial interpretations are undefeated by further empirical knowledge.

[11] He writes, "If one thinks that the mind-body problem is a real problem with a correct answer…then…[w]hat can one do but insist that knowable, objective modal truth must be possible and hope that someone will someday explain how it is possible? Forlorn as this hope is, nothing better is evident." (1988, pp. 13-14).

[12] One might call this feature of these mental states, with apologies to Stephen Colbert, their "truthiness".

[13] See Bealer (2000), and Kung (2005).

[14] Hart brings up the special phenomenology of sensory imaginings, in his (2003), primarily to distinguish his modal epistemology from Quine's argument that we need mathematical knowledge for science, though he notes that that our sensory imaginings, like our sense-experiences, have a special phenomenology with "evidential weight", adding that "[i]t is the plain natural history of much modal belief, to which history any modal epistemology owes respect, that modal belief tracks imagination." But he doesn't say much about how modal epistemology is to respect that history.

[15] For example, it is not clear that traits that endure in a species have a selective advantage, instead of being (not overly destructive) by-products of other traits that do. It also not clear that the capacity to acquire *true* beliefs is best conducive to survival. See (among many others who make such criticisms) Stich (1984/1994).

[16] See also his (2004, p. 357).

[17] See Kung (2005) for an attempt at this strategy.

[18] This also provides an answer to the worry, expressed in the text, that no attempt to use an "argument from need" to conclude that modal knowledge is possible will succeed, because the "need" is recognized by so few. The answer is that "the many" just haven't put themselves in situations in which the need would be felt.

[19] Hart, for one, denies this. He writes (2003), "…one who takes modality seriously should be a dualist, and a thorough-going materialist should dismiss modality". David Chalmers is perhaps the best-known proponent of this view.

[20] See particularly Loar (1990/1997), Nagel (1974), and Hill (1996).

[21] But see Loar, Nagel, Hill, cited in previous note, and also Levin (2004).

[22] For example, in Hart's (1988), he worries about an objection to his principle linking imaginability to possibility that rests on the premise that

(P) It is possible that a person who could not be disembodied imagine being disembodied.

The only evidence for belief in (P), though, says Hart, is our ability to imagine it. But this, he continues, "seems to require us to imagine an embodied person whom we are unable to imagine being disembodied, though he does. In order for us to imagine him imagining himself disembodied, it would be best and most convincing for us to put ourselves in his shoes and limn for ourselves the imaginings he has of himself disembodied [my emphases]. If we do so we will thereby have imagined him disembodied, thus frustrating the requirement that we be unable to imagine him disembodied." (33)

I'm not sure how important this strategy is, in the end, for making the point Hart is concerned with in these passages. But here is evidence, at the very least, of the recognition that these first-person abilities are germane to the argument.

[23] I am grateful to W.G. Hart and Janine Jones for their comments on an earlier version of this paper.

References

Bealer, G. (2000). A Theory of the A Priori: Intuition, Evidence, Concept-Possession. *Pacific Philosophical Quarterly*, 81.

Bermudez (2004). Non-conceptual Mental Content. *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/content-nonconceptual.

Descartes, R. (1641/1979). *Meditations on First Philosophy*. Indianapolis, IN: Hackett.

Evans, G. (1982). *Varieties of Reference*. Oxford: OUP.

Hart, W.D. (1988). *The Engines of the Soul*. Cambridge: CUP.

Hart, W.D. (2003). The Music of Modality. *Topoi*, 22.

Hill, C. (1996). Imaginabiliy, Conceivability, Possibility, and the Mind-Body Problem. *Philosophical Studies*, 87.

Kung, P. (2005). Is Imaginability a Guide to Possibility? unpublished.

Levin, J. (2004). The Evidential Status of Philosophical Intuition. *Philosophical Studies*, 121.

Loar, B. (1990/1997). Phenomenal States, in Block, Flanagan, Guzeldere (eds.), *The Nature of Consciousness*. Cambridge, MA: MIT Press.

McDowell, J. (1994). *Mind and World*. Cambridge, MA: HUP.

Nagel, T. (1974). What is it Like to Be a Bat? *Philosophical Review*, 83.

Peacocke, C. (1992). *A Study of Concepts*. Cambridge, MA: MIT Press.

Pryor, J. (2000). The Skeptic and the Dogmatist. *Nous*, 34.

Pryor, J. (2004). What's Wrong with Moore's Argument? *Philosophical Issues*, 14.

Sellars, W. (1956/1967). Empircism and the Philosophy of Mind", in Sellars, W. *Science, Perception, and Reality*. London: Routledge.

Stich, S. (1984/1994). Could Man Be an Irrational Animal? in Kornblith, H. *Naturalizing Epistemology* 2nd Edition. Cambridge, MA: MIT Press.

Janet Levin
University of Southern California
levin@email.usc.edu

Marco Mazzone
University of Catania

Imagery, Language and the Flexibility of Thought

Abstract In two recent papers, Dan Sperber and Peter Carruthers have addressed the issue of cognitive flexibility, giving us different but somehow complementary accounts of it. Here I intend to focus on another cognitive mechanism which plays some role in allowing flexibility, and has been given little emphasis in their accounts. This mechanism is sensory imagination. In so doing, I have to confront with the assumption, which is widespread in the philosophical domain, that perceptual representations cannot convey any thought process. In the first place, I argue that this assumption rests more on the notion of accessibility than on that of systematicity. In the second place, I argue that – as an argument from the frame problem seems to show – accessibility of propositional representation has been largely overestimated; besides, there are reasons to think that systematicity and accessibility of perceptual representations have been largely underestimated. Those arguments are held to support the conclusion that people, and non-human animals too, can make use of sensory imagination as a cognitive strategy to confront with novel state of affairs. This conclusion fits well with a large amount of research in comparative psychology, and converges with current models of controlled thought processes which are based on the notion of a mental global workspace. The differences with the accounts of Carruthers (2005) and Sperber (2005) are briefly investigated.

Key words Imagery, flexibility, language, sistematicity, controlled thought processes.

Introduction

In the philosophical field there is a large consensus on the thesis that thought is an essentially propositional affair. Of course, the most famous formulation of this view is the Language of Thought (LOT) hypothesis, sponsored by Jerry Fodor (1975). It's worth noting that according to this hypothesis LOT is also the explanation of the fact that babies and animals can think, even though they do not master a verbal language. So, thought processes would be embedded in a propositional medium even in non-human animals.

Standard as it can be amongst philosophers, this account of thought conflicts with a large amount of research in empirical psychology and neurophysiology. In these areas of research it is generally held that, while performing reasoning tasks, people do make use of analogical representations, strictly related though probably not identical to perceptual and motor ones. In short, mental images enables some sort of thought processes.

However, there is a second point on which Fodor's view has been contended even amongst philosophers: it is when he declares that central processes of thought are condemned to remain a mystery for cognitive science, in that cognitive science needs modularity, whilst, in his view, central processes are paradigmatically non-modular. More exactly, Fodor thinks a computational theory of thought has to face the well-known "frame problem": if we define thought as the ability to assemble pieces of information properly, in the central processes every piece of information, which is stored in the system or comes from input subsystems, is in principle accessible and can play a role. The problem of selecting precisely the relevant information for any given situation appears to be computationally intractable, at least if we want to preserve our present ideas about what a cognitive science has to be. In other words: in (central) thought processes, the accessibility of all the pieces of information which are couched in a propositional format is so overwhelmingly great that it's hard to imagine what could rationally constrain the selection of the pieces which are worthy of being considered.

Recently, both Dan Sperber and Peter Carruthers have proposed to reject Fodor's idea that central processes and modularity cannot be reconciled, and have tried to sketch an explanation of the way a modular architecture could grant – in Sperber's (2005) words – flexibility and context-sensitivity to thought processes. (More exactly, Sperber (2005:54) discriminates between "flexibility" and "context-sensitivity", using the first to refer to "context-sensitivity in the longer run", that is, "when humans in general are described as a particularly flexible species […] adapted to very diverse natural and humanmade environments". For the sake of brevity, I will use "flexibility" in a more general way, as covering both Sperber's expressions.) The speculations of Carruthers and Sperber are very stimulating, and in some way they complement each other. Carruthers puts his emphasis on high-level processes which rely on language faculty; Sperber focuses on aspects which are built in the general architecture of the brain, resulting in low-level unconscious processes. Besides, Carruthers incidentally mentions mental imagery as a factor that could have a role in the increasing complexity of a creative and flexible mind. Mental imagery occupies an intermediate position between linguistic and conscious processes on the one hand, and non linguistic, unconscious ones on the other. In this paper I intend to focus on the role this factor plays in allowing flexibility.

In practice, I will proceed this way. First, I will distinguish between the notions of systematicity and accessibility of mental contents, arguing that it is the latter which in fact seems to prevent perceptual representations from being relevant for thought processes (section 1). Second, I will argue that accessibility of propositional representations has been largely overestimated, and anyhow misunderstood (section 2); and conversely that accessibility of perceptual representations has been largely underestimated (section 3). Third, I will argue consequently that perceptual representations can in fact support both beliefs (section 4) and forms of reasoning based on sensory imagination (section 5). Finally, I will compare my account of flexibility with the ones given by Carruthers (2005) and Sperber (2005).

1. Systematicity and accessibility

Let's start from a clarification of Fodor's LOT hypothesis. The genuine intuition underlying it is that, in order to convey thought processes, a representational system has to be systematic. The notion of systematicity appears at a first sight very clear, as shown by one of the formulations that can be given of it: we cannot credit a subject with the ability to think that *Fa* and *Gb* (i.e., that the individual entity *a* is an *F* – fall under the concept *F* – and the individual entity *b* is a *G*), unless the subject can entertain the thoughts that *Fb* and *Ga* too. Obviously, this capacity of mental representations to combine with each other cannot be ruleless: for instance, something must preclude the formation of the strings '*ab'* and *'FG'*, which have no coherent meaning. Then systematicity is a capacity of mental representations to combine with each other *freely within the limits imposed by certain regular constraints*.

Systematicity is the reason why, according to Fodor, thought processes cannot be implemented through analogical representations, in that representations of that kind would lack any capacity to combine with each other systematically. By and large, analogical representations are told to serve only a very limited function, that is, assuring the connection between propositional mental contents and (action in) the world (Fodor 1987). But, as I showed elsewhere more at length (Mazzone 2004), this assumption is largely unwarranted. Let us suppose that a subject is able to represent perceptually the fact that *a* is a cat, and that *b* is black. This ability normally enables her to represent perceptually the fact that *b* is a cat and that *a* is black, too – assuming it is so. And, for sure, perceptual representations do not allow any possible combination of features or entities whatsoever: for instance, there is no way to amalgamate individual entities *a* and *b* in a single visual configuration, as we do instead with *a* and the property of blackness. Therefore, in a sufficiently clear sense perception *is*, after all, systematic. We could elaborate much more on this thesis, but presently what concerns us more is the following general point. Leaving aside problems such as concept combination and prototypes, which can be solved along the lines of Prinz (2002), the real problem we come across when conceiving thought processes as embedded in perceptual representations is *neither* that these representations lack the capacity to combine with each other, *nor* that they lack relevant constraints on their possible combinations. Quite on the contrary, perceptual representations are endowed with a structure which is apt to guide and constraint their combination. What they really lack is what I will call "(full) accessibility", that is, the standard capacity of our cognitive system to access all and only the information specifically needed to form a certain thought.

At least, this looks fine as a first approximation. Doesn't language, after all, allow us to combine freely every part of it with every other (here too, of course, within the limits fixed by some sort of constraints)? And isn't it true that we do not have a similar complete and precise access to the single elements of perceptual representations, in order to form this thought rather than that one? Well, I agree that language helps us to explore possible thoughts, that is, possible combinations of representational elements. But I also believe that this characteristic has been strongly misunderstood, and also idealized, assuming that it is far more pervasive than it really is. Conversely, I believe that accessibility of perceptual representations has been underestimated, and so has been the role this accessibility plays in the emergence of humans' complex cognitive capacities. Let us begin with the first point.

2. The ways of language

As previously noted, Fodor is prepared to pay the price for the full accessibility of propositional representations. In its more ordinary formulation, this position has been put in terms of thought processes being potentially affected by any propositional information available to the system (i.e. every belief that happens to be fixed by the organism). But it is reasonable to assume that the very same capacity is responsible for the fact that we can freely access every mental symbol, and combine symbols with each other in order to explore the range of possible thoughts. So, if we trust Fodor, the frame problem – and consequently, the "mystery" of central processes of thought – would be the price to pay for the full accessibility of propositional representations, both relative to accessibility of beliefs *and* symbols.

In a recent contribution, Sperber and Wilson (1996) pointed to the fact that Fodor's picture of rational processes is a strongly idealized one. His commitment to the thesis that central thought processes can potentially access every bit of (propositional) information stored in the system appears much too strong. We are not rational in that sense, they observe. At most, full accessibility is the limit to which knowledge as an inter-subjective enterprise tends in the long run: "Enduring collective cognitive enterprises, where, through communication, relevance can be better targeted, may begin to display shades of the kind of rationality that Fodor attributes to individual human cognition". I find this observation very enlightening in the present context. It tells us that, in a sense, accessibility of mental representations is more a social than an individual issue. (This does not mean, though, that individual cognition becomes irrelevant: social factors are nonetheless implemented in individual minds.) Plausibly, this fact depends on the public nature of language. In fact, we can isolate two different aspects in the way language contributes to accessibility of thoughts.

