VIEWPOINT
COVERT RECOGNITION AND THE NEURAL SYSTEM FOR FACE PROCESSING

Stefan R. Schweinberger and A. Mike Burton
(Department of Psychology, University of Glasgow)

ABSTRACT

In this viewpoint, we discuss the new evidence on covert face recognition in prosopagnosia presented by Bobes et al. (2003, this issue) and by Sperber and Spinnler (2003, this issue). Contrary to earlier hypotheses, both papers agree that covert and overt face recognition are based on the same mechanism. In line with this suggestion, an analysis of reported cases with prosopagnosia indicates that a degree of successful encoding of facial representations is a prerequisite for covert recognition to occur. While we agree with this general conclusion as far as Bobes et al.’s and Sperber and Spinnler’s data are concerned, we also discuss evidence for a dissociation between different measures of covert recognition. Specifically, studies in patients with Capgras delusion and patients with prosopagnosia suggest that skin conductance and behavioural indexes of covert face recognition are mediated by partially different mechanisms. We also discuss implications of the new data for models of normal face recognition that have been successful in simulating covert recognition phenomena (e.g., Young and Burton, 1999, and O’Reilly et al., 1999). Finally, in reviewing recent neurophysiological and brain imaging evidence concerning the neural system for face processing, we argue that the relationship between ERP components (specifically, N170, N250r, and N400) and different cognitive processes in face recognition is beginning to emerge.

Key words: faces, covert recognition, prosopagnosia, ERPs, modelling

INTRODUCTION

The two papers by Bobes and colleagues and by Sperber and Spinnler (2003, this issue) contribute intriguing new findings to the literature on covert face recognition. The initial observation, in patients with Capgras delusion and patients with prosopagnosia, that patients who are completely unable to recognize faces when tested overtly nevertheless can show some preserved recognition processes when tested indirectly has attracted much interest, as had similar findings of preserved implicit processing in other neurological disorders such as memory, language, or attention (Schacter et al., 1988).

Ever since the striking phenomenon of covert face recognition was first published, a key question has been whether overt and covert face recognition are produced by the same functional system (Burton et al., 1991). Alternatively, covert recognition might reflect the function of either a distinct secondary face recognition system (Bauer, 1984), or a disconnection of an intact system of face recognition from the processes that signal recognition to an awareness system (DeHaan et al., 1992). Both current papers make a contribution to this continuing controversy.

Bobes and coworkers (2003, this issue) investigated ERP correlates in identity matching for unfamiliar faces. The task was to decide whether two sequentially presented different pictures showed the same or a different person. Although FE (a 69-year old patient with prosopagnosia resulting from traumatic brain injury) performed at chance levels, his ERP response gave evidence of covert matching of unfamiliar faces, in terms of differences in the ERP for mismatching (non-repeated) as compared to matching (repeated) second faces. Importantly, these differences in FE had a very short onset latency (in fact, towards the lower limit of what was seen in 10 healthy controls matched for age and education). To Bobes et al. this suggests that FE is unimpaired in early face perception, and can extract structural codes from faces with normal speed. This finding has potentially important implications. The authors argue that the very same impairment that causes prosopagnosia may explain an impairment in sequential (but not simultaneous) matching of unfamiliar faces. While this may well be the case, we wish to make two small comments here. First, as Bobes et al. used no time delay between the sequentially presented faces, it is more difficult to exclude a perceptual matching strategy, especially as no backward masking was used to eliminate pictorial codes from the first face. A second and controversial issue (that the authors acknowledge) is whether unfamiliar and familiar faces are processed by the same mechanisms. Currently, considerable evidence favours the idea of different mechanisms in the processing of identity from familiar and unfamiliar faces (Hancock et al., 2000; Herzmann et al., 2002). In this respect, it would have been interesting to see whether ERP findings would support a normal timing of familiar face perception in FE as well. Similarly, other aspects of face processing appear to be relatively independent of identity, for example computations of sex or expression (Le Gal and Bruce, 2002), but see also Schweinberger and Soukup (1998). The question of which structural codes are being extracted here, and what information is necessary in order to make different decisions to faces, is one to which we will return.

Bobes and coworkers emphasized that in most of their participants, the difference between matching and mismatching second faces yielded two different ERP effects - an early modulation at a latency of approx. 180-300 ms, and a later modulation at a latency around 400-650 ms. Due to substantial individual variability in the topography and latency of these ERP modulations, it is not easy to ascertain their functional significance, or to relate them to known ERP components. Below we will comment in more detail on the potential significance of these interesting findings, and tentatively relate them to ERP modulations recently identified in immediate face repetition priming.

