tence. The existence of the strawberry has no bearing on my access to my own mental states. If visual phenomenology reveals nothing about the ontology representation, there is no reason to think that imagery does.

NOTE

1. The modern classic is Harman (1990).

Loss of visual imagery: Neuropsychological evidence in search for a theory

Georg Goldenberg

Neuropsychological Department, Munich Bogenhausen Hospital, D 81925 Munich, Germany. **Georg.Goldenberg@extern.lrz-muenchen.de**

Abstract: Observations on patients who lost visual imagery after brain damage call into question the notion that the knowledge subserving visual imagery is "tacit." Dissociations between deficient imagery and preserved recognition of objects suggest that imagery is exclusively based on explicit knowledge, whereas retrieval of "tacit" visual knowledge is bound to the presence of the object and the task of recognizing it.

Pylyshyn concludes that neuropsychological evidence does not support the contention that mental images are based on retinotopically organized neural representations. This argument is convincing but does not exhaust the contribution of neuropsychology to the theory of mental images. Observations of patients who lost visual imagery after brain damage call for a refinement or revision of the "tacit knowledge" hypothesis, too.

There are at least five visual categories for which imagery can be selectively affected by brain damage: shapes and colors of common objects, shapes of faces, shapes of letters, and topographical relationships (review in Goldenberg 1993). A patient who is unable to answer imagery questions about the shape of the ears or the length of the tail of animals (Kosslyn 1983) may do perfectly well on imagery questions like those shown in Figure 4 of the target article concerning the shape of letters (Goldenberg 1992). Such dissociations can hardly be explained by damage to a visual buffer or any other structure subserving generation of visual images independent of their content. The more likely hypothesis that these patients have lost knowledge of the visual appearance of only one category of things calls for a theory of that knowledge. How is it organized that it can break down for only one category? How is it related to knowledge of non-visual properties?

Another challenge to the tacit knowledge hypothesis is constituted by patients with loss of visual imagery and preserved visual recognition (Basso et al. 1980; Farah et al. 1988; Goldenberg 1992). The proposal that these patients have preserved knowledge of the visual appearance of objects, but are unable to employ an "image generation process" transforming knowledge into mental pictures (Farah 1984), has been criticized on two grounds. First, as already mentioned, the imagery deficit can be restricted to only certain categories of things. Second, it has been shown that these patients make errors when they are shown pictorial versions of imagery questions, although in this condition the crucial images are before them and need not be generated before the "mind's eye." For example, when shown images of bears with rounded and with pointed ears they cannot decide which of them is correct. Obviously, they lack knowledge of the shape of the bear's ears. Nonetheless they readily recognize that these are bears. Their visual recognition must have access to knowledge of the global shape and the characteristic features of bears to distinguish them from lions or dogs. But they are completely unable to imagine the visual appearance of a bear and not just that of its ears! The knowledge they use in recognition cannot be used for imagery.

Based on these lines of evidence I proposed that there are two kinds of knowledge of the visual appearance of things (Goldenberg 1992; 1998; Goldenberg & Artner 1991). Knowledge used in recognition is restricted to those features which permit a reliable identification of an object under varying circumstances. It neglects details like the shape of the bear's ears. There is a second store of visual knowledge within semantic memory. This knowledge includes information on features not necessary for recognition in addition to those used for recognition. The source of semantic visual knowledge may be an active interest in the visual appearance of objects, possibly enhanced by the high value given to visual arts in our culture and education (Armstrong 1996; Farah 1995a). The crucial point of this hypothesis is that knowledge used for visual recognition is completely embedded in visual recognition and cannot be used for any other purpose. Visual imagery is based exclusively on visual knowledge within semantic memory. If only this knowledge is lost, patients are unable to imagine the visual appearance of objects although the knowledge embedded in recognition enables them to recognize the same objects.

This hypothesis calls into question the idea that visual imagery is based on "tacit" knowledge. Pylyshyn states that "knowledge is called 'tacit' because it is not always explicitly available for . . . answering questions" (target article, sect. 3.1). Presumably "not always" means that retrieval is bound to a certain context or task. This applies to the knowledge used for recognition: Its retrieval is bound to the presence of the object and to the task of recognizing it. By contrast, you can form mental visual images in the absence of the object and in response to a wide variety of questions (or just for fun), that is, in principle, always!