To begin with, through explicit communication we can continuously and collectively explore complex theoretical and practical issues, and then fix the results of our explorations so that this collective wisdom gets transmitted from a generation to the next. On the one hand, this fact allows individual minds not to move blindly through the range of possible thoughts: it puts some pressure on restricting our chains of thought to (socially) relevant ones, greatly lowering the necessity to access the multitude of our mental representations. But, on the other hand, this collective fixation of beliefs makes individuals much more expert of the complexity of the world than they could have been by their experience alone: they can get a richer knowledge of the world, which means a richer network of connections between their representations, and therefore a larger reciprocal accessibility of thought contents. In other words: largely through social exposition, subjects arrange their conceptual networks so that, starting from a certain content, other mental contents are easily (and conditionally: "given condition *C*, starting from mental content *a* access mental content *b*…") accessed. In a number of cases, nothing else is needed.

But not in any case. And this leads us to the second point: there is (at least) another way in which a public language can contribute to the accessibility of mental contents. People argue with each other frequently, sometimes ending up with personal beliefs and intentions being modified. Public discussion can therefore lead to permanent or just occasional incorporation of other's beliefs in our conceptual network, even when those beliefs do not belong to the conventional wisdom of human groups. In this sense, the range of thoughts we can access is larger than the number of those enclosed within our individual minds. But inter-subjective dialogue can also have an important effect on subjects' access to *their own* thoughts: it can create a habit to employ a distinctive thought process, which consists of a sort of (more or less conscious) deliberation. That is, humans are trained to discuss with each other in order to envisage the foreseeable effects of different lines of behaviour; as a consequence, they can succeed (though in a variable measure) in internalizing that process, and then they can use it as an individual cognitive resource for the scheduling of actions. Specifically, since they internalize inter-subjective dialogue, individual minds become able to purposely search through stored (and incoming) information, so as to envisage potentially new (i.e. not conventionalized) chains of means and effects, in order to select the preferred course of action. Or could it be the case that linguistic social training just increases such ability, which is already in place? We will come back to this immediately. Let us first summarize what we have found out in this section.

The assumption that in central thought processes we can access any stored (or incoming) information is almost certainly too strong: it would settle the standard for rational thought too high, and as a consequence it would charge us with the frame problem. It is more reasonable to assume that we have only a very limited access to our mental representations, even to those dressed up in a propositional format. So to speak, we should not think of accessibility of mental contents as a general and free by-product of propositional format; we should rather think accessibility as an expensive and always partial product of certain specific cognitive procedures. We isolated two kinds of these procedures. The first kind is the sort of "frozen" accessibility which is guaranteed by exposition to social routines, be they verbal or not. Learning through explicit instruction is just an instance of this type of procedures. In such cases, the individual ability to access mental contents is constrained by the very same social conventions which in some measure promote it. In the second kind of procedures, on the contrary, there may be some exploration of non-conventionalized trains of thoughts. We have considered two forms in which this second procedure occurs: through inter-subjective dialogue; and through an internalized version of it that amounts to individual deliberation.

Amongst the mechanisms we considered, inter-subjective dialogue and individual deliberation are the ones which most approximate the idealized notion of (virtual) full accessibility. We can in principle try to explore all the evidence we happen to possess as far as a given problem is concerned. However, how much, in the short term, our trial will be effective is another issue. By everyone experience, we can persistently ignore the evidence we have at our disposal, either because we can't access it or because we can't see that it is relevant for (we can't access its relation with) our purposes. This is true both in practical and theoretical reasoning. In general, thanks to its public nature, language seems to make possible explicit exploration of our conceptual network. But this exploration is neither free nor complete.

3. The ways of imagery

Let us go back to the question we left unaddressed above: does the ability to explore our mental representations – in order to envisage previously unforeseen connections of thoughts – come with language and discourse? Or is it already in place, at least in a rudimentary form, and language just enhances it?

In Frege's tradition, a strong prejudice against perceptual representations has been widespread among philosophers. In fact, perceptual representations have been conceived of as essentially idiosyncratic and inert "mental images", and therefore they have been deemed to be entirely unsuited for thought processes. Fodor's LOT hypothesis appears to be an heir of this tradition, in that it assumes that thought processes can only take place in an amodal format of representation. We are not going to summarize the whole debate about mental imagery. Psychological experiments such as the ones on mental rotation of geometrical objects reported by Shepard and Metzler (1971), or the ones on mental scanning of mental images referred by Kosslyn (1994), together with more recent neurological evidence, have convinced almost everyone that we must be realist about mental imagery (and sensory imagination in general). There can be little doubt nowadays that perception and imagination share some cerebral mechanisms, so that mental imagery could be pictured as a sort of perception *in the absence of* a percept. Someone speaks of "off-line" perception (or off-line motion, in cases of motor imagery). The discussion nowadays is rather on how large exactly is the correspondence between sensorial and imaginative areas of the brain (Mellet et al. 1998). Besides, recent studies concerning conceptual knowledge (e.g. Barsalou et al. 2003) have strongly supported the thesis that there isn't any amodal representational system of the kind LOT was supposed to instantiate (of course, with the exception of verbal language itself): this means that, except for thought processes which occur in verbal language, mental operations have to take place in modality-specific systems of representation, that is, they exploit the very same representations we are supplied with by our perceptual systems.

For our purposes, the relevant point is that sensory imagination can be seen in fact as a form of reasoning, which allows people to put in connection means with ends. Take, for instance, mental rotation experiments. People were asked whether a given geometrical object was the same as another, once it was opportunely rotated. The results strongly suggest that, quite literally, subjects mentally rotated the first object in order to find out the correct answer. Now, let us think of the same cognitive process of mental rotation as being framed in a different, ecological rather than experimental, context. A subject is looking for a tool with a particular configuration, which is necessary to perform a given action. She remembers an object she saw before, and then she asks herself whether it could perform the desired action; but the point of view she looked at it from was not apt to appreciate whether it had exactly the right configuration. In this situation, mental rotation of the object would be part of a practical reasoning procedure, where the intended tool (or the intended function of it) is the purpose, and the subject has to find out whether the given object plus the rotation allowed to achieve that purpose (i.e., whether they were an appropriate mean for that end).

There are two points here that deserve particular mention.

In the first place, we have to remark that mental operations of the kind we have just considered, though not propositional in the sense called for by Fodor's LOT hypothesis, are nonetheless propositional in a much weaker sense – that is, in the sense of "having combinatorial structure of some sort" (Carruthers 2005). In other words, they can be regarded as sort of beliefs, though they do not exhibit a full-blown propositional status. As previously noted, perceptual representations, and mental operations which are executed on them, have not to be conceived of as if they were unstructured, or chaotic (i.e. unruled): in fact, it does make sense to speak of *beliefs* couched in perceptual, modal-specific representations. This assertion needs some further justification, which will be given in section 4.

In the second place, our ecological example of mental rotation is very opportune in that it brings us to mind that mental imagery is not necessarily triggered by language, and this is what makes it relevant also when talking about non-human thinking. We will come back to this later, in section 5.

4. Perceptual representations and beliefs

As we noted above, the reason why perceptual representations have been deemed unsuitable to convey beliefs is *not* primarily that they are supposed to be wholly unstructured. Rather, the problem is that their structure does not seem to be selectively accessible from central thought processes. For instance, a representation of a black cat is not the same as the belief that the cat is black, for the reason that, without a propositional re-description, we are not able to access precisely the thought THE CAT IS BLACK in order to make some subsequent computation on it – or so we are told. In other words, it would not be the blackness of the cat in particular – rather than, for instance, its furriness – what we represent through a given mental image. But let us imagine a non-human animal feeding itself with a single species of berry, and let us suppose moreover that just one feature distinguishes that kind of berry from others, which are otherwise wholly similar: it is black. In this case, seeing a *black* berry would be sufficient (in the right circumstances) to trigger a feeding behaviour. This means that the animal discriminates precisely the blackness of the berry as being a sufficient condition (in the right circumstances) to prompt a certain behaviour. What I am drawing our attention to is the following difference: on the one hand is the fact that both human and non-human animals do not selectively access a great number of their perceptual representations, in that these representations do not subserve any function in their actual reasoning; quite different is assuming that, as a matter of principle, perceptual representations do not admit to be selectively accessed by thought processes. We should not confuse the former idea, which is trivial, with the latter, which instead is quite problematic. Nor should we be tempted to conclude that, if any selective access has taken place, then obviously perceptual representations must have been re-described in a propositional format. This would beg the question. Everyone concedes that thought processes need procedures which access specific mental contents and properly combine them with each other. What should be demonstrated is that this necessarily requires a propositional system of representation.

Then, since plausibly perceptual representations appear to be both structured and selectively accessible to thought processes, we should better admit they can express beliefs. When Carruthers (2005: 74) credits mammals with beliefs of the form IF X THEN Y, which can subserve as input for practical reasoning modules, he is in fact ascribing them representations which are propositional in the weak sense he himself pointed to. We are not at all obliged to think of them as couched in amodal representations.

5. Mental imagery and problem solving

Let us now return to our "ecological" example of mental rotation. We noticed that such an episode of mental imagery could plausibly occur without being triggered by language; on the other hand, it clearly involved a belief of the form IF X THEN Y: "if I take that object and rotate it in such and such way, then I can get the tool I needed". But assuming beliefs can be vehicled by perceptual representations, as we suggested a few lines above, then it follows that a mental imagery process of that kind could nonetheless be performed by non-human animals, too – no matter if they are not endowed with any propositional system of representation.

However, our example involved in fact a more sophisticated mental aptitude than just that of expressing, or using, a belief. We should better think of it as a case of *formation* of belief ("fixation", in Fodor's terms). The subject had not from the beginning the belief we cited above; rather, she formed it as a result of an episode of mental rotation, once she noticed that that manipulation produced the intended effect (supposing it did). This fact is fairly important, in that it shows a way to answer the question we raised above: does the ability to explore mental representations and combine them creatively come with language and discourse, or is it already in place before (and independently of) the emergence of language? In fact, we can think of our example as an instance of a more general phenomenon: the use of mental imagery procedures in episodes of tool invention – or, even more generally, in episodes of problem solving. When successful, such procedures would cause the formation of beliefs with the following general form: “if I take objects of that sort and I manipulate them in such and such way, then I can get the tool I needed”. So, the question becomes: is there any evidence that such mental procedures really exist? In other words, does any cognitive organism really make use of mental imagery in order to solve practical problems of all sorts?

Obviously, I do not presume to settle here the issue. The ascription of mental processes is always a delicate matter; it brings with it a well-known repertoire of long-standing epistemological troubles. However, we can frame the issue so as to involve a much less demanding task. One could be satisfied if she can show that there is a class of non-human animals’ behaviour which has the following feature: it (comparatively) appears to be creative, that is, it exhibits some flexibility. Under the general assumption that perceptual representations allow selective access and mental manipulations of sorts, so as to be suited to express beliefs, that would be enough to strongly suggest the existence of a mechanism of the sort envisaged above. In other words, in the presence of creative behaviour amongst non-human animals, one would sensibly be urged to admit the existence of mental imagery processes which allow those animals to explore new connections between means and ends: that would be the best explanation on the market. (Though, one could argue that we have other proposals on the market, to be true. For instance, Sperber (2005) proposes a different explanation of creativity/flexibility, which, just as the present one, does not presuppose any propositional code. We will return to this point below.)

In fact, there is a kind of seemingly creative animal behaviour that comparative psychologists have traditionally explained through mental simulation. The ability to solve fresh problems is by definition the main feature of that behaviour. Moreover, non-human animals which are engaged in it typically stop moving for a while and look at the problematic state of affairs; then they suddenly start to act as if they had a clear image of what they are going to pursue. In psychological studies, the involved mental process has been usually conceived of as a form of insight. The chimpanzees of Köhler (1921), which solved some simple problems of getting food through object manipulation, gave the paradigm of that line of research; but I prefer to exemplify it with a more recent case study carried out by Heinrich (2000). He has studied the reaction of crows and ravens when compared with an unusual getting-food problem, and he has observed in the latter animal a behaviour which deserves to be interpreted as a case of insight.

The main experimental set was as follows. A piece of meat was tied to a twine, and the other end of the twine was tied to a trapeze. In order to get the meat, a bird’s only successful strategy would be to draw up a length of twine with its beak while sitting on the perch, then step on the twine, and repeat the procedure about seven times until the meat is level with the perch. Of a dozen crows observed none succeeded on the task. Every one of them flew towards the meat and tried to snatch it: a strategy which was awfully hopeless. Of the six ravens involved, on the other hand, five were successful; in fact they solved the problem within five minutes following the very same strategy. First, they looked at the setup for a little while; second, they pecked to the string where it was attached to the trapeze as if trying to sever it; third, they seemingly tried to break the string off by violently twisting it from side to side; fourth, they applied the correct procedure without any further hesitation.

As Patricia Churchland (2002: 87) has noted, "that none of the ravens got its neck jerked by going for the attached meat while flying suggests that their brains expected what would happen were they to do that and decided against it. That the ravens turned to the pull-up strategy within minutes and were successful in pulling up the meat *in one trial* strongly suggests that the ravens used body-image manipulation in causal problem solving". There could be a danger of theoretical optimism in this conclusion, especially in the first part of it. After all, it is very difficult to rule out the possibility that ravens simply do not have the unsuccessful behaviour of crows in their motor repertoire, or the like. But a point should be recalled here: Heinrich elaborated the experimental set with the deliberate purpose to arrange a thoroughly novel problem, which birds do not encounter in the wild. So, it is highly improbable that the uniform behaviour exhibited by the successful ravens could be explained by an inborn predisposition. Besides, we shouldn't forget that there is a step of the procedure the ravens have to repeat about seven times, without having any benefit until the whole procedure is accomplished. Therefore, at least the second part of Churchland's conclusion is difficult to resist. Ravens' behaviour plausibly instantiates a general kind of mental process, in which imagination is put at the service of simulating possible lines of action, so as to foreshadow what their results could be. In this sense, Churchland (2002) talks of an "emulator hypothesis", which receives – she notes – further support from the fact we have independent evidence about the neural location of such an "emulator". She has in mind exactly the previously cited neural correlation between sensory-motor processes and imaginative processes.