I propose that visual imagery is equivalent to the explicit recall of semantic knowledge of the visual appearance of things. This position is not meant to be a theory of imagery, but a request for such a theory. An adequate theory of imagery should explain how such knowledge is acquired, how it is organized, and how it differs from knowledge of other properties of things. It seems to me that imagery is still in search of a theory.

You are about to see pictorial representations!

Frédéric Gosselin[a] and Philippe G. Schyns[b]

[a]*Département de Psychologie, Université de Montréal, Montréal QC, H3C 3J7, Canada;* [b]*Department of Psychology, University of Glasgow, Glasgow G12 8QB, United Kingdom.* **frederic.gosselin@umontreal.ca**
philippe@psy.gla.ac.uk
http://mapageweb.umontreal.ca/gosselif/cv.html

Abstract: Pylyshyn argues against representations with pictorial properties that would be superimposed on a scene. We present evidence against this view, and a new method to depict pictorial properties. We propose a continuum between the top-down generation of internal signals (imagery) and the bottom-up signals from the outside world. Along the continuum, superstitious perceptions provide a method to tackle representational issues.

In a memorable courtship scene from the movie "A beautiful mind," John Nash asks his future wife to think of an object. "Anything!" he says. She chooses an umbrella. He then turns toward the starry sky and, connecting some stars one by one with his finger, shows her a sparse, but nonetheless recognizable, umbrella. You might not be capable of performing this feat on demand, but you have surely seen sparse versions of objects or scenes in the sky or elsewhere at one time or another. On a continuum extending from pure top-down mental images (internal signals) to strong bottom-up signals, these extremely sparse objects (we call them *superstitious perceptions* in reference to Skinner's celebrated 1948 article) are closer to mental images than extraneous signals. More importantly, we will demonstrate that they provide a powerful analytic tool to address the issue of internal representations.

We have recently produced a situation similar to the "umbrella in the stars" in our laboratories (Gosselin & Schyns, in press). In one experiment, we instructed two observers (MJ and NL) to detect the presence of a letter "S" (for Superstitious) inserted in white noise (black and white pixels peppered across the image field). The observers were instructed that the letter "S" was black on a white background, filled the image, and was present on 50% of 20,000 trials. No more detail was given regarding the attributes of the letter. Unbeknownst to the observers, each trial only consisted in the presentation of a 50 x 50 pixels white noise image (see our Fig. 1a, for one example) with a black-pixel density of 50%. *Crucially, no bottom-up external signal (i.e., an "S") was ever presented.*

At first, the observers found the task rather difficult, but, soon, they said, they responded with ease. In fact, observer NL said that after about 1,000 trials the "S" popped out when it was present. In any case, the observers detected an "S" in noise on 46% (NL) and 11% (MJ) of the trials, respectively.

What did the observers respond to? As already stated, no external signal was ever presented, and the observer only saw white noise. One possibility is that observers generated an internal signal via imagery, and tried to superimpose this signal onto the incoming white noise. Sometimes, this internal signal will be weakly correlated (here, a correlation smaller than .026) with the external white noise and the observer will detect the letter corresponding to his or her imagined signal. On the remaining trials, the mismatch will simply be too large and the observer will reject the noise as being what it is – noise. However, and this is important to stress, the observer must first generate an internal signal via imagery to be able to perform this detection task, and attempt to superimpose this internal signal to external noise. What is the internal signal of the observer? We will contend that whatever it is, it represents pictorial properties of the imagined letter.