In sum, there seem to be neural circuits engaged in mental simulation of objects' movement and bodily action, which in large measure coincide with the circuits involved in real ("on-line") perception and motor control; there is a large evidence, both introspective and experimental, that we make use of sensory imagination in order to recover, and explicitly access, information so as to establish new connections of thought, sometimes as part of problem solving procedures; and there seem to be examples of creative behaviour amongst non-human animals, plausibly to be explained through the attribution of sensory imagination processes. Altogether, these facts give sensible support to the thesis that some creativity (that is, flexibility, that is, in turn, capacity to combine mental representations in a way which is sensitive to current purposes) is present also before, and independently from, the emergence of language: in fact, it is plausible that this (limited) flexibility depends on sensory-motor imagination processes, which employ motor and perceptual representations in the service of off-line simulations.

6. Carruthers: flexibility and language

What is the relation between the just proposed account and the suggestions of Carruthers (2005) and Sperber (2005) concerning flexibility of thought?

To begin with, Carruthers puts the issue in a slightly different way than we did. His question is how information coming from distinct central, domain-specific, modules (naïve biology, naïve physics, mind-reading, etc.) can be put together. And his answer is roughly that this could be accomplished through the module of language, in that it can access the outputs of each of those other modules. The importance so attributed to language is something everybody intuitively would agree with, though it obviously cannot be the whole answer. The mere fact that beliefs couched in perceptual formats, coming from different cognitive domains, can virtually be re-described in the very same propositional format is not sufficient by itself to explain flexibility, as shown in section 2. The real issue is how the system "decides" which information in fact has to be accessed and re-described, that is, which procedures allow the system to connect the right pieces of information – in particular (but not only) when novel combinations are produced. From this point of view, it is of little importance whether the pieces of information come from different cognitive domains or not. The problem of how the system can obtain context-sensitive combinations rises in both cases.

We can distinguish two points, which connect with our previous discussion. Firstly, for flexibility to be achieved the existence of a mental space in which pieces of information can be combined is not sufficient; the problem is in fact, rather, which pieces of information precisely have to be combined together on a particular occasion. Secondly, we should ask if only language faculty supplies us with a mental space which allows integration of scattered information.

6.1. Selection of appropriate inputs

As to the first point, Carruthers himself admits the difficulty. In the section entitled "Outstanding Problems", he recognizes that "there remains the question of how some central-modular outputs rather than others get selected for encoding into language" (Carruthers 2005: 87). Later on, he suggests that the problem could be solved along the lines of Sperber (2005): we will come back to this suggestion in a moment. What should not be overlooked is the fact that in this context Carruthers' emphasis on the term "flexibility" is perhaps a bit confusing, and his main thesis about the contribution of language faculty to flexibility is at risk of being trivialized. (Of course, his proposal has many other aspects which continue to be worth discussing.) In fact, that thesis seems to simply come down to the consideration that we can speak *in principle* of everything, so that language creates a virtual space in which information coming from different sources can be put together: a thesis that nobody is willing to deny. The real question we have to address when dealing with the problem of flexibility is: how is it that the right information gets selected for encoding into language – as Carruthers would have said – or for being placed in whatever space is responsible for thought processes – as I would prefer to put it –? On that issue, however, not only has Carruthers (2005) nothing to say himself; but moreover, he explicitly defers to a solution which points to low-level mechanisms. There seems to remain no interesting role for language, when addressing the issue of flexibility.

At a more accurate recognition, though, Carruthers argues that flexibility ("creativity") depends on language faculty in another sense, too. This concerns what Carruthers calls "the supposer": a device which has "to generate new sentences for consideration – either at random or drawing on similarities and analogies suggested by perceptual or other input" (Carruthers 2005: 85). Still, Carruthers' assumption that the supposer has to be a linguistic device, necessarily dealing with sentences, is in my opinion wholly unwarranted. The function it is meant to perform is to explore possible combinations of mental contents: precisely what mental imagery allows us to do, as shown in sections 3 and 5. I do not mean to say that language has not any role here; on the contrary, it largely strengthens the ability to explore thoughts. But, first, language faculty is plausibly not necessary in order to have that ability; and second, language faculty in itself is not sufficient to strengthen that ability. It is not just a matter of language faculty (that is, of a linguistic space in which thoughts can be entertained); it is *using* language in social contexts that enhances our ability to explore possible thoughts (see section 2).

We had better note that, no matter whether the supposer is (exclusively) a linguistic device or not, it would support anyhow (more or less conscious) cognitive processes: the system has to search for thoughts which fit in in some given context; in practice, it is endowed with some procedures for thoughts activation (for instance, "at random or drawing on similarities and analogies", as Carruthers suggests), then it contrasts the activated thoughts with the context in order to discard the inappropriate ones and/or select the best candidate. This kind of cognitive process should not be confused with the sort of non-cognitive procedures Sperber (2005) focuses on. In this sense, both my suggestions and Carruthers' notion of "supposer" lie on the same side of the issue: they aim to give an account of creativity/flexibility as a cognitive process.

In sum, Carruthers (2005) seems to acknowledge two distinct aspects of flexibility. As to the first, he ends up recognizing that language faculty is not sufficient to account for it. Having a mental (linguistic) space in which scattered mental contents can be put together is far from being the whole issue; we should rather understand how mental contents are selected so as to be put together effectively. As far as this aspect is concerned, Carruthers defers to Sperber's (2005) proposal, based on non-cognitive processes. As to the second aspect, Carruthers singles out a kind of cognitive flexibility which, according to him, depends on the activation of *linguistic* suppositions. But this assumption is largely unwarranted: it does not even consider the existence of non-linguistic mental simulation – that is, of processes of sensory imagination. And this leads us to the second point I meant to touch: is language faculty the right place to search, if we are looking for a mental space in which mental contents can be combined together to support reasoning?

6.2. Mental workspaces

We argued that motor and perceptual representations are selectively accessible from sensory-motor imagination (and from other mental processes too); and we also reported that there are neural circuits which can accommodate those imaginative processes. This tells us that, after all, there is some sort of mental space in which, after having been selected, perceptual contents can be combined together and then manipulated. Still, we did not directly address that issue; in particular, we did not ask whether there is in fact a distinct mental space for each cognitive domain, or there is just one global mental space, with all the information coming from different domains converging on it. If the latter is the case, then we would have a non-linguistic, domain-general, mental space which is a perfect candidate to perform the role Carruthers assigns to language faculty.

In fact Carruthers (2003) takes this hypothesis into consideration, while discussing the global workspace model of Baars (Baars 1988; Baars 1997); but he rejects it in that "there is nothing in Baars' work to suggest that language will play the role of *integrating* the outputs of other conceptual systems" (Carruthers 2003: 510). To be true, Baars goes further than this: he explicitly maintains that global workspace takes language as just one of the inputs to be integrated (see Baars 2002; Baars and Franklin 2003). Anyway, as far as I can tell, Carruthers rejects the idea of a *non-linguistic* global workspace without any explicit argument, simply because it does not fit well with his commitment to the role of language.

It should be noted that the global workspace model of Baars is but one of a number of models which agree on the following general picture. In human cognition – not just in that, though – there is a clear difference between automatic and controlled processing. Subjects "can temporarily inhibit the automatic activation of some processors and enter in a strategic or 'controlled' mode of processing" (Dehaene and Naccache 2001: 13; see also Schneider and Chein 2003). This controlled mode of processing allows the subjects to "establish flexible links amongst existing processors" (*ibidem*): in other words, controlled processing allows flexible access to (the outputs of) each of a number of domain-specific processors in the system. In order to explain all this, we have to postulate "a distributed neural system or 'workspace' with long-distance connectivity that can potentially interconnect multiple specialized brain areas in a coordinated, though variable manner" (*ibidem*): in other words, this system can access a number of representations scattered through the brain, presumably including the motor and perceptual representations involved in simulation processes. Sensory imagination could therefore be thought of as one of the processes which occur in the global workspace just described.

Many cognitive devices are thought to play some role in the envisaged neural system – including the mechanism responsible for executive processes, working memory, and others. But none of the scholars which are working on that hypothesis assign any role to language as a mean for the integration of information.

7. Sperber: non-cognitive flexibility

There is not enough room left here to discuss the stimulating proposal of Sperber (2005) properly. I am just going to make a single general observation. As we noted, the aim of Sperber is to give a non-cognitive account of flexibility. At first sight, Sperber assigns to "non-cognitive" a quite specific meaning, which does not have any direct relation with the issue we discussed here. Still, if one also looks at the particular mechanisms Sperber evokes for the explanation of flexibility, the impression that his proposal intends to be non-cognitive in the following sense is strong, too: he assumes that the issue of flexibility can be accounted for without presupposing a global space in which mental contents can be accessed "for consideration" (as we said, this is the role Carruthers attributes to the supposer: it generates "new sentences for consideration"). In other words, Sperber's question is how a cognitive system can flexibly select the relevant information through wholly automatic processing, without even considering any explicit alternative. On the contrary, the issue we have been discussing here is how flexibility can be enhanced by controlled processes, through which subjects inhibit automatic reactions, try to imagine new possible solutions, and then compare those potential solutions both with the intended purposes and with each other.

My guess is that a sensible and complete account of the specific kind of flexibility exhibited by humans could not be given without considering the role of controlled processes, of which sensory imagination is but one particular kind.

References

Baars, B.J. (1988). *A cognitive theory of consciousness*. Cambridge, Cambridge University Press.

Baars, B.J. (1997). *In the Theater of Consiousness: The Workspace of the Mind.* Oxford, Oxford University Press.

Baars, B.J. (2002). The conscious access hypothesis: origins and recent evidence. *TRENDS in Cognitive Sciences*, 6 (1), 47-52.

Baars, B.J. and S. Franklin (2003). How conscious experience and working memory interact. *TRENDS in Cognitive Sciences*, 7 (4), 166-172.

Barsalou, L.W., Simmons, W.K., Barbey A.K. and Wilson C.D. (2003). Grounding conceptual knowledge in modality-specific systems. *TRENDS in Cognitive Sciences*, 7 (2), 84-91.

Carruthers, P. (2005). Distinctively Human Thinking. In *The Innate Mind: Struc-*

ture and Content. P. Carruthers, S. Laurence and S. Stich (eds.), Oxford, Oxford University Press.
Churchland, P.S. (2002). *Brain-Wise. Studies in Neurophilosophy*. Cambridge (Mass.), MIT Press.
Dehaene, S. and Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79, 1-37.
Fodor, J. (1975). *The Language of Thought*. New York, Crowell.
Fodor, J. (1987). *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*. Cambridge (Mass.), MIT Press.
Heinrich, B. (2000). Testing insight in ravens. In *The Evolution of Cognition*. C. Heyes and L. Huber (eds.), Cambridge (Mass.), MIT Press.
Köhler, W. (1921). *Intelligenzprüfungen an Menschenaffen*. Berlin, Springer.
Kosslyn, S. M. (1994). *Image and Brain*. Cambridge (Mass.), MIT Press.
Mazzone, M. (2004). Proto-concepts: on non-conceptual content of perception. *Cahiers Ferdinand de Saussure*, 57, 145-60.
Mellet, E., Petit, L., Mazoyer, B., Denis, M. and Tzourio N. (1998). Reopening the Mental Imagery Debate: Lessons from Functional Anatomy. *Neuroimage*, 8, 129-139.
Prinz, J. (2002). *Furnishing the Mind. Concepts and Their Perceptual Basis*. Cambridge (Mass.), MIT Press.
Schneider, W. and Chein, J. M. (2003). Controlled & automatic processing: behavior, theory, and biological mechanisms. *Cognitive Science*, 27, 525-559.
Shepard, R.N. and Metzler, J. (1971). Mental Rotation of Three-dimensional Objects. *Science*, 171, 701-3.
Sperber, D. (2005). Modularity and relevance: How can a massively modular mind be flexible and context-sensitive? In *The Innate Mind: Strucuture and Content*, P. Carruthers, S. Laurence and S. Stich (eds.), Oxford, Oxford University Press.
Sperber, D. and Wilson, D. (1996). Fodor's Frame Problem and Relevance Theory (reply to Chiappe & Kukla). *Behavioral and Brain Sciences*, 19 (3), 530-532.

Marco Mazzone
Università di Catania
mazzonem@unict.it

Michael J. Proulx - Petra Stoerig
University of Düsseldorf

Seeing sounds and tingling tongues: Qualia in synaesthesia and sensory substitution

Abstract In this paper we wish to bring together two seemingly independent areas of research: synaesthesia and sensory substitution. Synaesthesia refers to a rare condition where a sensory stimulus elicits not only the sensation that stimulus evokes in its own modality, but an additional one; a synaesthete may thus hear the word "Monday", and, in addition to hearing it, have a concurrent visual experience of a red color. Sensory substitution, in contrast, attempts to substitute a sensory modality that a person has lost by transforming the information it provided so that it can be accessed through another, intact sensory modality. To make visual information accessible to a blind person, for example, data taken by a camera would be transformed into tactile or auditory information. What do synaesthesia and sensory substitution have in common? Research in both of these areas contributes to our understanding both of cross-modal cooperation and of sensory sensations or qualia, asking under what circumstances these can arise in a modality that is not stimulated. Synaesthesia reveals that this "sensory cross-activation" is possible, and sensory substitution research hopes to induce it. In this article, we will review briefly the literature on synaesthesia, and discuss the issue of qualia for this domain of research. We will then address the evidence for synaesthesia and visual qualia in the blind, and the research on sensory substitution, to finally ask whether sensory substitution may induce a 'synthetic' form of synaesthesia by taking advantage of the nervous system's capacity for generating visual images in the absence of retinal input.

Key words Synaesthesia, sensory substitution, qualia, auditory, visual, tactile, crossmodal.

Qualia in synaesthesia

The term 'synaesthesia' derives from the Greek 'syn' (together) and 'aisthesis' (perception), and originally means 'perceiving together'. When we see a dog and hear him barking, we thus perceive his image together with the sounds he emits, combining inputs from two sensory modalities. In its more restricted, now common meaning, however, synaesthesia refers to a condition where a physical stimulus induces a sensation not only in the modality whose receptors it stimulates, but in an additional one as well. The stimulus that triggers the synaesthetic sensation is the *inducer*, and the extra sensations are called *concurrents.*

Although Marks (1975) reviewed reports that perhaps date back to Pythagorus and Aristotle, and Larner (2006) has uncovered evidence of a possible report of synaesthesia from the seventeenth century, many reviews date the first investigations of synaesthesia back to Galton (1880). Galton focused his investigation on the relation between numbers and mental imagery, and described correspondence from individuals who see black numbers (inducers) as colored (concurrents), such as "3" looks yellow and "4" looks red. Although this color-grapheme form of synaesthesia, where a number or letter is seen in a particular color, is probably the most common form of synaesthesia, many other types exist. Marks (1975) focused his review on colored-hearing synaesthesia, where sounds induce the experience of color (see also Harrison & Baron-Cohen, 1995). Associations between flavors and shapes (Cytowic, 1993; Ramachandran & Hubbard, 2003), between hearing and taste sensations (Beeli, Esslen & Jäncke, 2005), and between emotions and colors have all been documented in the literature, and it seems that almost any type of combination of inducer and concurrent is possible. Note however that even synaesthetes associating the same modalities do not have the same concurrent sensations evoked by the same stimulus. One synaesthete may describe a "3"as transparent yellow, another as reddish brown, showing that the associations are both idiosyncratic and very specific.

What are the concurrent sensations like? Do they resemble those that non-synaesthetes perceive when the appropriate modality is directly stimulated? Would the word 'hog', printed in oxblood red letters, evoke the same sensation in a non-synaesthete as the same word, only printed in black, evokes in a color-grapheme synaesthete who reports that 'hog' induces the experience of oxblood red? According to a distinction drawn by Dixon, Smilek and Merikle (2004), only a *projector* synaesthete experiences the concurrent color as if the grapheme was overlaid by that color. The word 'hog' would thus appear as if a somewhat transparent version of the same word was laid on top of the printed one, still allowing the synaesthete to see the black print of the grapheme. Although one would thus see both the color of the print and the concurrent one, the latter too would appear as actually present in the external world.

In contrast, an *associator* synaesthete describes her color experience as having an internal quality. Dixon et al. (2004) draw an analogy to a normal individual's experience of seeing a black-and-white photograph of a stop sign. Although one does not see the color red in the picture, one may have an internal experience of red evoked by the picture. Thus these synaesthetes have no external experience of the concurrent color, but rather see the word in the color of its print. Whether this division between types of synaesthesia runs on a continuum, such as the ability to produce vivid visual imagery does, is an open question for future research. For example, it might be the case that projector synaesthetes are in fact just experiencing a stronger version of the induced color experience than the associator synaesthetes, which leads to a specific percept of color, and consequences associated with that percept (Dixon et al. 2004). The distinction between projector and associator synaesthetes might thus correspond to the "higher" and "lower" manifestations of synaesthesia described by Ramachandran and Hubbard (2001b). Similarly, the experience of a color induced by a sound, such as a musical note, may vary between individuals. Unlike colored-grapheme synaesthesia, there is no external, visual object to be colored in colored hearing. What is the quality of the visual experience of color that is induced by sound? Ward, Huckstep and Tsakanikos (2006; see also Zigler, 1930) describe some synaesthetes who experience visual shapes that are colored according to the sound they hear, and others that just experience a general color sensation, perhaps similar to the associator synaesthetes described by Dixon et al. (2004). It thus appears as if the perceptual density of the concurrent qualia may differ. Nonetheless the concurrent qualia are commonly of the kind that non-synaesthetes perceive when the concurrent modality is directly stimulated.

On the automaticity and perceptual reality of the synaesthetic experience

Research on synaesthesia has evolved from merely describing cases, as with Galton (1880), to testing the consistency, perceptual reality and automaticity of the synaesthetic experience. This shift has taken place recently, although tests of consistency of one's synaesthetic experience through test-retest means was already used by Starr (1893). Taken up again by Baron-Cohen, Wyke, & Binnie (1987), and recently revised by Asher, Aitken, Farooqi, Kurmani, and Baron-Cohen (2006), this 'Test of Genuineness' distinguishes synaesthetes from their non-synaesthetic brethren by asking both for associations with inducers, and finding identical responses that average from 70 to 90% across time delays ranging from months to years, and without prior warning of testing, in the synaesthetes. This is about two- or threefold the average replication rate seen in non-synaesthetes, and shows that associations are commonly very stable over time. However, as pointed out by Ward and Mattingly (2006), 'genuineness' of synaesthesia would not need to be reflected in such stability of associations: A person who at different points in time experiences different extra sensations in response to the same stimulus would still be a synaesthete, if of an 'inconstant concurrent' kind.

Such a person would fail the test-retest 'test', but could still pass the test devised to tackle the perceptual reality of the synaesthetic experience. Ramachandran and Hubbard (2001a) first demonstrated that synaesthetic colors resulted in perceptual grouping in two grapheme-color synaesthetes, which suggested that the experience of the concurrent colors resembled that of real grouping by color. They also found that when the inducing letter or number was placed far in the synaesthete's peripheral vision, the synaesthetes had no experience of color even though they could still recognize the identity of the inducer. The experience of color thus was not just a metaphor or a memorized association, but rather quite similar to normal color perception. Later work by Palmeri, Blake, Marois, Flanery and Whetsell (2002) found that the synaesthetically-experienced colors could affect the speed by which one visually searched for a grapheme. Although visual search for a digital clock style *2* among *5*s is normally inefficient, and takes longer with each distractor *5* added to the display, a synaesthete with different colors induced by the *2s* and *5s* could find the target more rapidly than a non-synaesthete. The number *2* would pop-out of the display and be found quickly, much as one could easily find a singular red flower in a field of green grass (cf. Treisman & Gelade, 1980). There is even evidence that synaesthetically-induced colors perceptually group with real-colored items, suggesting that synaesthetic colors are indeed experienced in the external world, much as actual printed colors are (Kim, Blake & Palmeri, 2006).

Synaesthetic sensations can thus be experienced consistently and color one's perception of the world. But how automatic is the experience of color that is induced by a grapheme? Recent research has approached this question in a variety of ways. The most popular is to use a Stroop interference paradigm, where the synaesthetic experience of color, if automatic, can help or hinder one's performance on a task. For example, Palmeri et al. (2002) presented words to a synaesthete that either appeared in the color that the words evoked for that synaesthete or in a different color. The task was to name the color of the letters on the screen, not the induced concurrent color. For congruent trials, the synaesthete saw the word "moose" with pink letters, which is the same color that the word "moose" induces in the synaesthete. For incongruent trials, the word would be shown in a different color, such as green. The synaesthete was significantly slower to name the color of the letters on incongruent trials than on congruent trials, just as non-synaesthetes are slower to read color words that are printed in colors different from the one the color names (cf. Stroop, 1935). Just like the printed colors, those evoked by the words were automatically induced and conflicted with the perception of the actual color of the letters in the synaesthete. Others have demonstrated Stroop-like interference for colored-hearing synaesthetes as well (Ward et al., 2006). Note, however, that the automaticity these findings suggest does not imply that the concurrent sensation is evoked pre-attentively; after all, the tasks used all require attention being deployed to the target words or graphemes (Sagiv, Heer, & Robertson, 2006).

All of the above studies are concerned with what is called *developmental* synaesthesia (Grossenbacher & Lovelace, 2001). Developmental synaesthesia may have a genetic basis, as indicated by studies on its incidence in families, and especially monozygotic twins (Baron-Cohen, Burt, Smith-Laittan, Harrison, & Bolton, 1996), as well as a possible difference in its prevalence among males and females, with a higher prevalence in females (0.087% of females and 0.014% of males; Rich, Bradshaw & Mattingley, 2005). Developmental synaesthetes do not remember ever not having had their concurrent sensations.

Synaesthesia arising from visual deprivation

However, synaesthesia can also be acquired. Hallucinogenic drugs often induce synaesthesia, and, of even greater interest in our context, sensory deprivation can have a similar effect. Sensory deprivation, whether temporary, as through blindfolding of normal sighted volunteers, or permanent, as through sensory deafferentation, very often induces hallucinations in the modality deprived of its input (Merabet et al., 2004; Bonnet, 1760). In addition to such release phenomena, Armel and Ramachandran (1999) described a patient who gradually lost his sight due to degeneration of the retina. A couple of years after becoming totally blind, he began to have visual experiences induced by tactile sensations. Blind synaesthetes have also been studied by Jacobs, Karpik, Bozian, and Gothgen (1981) and Steven and Blakemore (2004), although the latter authors stress that their six subjects had developmental synaesthesia that persisted despite continued blindness. In all cases, colors were the concurrent sensations which were most often elicited by letter phonemes, numbers, and days of the week as well as months; only one also had colors for Braille. Descriptions of colored hearing in the blind go back to the late 19th century (Galton, 1880; Phillipe, 1893; Starr, 1893), and continued through the peak period of interest in synaesthesia in the 1920s (e.g. Wheeler and Cutsforth, 1921; Cutsforth, 1924; Voss, 1929). Even if some of these cases may have been of the developmental kind, sensory deprivation is likely to have contributed. If so, synaesthetes should be much more common among the blind, deaf, or numb, than among the sighted, hearing, and feeling. Indeed, in the latter the overall prevalence is estimated at a variable but rather low 0.02 to 5% of the population (Baron-Cohen et al., 1996; Rich et al., 2005, Simner et al., in press), while Phillipe (1893) estimated its incidence among the blind as ~33%!

Reports of synaesthesia in patients with central rather than peripheral neuropathology are rarer. Vike, Jabbari & Maitland (1984) reported that a patient with a large tumor in the left medial temporal lobe experienced visual phenomena in response to sound. The induced visual experience was only in response to sounds presented to the left ear, and only appeared in the left eye, both ipsilateral to the tumor. When the tumor was removed operatively, the synaesthetic experiences ended.

Acquired synaesthesia in patients with brain tumors and in the blind arose as a result of different types of nervous system modification: one central and one peripheral. However the hypothesis in both cases was that sensory deafferentation could have possibly resulted in the experience of synaesthesia. In the case of the blind patient, the lack of visual input that might normally override auditory-related response in visual cortex could have given rise to the visual experience of sounds. In the case of the patient with a brain tumor, who had an intact visual pathway, the tumor may have blocked some non-critical visual processing pathway that allowed the normally suppressed auditory-related responses to give rise to that patient's visual experience of sounds. Although developmental synaesthesia might arise through different means, these cases do provide the hypothesis that synaesthesia from drugs or pathology employs normal neural connections that somehow become disinihibited and produce experiences that normally would be suppressed (Grossenbacher & Lovelace, 2001).

Visual qualia in the blind

Experiences associated with the deprived modality, both in the form of concurrent sensations and hallucinations, are amply documented both in patients with peripheral and central pathology. Concurrent synaesthetic sensations tend to be simple in appearance, and often take the form of color patches, although quite complex three-dimensional shapes have also been reported. Hallucinations that occur frequently both in patients with peripheral and central blindness range from simple photisms to complex objects (Kölmel, 1985), and even include people moving through the lost visual field (Gloning, Gloning, & Hoff, 1967). Their appearance seems linked to specific cortical regions whose direct electrical stimulation during neurosurgery also generates phenomena of different complexity and kind (Wieser, 2003). Evidence from functional neuroimaging supports this inference both for hallucinations (ffytche et al., 1998) and for synaesthetic concurrents (Nunn et al., 2002; Paulesu et al., 1995). Moreover, Transcranial Magnetic Stimulation (TMS involves the use of a non-invasive, strong magnetic field to modulate cortical activity) applied over the visual cortices can induce visual phosphenes (Cowey & Walsh, 2000), and so do prostheses implanted in the visual cortex (Dobelle, 2002), the optic nerve (Duret et al., 2006) or the retina (Gekeler, Messia, Ottinger, Bartz-Schmidt, & Zrenner, 2006). Together, these data indicate that electromagnetic stimulation (phosphenes), endogenous activation (hallucinations), and activation induced by stimulation of an alternative intact sensory pathway (concurrents) can induce conscious visual sensations in the blind who have lost sight due to loss of retinal input or to lesions of the primary visual cortex. The patients' lesioned visual system must thus retain the potential to generate visual qualia.