From Wiener (1958), we know that systematic responses of a black box to white noise can be used to analyze its behavior. We are thus looking for a systematic correlation between the noise fields (**xi**) and the detection responses (**y**). This is what *reverse correlation* does (see also Ahumada & Lovell 1971). The first Wiener kernel (the linear component) is equal to $k^{-1}\sum_t y(t)x_t$, where k is a constant and t is variables indexing all the trials. Leaving aside k, this amounts to subtracting the sum of all the noise fields that led to a rejection response from the sum of all the other noise fields (see Fig. 1a, NL and MJ). For each observer, we best-fitted a Gaussian density function (see Fig. 1b, the solid lines) to the energy distribution of his or her first order Wiener kernel (Fig. 1b, the open circles). This kernel (called the "classification image") represents the template of information that drives the detection of the target "S" letter for this observer. In other words, the first order kernel provides a first approximation of the representation of the imagined internal signal for the letter "S." To better visualize this representation, we sought an information peak in a spectral analysis of the kernel, and filtered out all spatial frequencies one standard deviation away from the mean (i.e., keeping a bandwidth of 0–3 cycles per letter). The outcomes are black "S"s on a white background filling the image (see Fig. 1c, NL and MJ).

The first order kernel predicts the detection response from each pixel, individually. However, it is likely that observers used higher order relationships between the elements of the internal signal – for example, combinations of two pixels. The second Wiener kernel examines what these second order relationships are. It is equal to $k^{-2}\sum_t y(t)x_t'x_t$. Leaving aside k, this is equivalent to subtracting the sum of all the autocorrelations of the other noise fields (i.e., the outer product of each noise field vector with itself) that led to a rejection response from the sum of all the autocorrelations of the other noise fields. Figure 1f (NL and MJ) shows the regions of the second-order kernels that are statistically significant ($p < .01$). The number of significant regions far exceeds what would be expected by chance for both observers (937 pixels for NL, $p < .01$, and 1,318 pixels for MJ, $p < .01$), revealing that the imagined internal signal imagined *did* include nonlinear relationships.

What conclusions can be drawn from this study? We have induced superstitious perceptions of an "S" by instructing observers to detect this letter in noise. Unknown to them, the stimuli never comprised the letter, but only white noise. If the observers had been performing only according to an external signal (i.e., in a bottom-up manner), their kernels should have had the same properties as averaged white noise – that is, zero energy across all spatial frequencies. However, there was a marked peak of energy between 1 and 3 cycles per letter that could only arise from top-down influences arising from an internally generated signal – that is, a mental image. Further analyses revealed the properties of the internal signal driving the detection behavior. With white noise as inputs, the revealed letter could only depict the observer's imagined letter "S."

Is the internal signal pictorial in nature? Functionally, yes, because, if not from a matched internal signal, where else would the pictorial properties present in the kernels come from? Does this imply that the observers actually used an image of a "S" from their memory? Not necessarily, but they had to have knowledge of *all* the pictorial characteristics of an "S," functionally isomorphic to an actual image of an "S." We believe that you have just seen representations with pictorial properties!

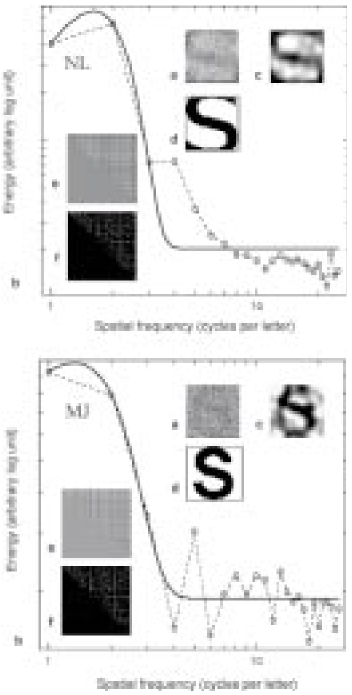

Figure 1 (Gosselin & Schyns). Adapted from Gosselin and Schyns (in press, Experiment 1). (a) Raw first order Wiener kernels. (b) Distributions of the average squared amplitude energy for different spatial frequencies (collapsed across all orientations) of (a) (expected energy = constant). The solid lines are the best Gaussian fits. (c) (a) filtered with a smooth Butterfield low-pass. We squeezed pixel intensities within two standard deviations from the mean. (d) Best matches between (c) and 11,284 letters. (e) Raw second order Wiener kernels. (f) Statistically significant ($p <$ .01) pixels of (e).