Normal synaesthesia

All of us possess extensive connections between the different sensory modalities. They are found in cortical as well as sub-cortical structures of our brain, and induce the facilitating effects that stimulation with corresponding images and sounds – image of a dog and sound of a bark, image of a bird and sound of its song – exert both at the level of the neuronal and the behavioural responses (Stein and Meredith, 1990). Some researchers have examined the possibility of 'natural' cross-modal mapping extending past the object level, such as a dog barking, to simple features, such as hearing a musical note and seeing a flash of light. Our language often suggests such equivalences, as when a frequency is described as either high or low pitch. To learn whether the terms "high" and "low" influence one's perception of space, Walker and Smith (1984) used a variant of the Stroop test to examine how incidental tones that were either high or low pitch affected a lexical decision task. Subjects were to respond only to words shown on a computer screen. If the words "up" or "top" appeared, they were to press one key; if the words "down" or "bottom" appeared they were to press another. The tones played simultaneous with the appearance of the words, however, were to be ignored. Walker and Smith found that subjects were quicker to respond when the auditory tone and the word were congruent (that is, a high pitch occurred at the same time as the word "up"), and slower to respond when the tone and words were incongruent (that is, a low pitch occurred at the same time as the word "up"). This supports the idea that the labeling of an auditory pitch as high or low is not just metaphorical, but actually has an automatic impact on one's perception of the world. Importantly, the subjects were not synaesthetes, and this effect seems to derive from some crosswiring between auditory pitch and spatial representation that exists normally. Mudd (1963) performed a qualitative study to examine the natural correspondences that exist between a variety of auditory properties and visual space. Besides also noting the relation between auditory pitch and vertical location, he also reported the same relationship for auditory intensity; that is higher pitches and louder sounds were associated with higher vertical locations, but lower pitches and quieter sounds were associated with lower vertical locations. A recent report on 'implicit synaesthesia', a term coined to describe people who rank high on the 'Test of Genuineness' but do not report concurrent sensations (Steven, Hansen, & Blakemore, 2004), may suggest that normal, implicit, associator, and projector synaesthesia are spaced out over a continuum at whose end the concurrent sensation is explicitly phenomenal. If we all share the propensity for natural crossmodal synaesthesia, its basic mapping dimensions could possibly be exploited in the context of sensory substitution.

Sensory substitution for the blind

Sensory substitution for the blind converts visual information so that it can be processed by an intact sensory pathway. One of the first devices converted the signals from a video camera into tactile stimulation applied to the back of the subject who was seated in a chair designed for this purpose (Bach-y-Rita, Collins, Saunders, White & Scadden, 1969). Recent advances have allowed much smaller devices to provide the tactile information, most notably via stimulation of the tongue or forehead (for a review, see Bach-y-Rita & Kercel, 2003). The auditory modality rather than the somatosensory one was first targeted by Meijer (1992). His system, dubbed The vOICe, is mobile, and in addition to a video camera providing the visual input requires a small computer running the conversion program, and stereo headphones to provide the resultant sound patterns to the user. Meijer's program uses three major principles: horizontal location is coded by stereo panning and the time provided by the left-to-right scanning transformation of each image; vertical location is coded by frequency, so that up is represented by high frequencies and down by low frequencies; pixel brightness is coded by loudness, such that a bright white pixel is heard at maximal volume, and a dark pixel is silent. A more recently developed device, called the Prosthesis Substituting Vision with Audition (PSVA; Capelle, Trullemans, Arno, & Veraart, 1998), uses somewhat similar transformations for vertical location as well as brightness, but magnifies the center of the rather small image to simulate the visual system's emphasis on the macular region of the retina, and uses different tones to provide horizontal location directly, rather than through the left-to-right scan that The vOICe employs. The third substitution device is similar to The vOICe in terms of how vertical and horizontal locations are coded, but differs in the style of presentation and in how much pre-processing of the image takes place. SmartSight (Cronly-Dillon, Persaud, & Gregory, 1999; Cronly-Dillon, Persaud, & Blore, 2000) also codes vertical location by pitch; however, it presents the information in terms of musical notes, with the center of the image corresponding to middle C. Horizontal information is coded by time from left-to-right. The musical quality of the sound patterns is the most obvious difference when compared with The vOICe. Whereas The vOICe and PSVA depend on the ability of the user to learn to associate the complex sound patterns with objects and features in the visual world, SmartSight divides the image into different features. For example, a user could first listen to only the vertically-oriented edges, and subsequently listen to the horizontally-oriented edges, to then put the vertical and horizontal edge locations together to understand what the pieces combine to create.

All of these devices, somatosensory as well as auditory, have been tested on blindfolded and/or blind subjects, and the studies to date show that subjects can learn to interpret the information, and distinguish patterns of dots (Arno, Capelle, Wante-Defalque, Catalan-Ahumada, & Veraart, 1999) , orientations of Ts (Kupers et al., 2006), geometric shapes (Stoerig et al., 2004), and acquire knowledge of the spatial location of objects (Auvray, Hanneton, & O'Regan, 2003; Proulx, Stoerig, Ludowig & Knoll, 2006); even real objects (Auvray, Hanneton, Lenay, & O'Regan, 2005) or images of objects and scenes (Cronly-Dillon et al., 1999, 2000; Stoerig et al., 2004) seem to become classifiable or identifiable with training or extended use of the device (Poirier, Richard, Tranduy, & Veraart, 2006; see also Bach-y-Rita and Kercel, 2003).

Recordings of the brain activity associated with learning to use the devices has shown that visual cortices seem to become recruited into the analysis of the novel inputs especially in the blind (Arno et al., 2001;Kupers et al., 2006; Ptito, Moesgaard, Gjedde, & Kupers, 2005). They thus agree with previous studies that demonstrate that blind subjects activate even early visual cortex when involved in tasks ranging from reading Braille (Sadato et al., 1996), or sound localization (Weeks et al., 2000), to solving verbal memory tasks (Amedi et al., 2003). Disruption of visual cortical activity, whether by means of Transcranial Magnetic Stimulation applied over the occipital lobe (Cohen et al., 1997; Cohen et al., 1999; for a review see Floel & Cohen, 2006) or by a brain lesion (Hamilton, Keenan, Catala, & Pascual-Leone, 2000) interferes with the processing of these non-visual tasks, indicating that these cortices, when deprived of their normal retinal input, do in fact play a functional role in non-visual problem solving. It is important to note that synaesthetic or veridical visual sensations have generally not been reported in these studies, and this unfortunately indicates that visual cortical activation as such cannot be used to attest the presence of visual qualia. However, the fact that visual (V1, V2, V4, V8) and parietal cortical activation was also found in a late-blind synaesthete when he listened to time-words that induced spatially localized colors (Steven, Hansen & Blakemore, 2006) demonstrates that even the brain of a late-blind individual – he had been blind for 10 years – can still employ these areas to generate visual qualia in response to certain words. Conceivably, the visual cortex of the early-blind helps them to excel in a variety of non-visual tasks; in fact, several studies have demonstrated that the visual cortices of the early-blind appear to be more involved in these tasks than in the late-blind (Cohen et al., 1999). However, by virtue of not having been appropriately stimulated by retinal input, the visual cortex of the early-blind may not develop or maintain its propensity to provide phenomenal visual representations. In contrast, people who lose sight later in life may respond to visual deprivation with synaesthetic or hallucinatory vision.

Goals of visual substitution

By providing information about silent objects that are out of reach, sensory substitution ought to be useful for early and late-blind subjects (who can still hear or feel). As the early-blind have no or relatively short periods of vision, they usually face less problems in daily life than the late-blind. Moreover, the late-blind miss the easy access to information about objects and people outside of extra-personal space that sight provides. Visual substitution devices deliver this information, even though they are still a long way from allowing the ease of vision. A lot of perceptual learning is required for people to learn to understand the meaning of the new sensations, and the early-blind may have an easier time of it if they start using the systems early in life.

Early-blind subjects would probably experience sensations associated with the modality through which the substitution device delivers the information even when they recruit their visual cortex for the processing. A recent study by Kupers and colleagues (2006) supports this prediction. The authors trained early- and late-blind as well as sighted subjects to use a substitution system that converts visual input into electrotactile patterns on the subject's tongue (see Bach-y-Rita & Kercel, 2003). Following successful training, the authors used Transcranial Magnetic Stimulation to examine what sort of qualia would be associated with visual cortex stimulation. When the sighted subjects, who had been trained with the sensory substitution device, received TMS, they experienced visual phosphenes (they saw flashes of light as a result of the activation of their visual cortex), confirming that the visual cortex had not substantially deviated from its established visual role. In contrast, three of the eight early-blind subjects reported sensations on the tongue, but none reported visual phenomena. How about the five late-blind subjects? A tingling on the tongue occurred in only one, but two reported visual sensations, indicating that sensory substitution may in fact be able to induce 'synthetic' synaesthesia with concurrent sensations derived from the deprived modality.

Normal (developmental and deprivation-induced) synaesthesia does not seem to reflect features of the external world. However, they can be used to identify something that is rendered invisible by presenting it in the visual periphery under conditions of 'crowding'. Ramachandran & Hubbard (2001b) asked their volunteers to fixate on a central spot, and presented a single digit off to one side. All subjects found it easy to identify the digit. However, when it was surrounded by four other numbers ('crowding'), one on each side, identification of the target fell to chance – in non-synaesthetes. The two color-grapheme synaesthetes they tested, however, still reported seeing the color induced by the target number, and correctly inferred its identity on the basis of this color. Their remarks, such as "I can't see that middle letter but it must be an 'O' because it looks blue" (p. 8, Ramachrandran & Hubbard, 2001b), and their performance demonstrate that synaesthetic sensations may in fact be useful. Note that Hubbard, Arman, Ramachandran, & Boynton, 2005, report that not all synaesthetes show this performance benefit in crowding tasks, again attesting to the inter-individual variability of the phenomena.

'Synthetic' synaesthesia

As described above, late-blind people may experience visual sensations in the form of phosphenes or fully-formed hallucinations. That sensory substitution devices can induce visual concurrents has been shown by Kupers et al. (2006) in two of their five late-blind participants. In addition to the pleasure visual sensations can give to the blind (Cowey & Walsh, 2000), and the enrichment of the perceptual world they cause, these concurrents may be useful *if* they can provide information about the external world. To serve this purpose, the concurrents would not need to correspond to what a sighted person may see. As shown by the 'crowding' experiment cited above, a concurrent may substitute for otherwise unavailable information provided it is firmly associated with a particular stimulus. Say if a round object always elicited an experience of green, while angular objects do not, then seeing green would carry veridical shape information.

Could it also be possible to induce concurrent sensations that correspond to the world yet more directly? Although visual substitution may well be successful without inducing synthetic synaesthesia, let alone a veridical one, we and others, like Peter Meijer, who developed The vOICe, are pursuing this work in the hope that it will prove to be possible. As yet, the negative evidence is strong. Research has shown activity in primary visual cortex in the brains of blind or blindfolded patients doing a tactile task has not been accompanied by concurrent visual qualia that is evoked by the tactile sensations (Sadato et al., 1996; Pascual-Leone & Hamilton, 2001). Studies of sensory substitution that demonstrated successful perceptual learning as well as visual cortical activation in blind subjects have not reported any visual sensations on the part of their subjects (Arno et al., 1999; Cronly-Dillon et al., 1999, 2000; Poirier et al., 2006; Ptito et al., 2005), and our recent work with sensory substitution has found that the ability to interpret the sounds created by image-to-sound conversion can be accounted for by the activity in auditory cortex alone (Pollok, Schnitzler, Stoerig, Mierdorf, & Schnitzler, 2005). However, there is also some positive evidence. Kupers et al. (2006) described tingling tongues, but, in the late-blind, some visual sensations from visual cortical stimulation as well, although their subjects had received only basic training in sensory substitution over a short period of time. This is remarkable because the two or three weeks commonly used for formal training and testing are too short to allow even the automatic reading-out of the system's information that Bach-y-Rita, pioneer of this line of research, invokes. Clearly, such automaticity is necessary to free the sensory modality used to provide visual input for its primary – hearing or feeling – tasks, and a lot of perceptual learning must go into establishing it. That time also plays an important part for visual substitution to induce visual concurrents is suggested by a late-blind user of The vOICe system. She has used it on a regular basis for several years, and described instances when she not only experienced visual phenomena, but actually saw what she was hearing (Fletcher, 2002). We have conducted research with a sighted subject who was blindfolded for 21 days during which he used the vOICe system. He too reported a few instances where saw ‘like a hallucination’ what he was hearing through the device, and correctly described it. In his case, the propensity of the deprived system to provoke visual hallucinations may have contributed to lending the information its visual gestalt (see also Merabet et al. 2004).

Whether these are isolated instances, or something that can be permanently ingrained with training requires much more research. Extended use of the device may be one necessary condition, and a visual system still prone to generate visual qualia, another. Third, substitution systems that take advantage of natural correspondences between vision and audition or touch (e.g., Walker & Smith, 1984) and allow the other sensory systems to exploit their natural connections with the visual system, may not only facilitate the extensive perceptual learning, but the induction of synthetic synaesthesia that provides qualia for the missing sensory modality as well. Sensory substitution challenges the brain’s plasticity. Should it, given time and training, be able to learn to generate visual concurrents that correspond to the auditory or tactile information the substitution system provides, the future will not only hold tingling ears and tongues for the blind, but look distinctly bright.

Acknowledgements

The writing of this review, and some of our research that we describe, was supported in part by the Volkswagen Stiftung (I/80 742) and the Anton-Betz-Stiftung.

References

Amedi, A., Raz, N., Pianka, P., Malach, R., & Zohary, E. (2003). Early 'visual' cortex activation correlates with superior verbal memory performance in the blind. *Nature Neuroscience*, 6, 758-766.

Armel, K. C., & Ramachandran, V. S. (1999). Acquired synesthesia in retinitis pigmentosa. *Neurocase*, 5, 293-296.

Arno, P., Capelle, C., Wanet-Defalque, M.C., Catalan-Ahumada, M., & Veraart, C. (1999). Auditory coding of visual patterns for the blind. *Perception*, 28, 1013-1029.

Arno, P., De Volder, A.G., Vanlierde, A., Wanet-Defalque, M.C., Streel, E., Robert, A., Sanabria-Bohorquez, S., & Veraart, C. (2001). Occipital activation by pattern recognition in the early blind using auditory substitution for vision. *Neuroimage*, 13, 632-645.

Asher, J. E., Aitken, M. R. F., Farooqi, N., Kurmani, S., & Baron-Cohen, S. (2006). Diagnosing and phenotyping visual synaesthesia. A preliminary evaluation of the revised test of genuineness (TOG-R). *Cortex*, 42, 37-

146.

Auvray, M., Hanneton, S, Lenay, C.., & O'Regan, J.K. (2005). There is something out there. Distal attribution in sensory substitution, twenty years later. *Journal of Integrative Neuroscience*, 4, 505-521.

Auvray, M., Hanneton, S., & O'Regan, J.K. (2003). Localisation and form recognition in sensory substitution. *Perception*, 32 supplement, 94.

Bach-y-Rita, P., & Kercel, S. W. (2003). Sensory substitution and the human-machine interface. *Trends in Cognitive Sciences*, 7, 541-546.

Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B. & Scadden, L. (1969). Vision substitution by tactile image projection. *Nature*, 221, 963-964.

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., & Bolton, P. (1996). Synaesthesia, Prevalence and familiality. *Perception*, 25, 1073-1079.

Baron-Cohen, S., Wyke, M. A., & Binnie, C. (1987). Hearing words and seeing colours. An experimental investigation of a case of synaesthesia. *Perception*, 16, 761-767.

Beeli, G., Esslen, M., Jäncke, L. (2005). When coloured sounds taste sweet. *Nature*, 434, 38.

Bonnet, C. (1760). *Essai analytique sur les facultés de l'âme*. Geneva, Slatkine.

Capelle, C., Trullemans, C., Arno, P., & Veraart, C. (1998). A real-time experimental prototype for enhancement of vision rehabilitation using auditory substation. *IEEE Transaction on Biomedical Engineering*, 45, 1279-1293.

Cohen L. G., Celnik P., Pascual-Leone A., Corwell B., Falz L., Dambrosia, J., Honda, M., Sadato, N., Gerloff, C., Catala, M. D., & Hallett, M. (1997). Functional relevance of cross-modal plasticity in blind humans. *Nature*, 389, 180-183.

Cohen, L.G., Weeks, R.A., Sadato, N., Celnik, P., Ishii, K., & Hallett, M. (1999). Period of susceptibility for cross-modal plasticity in the blind. *Annals of Neurology*, 45, 451-460.

Cowey, A., & Walsh, V. (2000). Magnetically induced phosphenes in sighted, blind and blindsighted observers. *Neuroreport*, 11, 3269-3273.

Cronly-Dillon, J., Persaud, K., & Gregory R. P. F. (1999). The perception of visual images encoded in musical form, A study in cross modality information transfer. *Proceedings of the Royal Society of London Series B – Biological Sciences*, 266, 2427-2433.

Cronly-Dillon, J., Persaud, K. C., & Blore, R. (2000). Blind subjects construct conscious mental images of visual scenes encoded in musical form. *Proceedings of the Royal Society of London Series B – Biological Sciences*, 267, 2231-2238.

Cutsforth, T. D. (1924). Synaesthesia in the Process of Reasoning. *American Journal of Psychology*, 35, 88-97.

Cytowic, R. E. (1993). *The Man Who Tasted Shapes.* London, Abacus.

Dixon, M. J., Smilek, D., & Merikle, P. M. (2004). Not all synaesthetes are created equal. Projector versus associator synaesthetes. *Cognitive, Affec-*

tive, & Behavioral Neuroscience, 4, 335-343.

Dobelle, W. H. (2002). Artificial vision for the blind by connecting a television camera to the visual cortex. *American Society for Artificial Internal Organs*, 46, 3-9.

Duret, F., Brelén, M. E., Lambert, V., Gérard, B., Delbeke, J., & Veraart, C. (2006). Object localization, discrimination, and grasping with the optic nerve visual prosthesis. *Restorative Neurology and Neuroscience*, 24, 31-40.

Ffytche, D. H., Howard, R. J., Brammer, M. J., David, A., Woodruff, P., & Williams, S.(1998). The anatomy of conscious vision, an fMRI study of visual hallucinations. *Nature Neuroscience*, 1, 738-742.

Fletcher, P. (2002). Seeing with sound. A journey into sight. Paper presented at the Toward a Science of Consciousness meeting, Tucson, Arizona.

Floel, A., & Cohen, L.G. (2006), October 4. Contribution of noninvasive cortical stimulation to the study of memory functions. *Brain Research Reviews* (DOI 10.1016/j.brainresrev.2006.08.006).

Galton, F. (1880). Visualised numerals. *Nature*, 22, 494-495.

Gekeler, F., Messias, A., Ottinger, M., Bartz-Schmidt. K., U., & Zrenner, E. (2006). Phosphenes electrically evoked with DTL electrodes, A study in patients with Retinitis Pigmentosa, Glacoma, and Homonymous visual field loss and normal subjects. *Investigative Ophthalmology & Visual Science*, 47, 4966-4974.

Gloning, J., Gloning, K., & Hoff, H. (1967). Über optische Halluzinationen. *Wiener Zeitschrift für Nervenheilkunde und deren Grenzgebiete*, 25, 1-19.

Grossenbacher, P. G., & Lovelace, C. T. (2001). Mechanisms of synesthesia. Cognitive and physiological constraints. *Trends in Cognitive Sciences*, 5, 36-41.

Hamilton, R., Keenan, J.P., Catala, M., & Pascual-Leone, A. (2000). Alexia for Braille following bilateral occipital stroke in an early blind woman. *Neuroreport*, 11, 237-240.

Harrison, J., & Baron-Cohen, S. (1995). Synaesthesia, Reconciling the subjective with the objective. *Endeavour*, 19, 157-160.

Hubbard, E. M., Arman, A. C., Ramachandran, V. S., & Boynton, G. M. (2005). Individual differences among grapheme-color synaesthetes, Brain-behavior correlations. *Neuron*, 45, 975-985.

Jacobs, L., Karpik, A., Bozian, D., & Gothgen, S. (1981). Auditory-visual synesthesia, sound induced photisms. *Archives of Neurology*, 38, 211-216.

Kim, C-Y, Blake, R., & Palmeri, T. J. (2006). Perceptual interaction between real and synesthetic colors. *Cortex*, 421, 195-203.

Kölmel, HW. (1985). Complex visual hallucinations in the hemianopic field. *Journal of Neurology, Neurosurgery and Psychiatry*, 48, 29-38.

Kupers, R., Fumal, A., Maertens de Noordhout, A., Gjedde, A., Schoenen, J., & Ptito, M. (2006). Transcranial magnetic stimulation of the visual cortex

induces somatotopically organized qualia in blind subjects. *Proceedings of the National Academy of Sciences of the USA*, 103, 13256-13260.
Larner, A. J. (2006). A possible account of synaesthesia dating from the seventeenth century. *Journal of the History of the Neurosciences*, 15, 245-249.
Marks, L. E. (1975). On colored-hearing synesthesia, Cross modal translations of sensory dimensions. *Psychological Bulletin*, 82, 303-331.
Meijer, P. B. L. (1992). An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering*, 39, 112-121.
Merabet, L.B., Maquire, D., Warde, A., Alterescu, K., Stickgold, R., & Pascual-Leone, A. (2004). Visual hallucinations during prolonged blindfolding in sighted subjects. *Journal of Neuro-Opthalmology*, 24, 109-113.
Mudd, S. A. (1963). Spatial stereotypes of four dimensions of pure tone. *Journal of Experimental Psychology*, 66, 347-352.
Nunn, J. A., Gregory, L. J., Brammer, M., Williams, S. C. R., Parslow, D. M., Morgan, M. J., Morris, R. G., Bullmore, E. T., Baron-Cohen, S., & Gray, J. A. (2002). Functional magnetic resonance imaging of synesthesia, activation of V4/V8 by spoken words. *Nature Neuroscience*, 5, 371-375.
Palmeri, T. J., Blake, R., Marois, R., Flanery, M. A., & Whetsell, W. (2002). The perceptual reality of synesthetic colors. *Proceedings of the National Academy of Sciences of the USA* 99, 4127-4131.
Pascual-Leone, A., & Hamilton, R. (2001). The metamodal organization of the brain. *Progress in Brain Research*, 134, 427-445.
Paulesu, E., Harrison, J., Baron-Cohen, S., Watson, J. D., Goldstein, L., Heather, J., Frackowiak, R. S., & Frith, C. D. (1995). The physiology of coloured hearing. A PET activation study of colour-word synaesthesia. *Brain*, 118, 661-676.
Phillipe, J. (1893). L'audition colorée des aveugles. *Revue Scientifique*, 1, 806-809.
Poirier, C., Richard, M.-A., Tranduy, D., & Veraart, C. (2006). Assessment of sensory substitution prosthesis potentialities in minimalist conditions of learning. *Applied Cognitive Psychology*, 20, 447-460.
Pollok, B., Schnitzler, I., Stoerig, P., Mierdorf, T., & Schnitzler, A. (2005). Image-to-sound conversion, experience-induced plasticity in auditory cortex of blindfolded adults. *Experimental Brain Research*, 167, 287-91.
Proulx, M. J., Stoerig, P., Ludowig, E., & Knoll, I. (2006). Seeing 'where' through the ears? Localization by sensory substitution. Paper presented at the Association for the Scientific Study of Consciousness meeting, Oxford, England.
Ptito, M., Moesgaard, S. M., Gjedde, A., & Kupers, R. (2005). Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind. *Brain*, 128, 606-614.
Ramachandran, V. S., & Hubbard, E. M. (2001a). Psychophysical investigations

into the neural basis of synaesthesia. *Proceedings of the Royal Society of London B*, 268, 979-983.

Ramachandran, V. S., & Hubbard, E. M. (2001b). Synaesthesia - A window into perception, thought and language. *Journal of Consciousness Studies*, 8, 3-34.

Ramachandran, V. S., & Hubbard, E. M. (2003). Hearing colors, tasting shapes. *Scientific American*, 288, 52-59.

Rich, A. N., Bradshaw, J. L., & Mattingley, J. B. (2005). A systematic, large-scale study of synaesthesia, Implications for the role of early experience in lexical-colour associations. *Cognition*, 98, 53-84.

Sadato, N., Pascual-Leone, A., Grafman, J., Ibañez, V., Deiber, M.-P., Dold, G., & Hallett, M. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature*, 380, 526-528.

Sagiv, N., Heer, J., & Robertson, L. (2006). Does binding of synesthetic color to the evoking grapheme require attention? *Cortex*, 42, 232-242.

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S A., Fraser, C., Scott, K., & Ward, J. In press. Synaesthesia, The prevalence of atypical cross-modal experiences. *Perception*.

Starr, F. (1893). Note on color-hearing. *American Journal of Psychology*, 5, 416-418.

Stein, B. E. & Meredith, M. A. (1990). Multisensory integration. Neural and behavioral solutions for dealing with stimuli from different sensory modalities. *Annals of the New York Academy of Sciences*, 608, 51-65.

Steven, M. S., & Blakemore, C. (2004). Visual synaesthesia in the blind. *Perception*, 33, 855-868.

Steven, M. S., Hansen, P., & Blakemore, C. (2004). Implicit Synesthesia, Non-synesthetes with reproducible color-word associations. *Society for Neuroscience Abstracts*, 664, 19.

Steven, M. S., Hansen, P., & Blakemore, C. (2006). Activation of color-selective areas of the visual cortex in a blind synesthete. *Cortex*, 42, 304-308.

Stoerig, P., Ludowig, E., Mierdorf, T., Oros-Peusquens, A., Shah, J. N., Meijer, P. B., & Pascual-Leone, A. (2004). Seeing through the ears? Identification of images converted to sounds improves with practice. *Society for Neuroscience Abstracts*, 177, 15

Stroop, J. R. (1935). Studies of indifference in serial verbal reactions. *Journal of ExperimentalPsychology*, 18, 643-662.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-136.

Vike, J., Jabbari, B., & Maitland, C. G. (1984). Auditory-visual synesthesia, Report of a case with intact visual pathways. *Archives of Neurology*, 41, 680-681.

Voss, W. (1929). Das Farbenhören bei Erblindeten. Untersuchungen über Wesen und Arten der Photismen bei blinden Synoptikern unter besonderer Berücksichtigung des Formproblems. *Archiv für die gesamte Psychologie*,

73, 407-524.
Walker, P., & Smith, S. (1984). Stroop interference based on the synaesthetic qualities of auditory pitch. *Perception*, 13, 75-81.
Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia, To what extent does it use cross-modal mechanisms common to us all?. *Cortex*, 421, 264-280.
Ward, J., & Mattingley, J. B. (2006). Synaesthesia, An overview of contemporary findings and controversies. *Cortex*, 42, 129-36.
Weeks, R., Horwitz, B., Aziz-Sultan, A., Tian, B., Wessinger, C.M., Cohen, L.G., Hallett, M., & Rauschecker, J.P. (2000). A positron emission tomographic study of auditory localization in the congenitally blind. *Journal of Neuroscience* 20, 2664-2672.
Wheeler, R. H., & Cutsforth, T. D. (1921). The Number Forms of a Blind Subject. *American Journal of Psychology*, 32, 21-25.
Wieser, H. G. (2003). Music and the brain. Lessons from brain diseases and some reflections on the 'emotional' brain. *Annals of the New York Academy of Sciences*, 999, 76-94.
Zigler, M. J. (1930). Tone shapes. A novel type of synaesthesia. *Journal of General Psychology*, 3, 277-286.

Michael J. Proulx - Petra Stoerig
University of Düsseldorf
Michael.Proulx@uni-duesseldorf.de
Petra.Stoerig@uni-duesseldorf.de

Frédérique de Vignemont
Institute of Cognitive Science, CNRS

Brainreading of perceptual experiences: a challenge for first-person authority?

Abstract According to a traditional Cartesian view of the mind, you have a privileged access to your own conscious experiences that nobody else can have. Therefore, you have more authority than anybody else on your own experiences. Perceptual experiences are self-intimating: you are aware of what you are consciously perceiving. If you report seeing a pink elephant, nobody is entitled to deny it. There may be no pink elephant, but you do have the conscious experience of such elephant. However, the progress in brain imaging might lead to the possibility that the scientist knows as well as you – or even better than you – what you are seeing, and even what you are hallucinating. What was only a thought experiment fourty years ago (Smart, 1962, Armstrong, 1963; Rorty, 1965) has become reality. Does brain reading challenge the privacy of the mind? Who has the most authority on your mind in case of conflict? You or the brain scientist?

Keywords Brainreading, reduction, privacy, first-person authority, introspection.

1. Introduction

You are in Siena in front of the main square during the Palio delle Contrade. I know what you are seeing because I am also there, next to you, and I can see the same colorful horses and the same beautiful buildings. Although we share the same environment, am I entitled to claim that I have access to your perceptual states? Do I know your experiences as well as I know my own? Obviously not. If rather than looking at the horses, you were paying attention to a pretty woman in the corner, I would not know what you see. You may even have a visual hallucination of a pink elephant running with the horses, without me being able to know it. We may share the same external world, but by no means do we share the same perceptual experiences. A skeptic might go further and argue that I am not even entitled to conclude that you do have experiences. There seems to be a gap between the self and the other such that one could never have a direct access to someone else's mental states. The problem of other minds arises from the Cartesian conception of mentality: mental states are private. One has a privileged direct knowledge only of one's own mental states. One does not have any immediate access to other minds, but only to their public behaviors. One can only infer what the other people think from what one can observe.

This traditional view of the mind has been recently shaken by recent results in neuroscience. Scientists now claim to be able to read in your mind. Using brain imaging techniques, they know what you see, based on your brain activations that they decode. You cannot hide anymore that you are looking at the pretty woman because they can detect it in your brain. Even your visual hallucinations would have no secret for them. Brain scientists could know them as well as you do. But could they know them even better than you? What would happen if you report seeing a pink elephant while they claim that you see a blue giraffe?

What is at stake here is the possibility that someone could know better than you your own perceptual experiences and mental images. If this were true, then it would challenge the Cartesian view of an incorrigible and privileged access to one's own experiences. There are two ways to address this question. The classical way consists in discussing the first-person authority principle and in assessing its validity (Jackson, 1973, Shoemaker, 1996). The analysis of brainreading per se would not then be of any relevance. The (in)corrigibility of introspection is an intrinsic property of the mind. It does not depend on the existence of any external mean to revise first-person judgments. But one can tackle this debate from a different angle. Even if introspection is corrigible in principle, brainreading might not be more reliable than introspection. Is brainreading grounded enough to revise first-person reports? I will focus here on the strength and the limits of brainreading.

2. Brainreading: from thought experiments to neuroimaging

Can one have access to your mental states by decoding your brain activity? Since their discovery, brain imaging techniques like electroencephalograph (EEG) and functional magnetic resonance imaging (fMRI) have improved the knowledge of the mind. Such methods have led some philosophers to imagine that one day, brain scientists could read in our brain and even, that they could know better than us our own mental states. But is brainreading only a thought experiment?

2.1 Thought experiments

More than forty years ago, some philosophers started to argue on the possibility that a brain scientist could read in our brain. It was not so much the technological progress that interested them than the challenge that such possibility would raise for the Cartesian view of the mind. Let us imagine that the subject sincerely reports that she sees vertical lines while the brain scientist claims that this is not true and that the subject has a visual experience of horizontal lines. With whom shall we take side? Could the brain scientist know better than the subject her own perceptual experiences? The answers varied. Armstrong (1963) argued that self-knowledge is not incorrigible, nor privileged:

> We must allow the possibility that somebody else (for example, a brain technician) reaches a true belief about my inner state when I reach a false one. And then what reason is there to deny that the technician is a better authority on my mental states than I am?

As for Rorty (1965), in a paper where he articulated for the first time his eliminative materialism, he even imagine that children could learn what they feel by looking to a portable scanner:

> Suppose, in particular, that we find it convenient to speed up the learning of contrastive observation predicates (such as "painful", "tickling", etc) by supplying children with portable encephalographs-cum-teaching-machines which, when ever the appropriate brain-process occurs, murmur the appropriate terms in their ears.

We can contrast Armstrong's and Rorty's view with Baier's and Smart's hypothesis:

> To say that one day our physiological knowledge will increase to such an extent that we shall be able to make absolutely reliable encephalograph-based claims about people's experiences, is only to say that, if carefully checked, our encephalograph-based claims about 'experiences' will always be correct, i.e. will make the same claims as a truthful introspective report […] However good the evidence may be, such a physiological theory can never be used to show the sufferer that he was mistaken in thinking that he had a pain, for such mistake is inconceivable. The sufferer's epistemological authority must therefore be better than the best physiological theory can ever be (Baier, 1962).

Smart (1962) went even further:

> I must, I think, agree with Baier that if the sort of situation which we have just envisaged did in fact come about, then I should have to reject the brain process thesis and would perhaps espouse dualism.

We will see that the sort of situation just described in the previous quotations has "in fact come about". Shall we espouse dualism with Smart? Why does such possibility feel so threatening?

2.2 Privacy of the mind

What is at stake here is the question of the privacy of conscious experiences. Privacy can be interpreted in four different ways (Ayer, 1963):

(1) My mental states can be detected only by me.
(2) There is a type of access that I am the only one to have.
(3) I am the best authority on my mental states.
(4) I cannot share them with anybody.

Brainreading does not threaten the privacy of the mind at all levels. For instance, it does not pretend to provide the same kind of access that one has toward one's mental states: the brain scientist looks at the brain activity recorded by a scanner, while the subject looks introspectively at her own mental states. Nor does the brain scientist experience your perceptual states as if he were you. The first interpretation is more relevant for us: brainreading would allow an individual to detect someone else's mental states. In this sense, our conscious experiences would not be private. However, this interpretation is too wide. One can detect the mental states of other people even without brain imaging. This is what is at the core of mindreading (or theory of mind). For example, I can know that you feel jealous based on the way you look at your husband. I do not need to decode your brain activity. What really distinguishes brainreading from mindreading is its bearing for first-person authority (3). Mindreading will always be less reliable than introspection to detect conscious experiences. But what about brainreading?

Two things must be said about first-person authority. First, one should not confuse incorribility and infallibility. Self-knowledge is infallible if and only if I cannot be mistaken about my own mental states. Self-knowledge is incorrigible if and only if nobody can show me that I am mistaken about my mental states, even I am indeed mistaken. First-person authority makes only the latter claim. According to such view, I am the best authority on my mental states, which does not imply that I am a perfect authority. Nobody is epistemically entitled to revise my introspective judgments. If I say that I see a pink elephant, one can tell me that there is no pink elephant, but not that I do not have a visual experience of a pink elephant. Second, if there is indeed first-person authority, then it holds only for conscious states. It is now well accepted that other people can know better than me my unconscious states. For example, you can detect that I am jealous, although I am unaware of this feeling. The core of the debate is thus the possibility that brainreading challenges the first-person authority for one's own conscious states. Here, I will focus on visual experiences.

2.3 Neuroimaging experiments

Brainreading is no longer only a thought experiment from the sixties. It has been shown that one can have access to someone else's conscious perceptual experiences by using brain imaging methods. Until very recently, these techniques were used to investigate and localize cognitive abilities, rather than mental states. The approach was similar to neuropsychology. They both aim at individuating functions of the mind, rather than states of the mind, whether it is by detecting the brain region responsible of a function or the lesion responsible of a functional deficit. The realization of the function in a specific mental state with a specific content was beyond the scope of brain imaging and neuropsychology. For instance, we have known for a long time that language understanding activates Broca's area, but we do not know yet what the understanding of the utterance "Tweety is happy" activates in contrast with the understanding of the utterance "Tweety is unhappy". In the same way, fMRI could show activation of V4 for color perception, but it did not reveal which color the subject actually perceived.

This is precisely the novelty of recent studies in neuroscience, which claim to be able to read the content of an individual's mental states by reading her brain. This new use of neuroimaging methods provides answers not only to the general question of functions like vision, but also to the very precise and detailed question of the content of perceptual experiences. The level of analysis has changed. Behaviorism considers the brain as a whole, like a black box, computationalism considers the brain as a set of cognitive functions, and this new trend of neuroscience considers the brain as a set of contentful states.

Let me provide a few examples (for a complete review, see Haynes and Rees, 2006). Some visual events are processed in very distinct areas of the brain. For example, the fusiform face area responds more strongly to faces, while the parahippocampal place area responds best to images containing views of houses. By measuring the activity in these two regions, one can track whether a person is currently seeing faces or houses with an accuracy of 85% (O'Craven & Kanwisher 2000). This method is limited as most of the percepts activate overlapping and distributed cortical networks, which are not as well segregated. However, one can still abstract the corresponding distributed pattern of brain activity after several presentation of the same visual event. Several types of perceptual content have already been decoded by such method: orientation, direction of motion and even perceived color. In addition, the content of visual states can be decoded although the subject is unaware of what he sees (Haynes and Rees, 2005). In this case, the brain scientist knows better than the subject her own mental states. However, this is limited to unconscious states.

But is it the perceptual experience *per se* that is read or merely a subpersonal stage of visual processing? A recent study has used binocular rivalry to show that one has access to what it is like for the subject to see the visual event. If you present two different images to the two eyes, one sees each image in turn for a few seconds while the other is suppressed. Binocular rivalry enables to dissociate between the visual processing, which remains the same, and the conscious perception, which alternates between the two images. Interestingly, recent studies using EEG or fMRI show that they could decode the fluctuation of perceptual experiences on a second-by-second basis (Haynes and Rees 2005). Brainreading provides access to conscious perceptual states.

There are some limits to such method. Until now, brainreading has been applied only to a very restricted scope of mental states, that is, perceptual states. Propositional attitudes like beliefs and desires are still far from being easily accessible by such method. Even for perception, brainreading works only for very simple visual stimuli. What would happen for a complex visual scene? In the hypothesis of the language of thought, one could combine and recombine concepts indefinitely; does the brain also follow the rules of compositionality? This might be just an empirical question. However, there are other difficulties that are more theoretical. I will briefly mention two of them here, but I will not enter in the debate. First, there is the question of externalism. If meaning is not in the brain, as claimed by Putnam (1975), then brainreading would never be able to give access to the content of perceptions and beliefs. Second, there is the question of holism. If the content of a specific belief can be understood only in relation with all the other mental states, then brainreading would have to read the whole brain and not only a few voxels. All these questions arise from the definition of mental representations and from theories of content. They deserve to be seriously addressed in order to know if brainreading is conceivable or not. I will not take side in the debate and focus rather on the epistemological question of the privacy of the mind. Does brainreading challenge first-person authority?

3. Who is right?

What is at stake here is the possibility that someone could know better than me my own conscious experiences. The recent results might help us to shed a new light on this debate.

What would happen if the brain scientist claims that you see vertical lines while you claim yourself that you see horizontal lines? We know that the brain scientist can know what you see although you are unaware of what you see. But this is a different situation. The subject has a conscious experience but the brain scientist disagrees on the content of this experience. With whom shall we take side? At first view, the situation is not as clear-cut as described by Armstrong and others. There are different parameters to take into account. For example, the answer partly depends on whether it is the first time that there is a disagreement or not. It seems to be useful to go further in the detail of the thought experiment and imagine different situations.

(a) Single-case scenario: the brain scientist disagrees only once with the subject. Otherwise, they perfectly agree.

(b) 50% chance scenario: one time the brain scientist and the subject agree, the next time they do not, and it goes on like that indefinitely.

(c) Repetitive error scenario: the brain scientist starts to always deny what the subject is claiming. They no longer agree.

Who is right in all these situations? The introspecting subject or the brain scientist? The single case scenario is the easiest one. The actual brainreading method does not pretend to possess 100% accuracy. One error on the side of the brain scientist remains acceptable. It becomes more complicated in the next two situations. We are confronted to three alternatives: the brain scientist is right, or the subject is right, or they are both right.

3.1 The brain scientist is right

What are the conditions for the brain scientist to be right? One plausible scenario described by Armstrong is that the subject displays a disruption of introspection. Like in perception, there can be introspective illusions and hallucinations. The underlying assumption is that introspection is a mechanism, which can be impaired like any mechanism. This "internal scanner" is not necessarily better than the scientist's brain scanner. Armstrong (1963) concludes that self-knowledge is not incorrigible:

> Consider again the case of a brain technician who has a perfect understanding of the correlation between states of my brain and inner experiences. Suppose, then, that I report, "I seem to be seeing something green," using the sentence as a phenomenological report on my visual experience. The brain technician is able to say from his knowledge of brain patterns that (i) I am not lying; (ii) my brain is in the appropriate state for some other experience; (iii) there are disturbances in the brain processes respon-

> sible for introspective awareness which would account for my mistake. On the evidence offered by the technician it ought to be concluded that I have made a mistake.

Armstrong imagines that the scanner shows some deficit of introspection that would explain why I report an experience that I do not have. Consequently, it makes sense to side with the brain scientist. This scenario may sound plausible for the repetitive error situation. We can imagine that a lesion has impaired my internal scanner leading to systematic introspective errors, while my perceptual system is preserved. The brain scientist could decode my visual experiences as before, while I am no longer able to do it. It is more difficult to imagine a plausible scenario in the 50% chance situation. An introspective mechanism, if such mechanism does really exist, rarely goes on and off like that. However, if the brain scientist does indeed detect disturbances of the introspective mechanism each time he disagrees with the subject, then he will be justified to conclude that he knows better than the subject.

However, the fact that introspection can be misleading does not logically imply that brainreading is more reliable than introspection. Brainreading is built in the first place on the subject's introspective reports. If these reports are fallible, how can they ground brainreading? Armstrong could reply that the brain technician knows when there is a "disturbance" of introspection, and thus, when he should not take into account the introspective reports. However, this assumes that the brain technician knows when the introspective reports are no longer reliable. But how could he know that? Once the correlation between the perceptual experience and the brain activity has been found, the technician is able to detect when the introspective report does not match the brain activity corresponding to a specific perceptual experience. However, to be able to correlate the perceptual experience and the brain activity in the first place, the technician has to rely on the introspective reports. But how can he know then if there is a mismatch between the report and the perceptual experience?

Let us imagine that the subject's introspective mechanism is already disturbed when the technician starts studying her brain. When the subject says 'I see the sky blue', she actually experiences the sky red. Because of her reports, the brain technician correlates the brain activity B_1 when she sees the sky red with the perceptual experience 'seeing the sky blue'. There is no way for the brain technician to know that her introspection is disturbed because he has nothing that he could check. Let us imagine now that the subject suddenly regains her normal introspective ability and reports seeing the sky red when she sees the sky red. The brain activity B_1 is the same as before. Therefore, the technician concludes that the subject is wrong and that she is actually seeing the sky blue. Yet, we should not side with the brain technician.

Alternatively, one might suggest that the brain technician does not need to know when introspection is impaired. As introspection is most of the time reliable, it should be enough to 'average' the reports in order to erase the mistakes. When the technician records the brain activity B_1, the subject reports 75% of the time that she sees the sky blue and only 25% of the time that she sees the sky red, then the correlation will be between B_1 and seeing the sky blue. But again, the technician has no guarantee that the subject is not right only 25% of the time and that she does not see the sky red.

Armstrong's scenario assumes that (i) there is an introspective mechanism that can break down and (ii) one can detect from the outside when this mechanism is impaired. The first assumption is already controversial but I will not go here into this debate, which goes far beyond the scope of this paper. The second assumption is more relevant for us. There is a vicious circle hidden behind Armstrong's scenario. To be able to know better than the subject her inner experiences, one has to be able to assess her introspective ability, and conversely, to be able to assess the subject's introspective ability, one has to be able to know her inner experiences.

3.2 Both the brain scientist and the subject are right

One way to escape the dilemma between the subject and the brain scientist is to argue that they are both right. Maybe the disagreement between the brain scientist and the subject lays only in the way they categorize the experience, as suggested by Rorty (1965). They detect the same experience, but they name it differently. According to Rorty, what is at stake here is the distinction between misjudging and misnaming. For instance, I could call 'pain' a state that brain scientists would not call 'pain'. We would not disagree on the state that I am in, but on the meaning of mental terms.

We can imagine at least two reasons for that. First, my conceptual repertoire is not rich enough to account for the fineness of grain of my perceptual experiences (Peacocke, 1992). I can report that I see the sky blue, while referring to many different shades of blue. My introspective report would not be precise enough. In this situation, both the brain scientist and the subject would be right, although the former would be more accurate than the latter. Second, I may use a proximate concept to refer to an experience that I am not familiar with. For example, I can report that the perfume smells like rose while it smells like lilac. The brain scientist will detect the brain activity for a lilac experience and he will be right.

These scenarios may sound plausible as long as one does not think of the source of the brain scientist's knowledge. How could he have a more fine-grained discrimination of my experiences than me? How could he decode the brain activity of perceptual experiences that I have never had before? If the brain scientist decodes the subject's brain activity based on the subject's introspective reports, then the brain scientist should have the same conceptual repertoire as the subject.

As an alternative, one could suggest that brainreading is based on the correlation between brain activity and external stimuli, rather than between brain activity and subjective reports. The advantage of this solution is that the brain scientist can discriminate between the different types of brain activity elicited by different shades of color. However, it does not solve the problem for unfamiliar stimuli. In addition, by giving up subjective reports, the brain scientist is no longer guaranteed that what he decodes is the phenomenal states. One could indeed imagine that in front of two shades of blue, I have the same experience.

One solution might be to correlate between the three dimensions: brain activity, external stimuli and subjective reports. But even this solution does not work in all the situations we described before. It is hard to imagine that one could keep changing one's mind about the way to categorize the same experience like in the 50% chance scenario.

To assess that the brain scientist is right, one has to be ready to assume that (i) introspective reports are not always reliable for conscious experiences, and (ii) brain scientists have a source of knowledge independent of introspective reports that is more reliable and more fine-grained. One solution might be that the scientist relies not only on the subject's introspective reports, but also on other individuals' reports. However, as we will see, this solution raises several difficulties.

3.3 The subject is right

If we are not entitled to assess that the brain scientist knows better than the subject, then we shall conclude that the subject is right. Interestingly, brain scientists themselves like Haynes would take the subject's side in case of conflict (personal communication). There are two possible reasons for the scientist's mistake, either because of the nature of introspection or because of the brainreading method. The first interpretation is defended by Smart (1962). He claims that the possibility of the brain scientist being right against the subject is not even conceivable. A priori if I claim that I have an experience, then I do have this experience. Mental states are private and therefore, cannot be accessible publicly. I would not go here into the detail of this view, as it seems more interesting to focus on our actual results and on their underlying method. As Armstrong (1963) notices, one should accept what the scientist says only if the brain theory is well founded. The first step is thus to show that the brain theory is well founded. But does it really deserve to be called a theory? More particularly, can we say that through brainreading, one has reached a reductive explanation of the mind? By reading the mind, do we have reduced it?

According to Ernst Nagel (1961), reductive explanation has two main features:

- Connectability: mental terms are definable in neural terms by using bridge-laws.
- Derivability: mental laws, once they have been translated in neural terms, are explainable by the brain theory.

Following such definition, we can quickly see that we are still far from a reductive explanation of the mind. First, brainreading does not give an explanation of higher-level phenomena (i.e. mental states) on the basis of an account of the lower-level phenomena (i.e. brain states). It does not appeal to a detailed and exhaustive brain theory to which one could reduce mental states. Second, one is still far from translating mental laws to brain laws. All that one can do is to translate some mental states to some brain states. Do these translations qualify as bridge laws?

One of the issues raised by Nagel's notion of bridge laws is the items that they are supposed to connect. Do they apply to tokens or to types of mental states? How specific are they? In the brainreading context, this question is all the more important that the actual method underlying brainreading cannot be generalized to types. Let me now describe this method.

Brainreading is based on repetitive correlations between a specific brain state B_1 and a specific perceptual state P_1 occurring in a specific individual subject S_1. In this sense, it is well founded. However, one cannot generalize from these correlations. Brainreading is based on a one-to-one matching, which raises two difficulties. First, it is limited to P_1 and it cannot be generalized to other perceptual states $P_2, P_3 \dots P_n$. Second, it is limited to S_1 and it cannot be generalized to the same perceptual state P_1 in another individual S_2. This is due to differences between subjects. The underlying brain theory is thus a collection of correlations between specific states for each individual. Standard brain imaging techniques require normalizing the data that they have from different subjects. Put it another way, no brain looks exactly like the other. In order to average the brain activity, one has first to map the subject's specific brain to a standard brain that will be used as a reference. By doing so, one loses fine-grained spatial differences, which are essential to brainreading. There is a dilemma: either one averages the brains so that the results are true of several individual people, but one cannot decode mental states, or one analyses each individual brain so that one can decode mental states, but the results cannot be generalized to other individuals.

Are these limits purely empirical and temporary, specific to the actual method underlying brainreading? Or are they more fundamental? We can of course expect some progress. However, the possibility of reducing the mind is beyond the reach of any scientific progress. What is at stake here is the classical problem of multiple realizability (Fodor, 1965; Putnam, 1965). The classical version of the problem highlights the variety of ways in which the same computational states can be implemented or realized in different systems (electronic, mechanical, hydraulic…). But even within the same kind of system (e.g. a human brain), one cannot identify any mental state with any specific brain state, since the same mental state can be realized without this brain state. Multiple realizability applies within and between individual people.

To reply to the problem of multiple realizability, one could suggest providing a disjunctivist account of bridge-laws to establish nomic equivalence between P_1 and B_1 or B_2 or B_3. However, Kim (1993) argues that such account cannot work because heterogeneous disjunctions are not projectible. He uses the example of jade. One is not entitled to conclude that jade is green just because one has found only green samples of jade. Indeed, one has maybe brought only jadeite and one cannot expect that the next sample of jade will also be green as it could be nephrite. Jade is a heterogeneous disjunction of two kinds, nephrite and jadeite. Similarly, finding that P_1 and B_1 are nomically equivalent does not provide any justification for believing that the next P_1 is a B_1. The fact that P_1 accompanies B_1 is a further fact, not explainable by simply telling the story about the correlations. In case of disagreement, we cannot side with the scientist.

Brainreading is not only a hot topic for sci-fi novels nor an exciting thought experiments from the sixties. It is now becoming the reality, convening with it a stream of dreams, but also of ethical worries. Mental images were thought for a long time to be private, available only to the subject. Privacy was even part of their definition. And it seems now to be no longer true. But does that mean that one could know better than us the content of our perception, dream, hallucination? There are two ways to refute this conclusion. First, one may reply that we cannot be wrong about our own conscious perceptual experiences. However, even if we could be wrong, one can still argue that brainreading is not reliable enough to challenge first-person authority. It would need to rely on a reductive explanation of the mind. However, such reduction of the mind has been shown to be impossible given the possibility of multiple realizability. Brainreading relies on a mere collection of brute correlations, which cannot provide any guarantee that they will go on forever.

The problem of multiple realizability however is not an argument against brainreading per se, but against the possibility of knowing better than the subject. The following example will illustrate this point. The same book, say, *A la recherche du temps perdu*, from Marcel Proust, can be "realized" in a variety of texts using different languages. An English reader although he is unable to speak french, is still capable to read *A la recherche du temps perdu*. His limits do not prevent him to decode this realization of the book. Similarly, the brain scientist can decode the subject's mental states. The difficulty arises when there is a disagreement between the subject and the brain scientist. Brainreading is not grounded enough to challenge introspective reports. Brain scientists can read your mind, but they cannot know better than you.

References

Armstrong, D. (1963). Is Introspective Knowledge Incorrigible? *Philosophical Review*, LXII, 417-432.

Ayer, A. (1963). *The concept of a person and other essays.* St Martin's Press, New York.

Baier, K. (1962). Smart on sensations. *Australasian Journal of Philosophy*, 40, 57-68.

Fodor, J. (1965). Explanation in Psychology. In M. Black (ed.), *Philosophy in America*, Muirhead Library of Philosophy.

Haynes, John-Dylan and Rees, G. (2005). Predicting the stream of consciousness from activity in human visual cortex. *Current Biology*, 15, 1301-7.

Haynes, John-Dylan and Rees, G. (2006). Decoding mental states from brain ac-

tivity in humans. *Nature Reviews Neuroscience*, 7, 523-34.
Jackson, F. (1973). Is there a good argument against the incorrigibility thesis? *Australasian Journal of philosophy*, 51, 51-62.
Kim, J. (1993). *Supervenience and Mind.* Cambridge University Press, Cambridge (Mass.).
Nagel, E. (1961). *The structure of science.* Harcourt, Brace and World, New York.
O'Craven, K. and Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stiimulus-specific brain regions. *Journal of Cognitive Neuroscience*, 12, 1013-23.
Peacocke, C. (1992). *A study of concepts*. MIT Press, Cambridge (Mass.).
Putnam, H. (1965). Brains and Behavior. In R. J. Butler (ed.), *Analytical Philosophy* Vol 2, Blackwell, Oxford.
Putnam, H. (1975). *Mind, Language and Reality. Philosophical Papers, vol. 2*. Cambridge University Press, Cambridge (Mass.).
Rorty, R. (1965). Mind-Body Identity, Privacy, and Categories. *The Review of Metaphysics*, 19, 24-54.
Shoemaker, S. (1996). *The first-person perspective and other essays*. Cambridge University Press, New York.
Smart, J. (1962). Brain processes and incorrigibility. *Australasian Journal of Philosophy*, 40, 68-70.

Frédérique de Vignemont
Institute of Cognitive Science, CNRS
fvignemont@isc.cnrs.fr

Finito di stampare nel mese di Dicembre 2006
presso la Stampa Editoriale srl
STRADA STATALE 7/BIS - ZONA INDUSTRIALE DI AVELLINO
83030 – MANOCALZATI (AV)
Tel. 0825/62.69.66 - Fax 0825/61.08.88