Sperber and Spinnler (2003, this issue) used the face-name relearning paradigm (Bruyer et al., 1983) to investigate covert person recognition in a patient with progressive fronto-temporal dementia. They initially observed abolished overt recognition but preserved covert recognition. However, as the disease progressed, covert recognition also faded away. These findings were taken as support for the idea that covert and overt face recognition are based on the same mechanisms, and that a partial and incomplete activation of semantic information for people may still support covert but not overt face recognition.
These data provide informative and novel contributions to our understanding of covert recognition. In this article, we would like to draw attention to four current questions that relate to the interpretation of these new data in the context of the functional and neuronal architecture of the normal human face recognition system. First, do all forms of covert face recognition reflect the operation of the same mechanisms, or are there dissociable forms of covert recognition? This question arises because covert face recognition has been demonstrated using a variety of methods, with variable results. It is also well known that prosopagnosia is not a unitary disorder in terms of the functional level of the deficit, as well as in functional neuroanatomy and etiology of the lesion. Second, we therefore ask what different forms of prosopagnosia should be distinguished, and how do these relate to findings on covert recognition? Third, how can covert recognition be explained with respect to functional models of normal face recognition? And finally, can we begin to relate functional processes to brain processes which can be observed with current functional brain imaging methods, such as fMRI or ERP?

1. DIFFERENT FORMS OF COVERT RECOGNITION?

Covert face recognition in the absence of overt recognition has been demonstrated with a variety of methods. Most frequently used are behavioural methods such as face-name relearning (Bruyer et al., 1983), face-name interference (DeHaan et al., 1987), or RT priming (Young et al., 1988). Other approaches use classical conditioning (Cole and Perez-Cruet, 1964), skin conductance responses (SCRs; Bauer, 1984), event-related brain potentials (ERPs; Renault et al., 1989), and eye-movement scanpaths (Rizzo et al., 1987).

Here we will focus on the three most frequently used methods: One approach is face-name relearning, where the patient is taught an association between a familiar face and a familiar name. Half of the items are “true” pairs and the other half are “untrue”, non-corresponding pairs (see also Sperber and Spinnler, this issue), and covert recognition is indicated by an advantage in learning true over untrue pairs. In RT priming tasks, speeded familiarity decisions are performed for familiar and unfamiliar names, which typically presents no problem for patients with prosopagnosia. Immediately preceding the target name, a prime is shown which can represent a related or an unrelated person. In healthy participants RTs to target names are faster for primed than unprimed target names, and this is independent of whether the prime is a name or a face (Young et al., 1988; Schweinberger, 1996). In prosopagnosic patients, prime faces may cause significant priming, even when these faces cannot be recognized overtly (Young et al., 1988). Finally SCRs can be measured while a series of faces showing familiar and unfamiliar people is presented. Typically, faces of familiar people elicit a larger SCR than unfamiliar people, even when overt recognition is at chance levels. This has been shown both for patients with prosopagnosia (Tranel and Damasio, 1985, 1988), and for healthy participants using subliminally presented faces (Tranel and Damasio, 1988; Ellis et al., 1993). In a variant of the SCR paradigm, larger SCRs were demonstrated in LF
when a familiar face was paired with the correct as opposed to an incorrect name. Again, the patient performed at chance levels when overtly choosing the correct name (Bauer, 1984).

A controversial question is whether covert recognition reflects the function of a distinct secondary face recognition system, perhaps mediated via a dorsal visual-limbic route (Bauer, 1984, 1986), or a disconnection of an intact system of face recognition from the processes that signal recognition to an awareness system - a view that had been advanced by DeHaan, Bauer, and Greve some years later (DeHaan et al., 1992). Alternatively, covert measures may be simply more sensitive to the output of the normal face recognition system, even when this output is reduced to the extent that overt recognition is no longer supported. In other words, the single dissociation between preserved covert and impaired overt recognition might reflect a resource artifact (Shallice, 1988). Supporting this idea, computational accounts have successfully modelled covert recognition in the absence of overt recognition, either in terms of partial damage to the links between visual face representations and modality-independent person representations (Burton et al., 1991), see also Figure 1 and Burton et al. (1990, 1999), or in terms of partial damage to visual face representations themselves (Farah et al., 1993). As we will see below, these two accounts differ strongly in several respects (Young and Burton, 1999), but they share the idea that both overt and covert recognition emerge from the same functional system.

Note that these models, favouring similar mechanisms for overt and covert recognition, focus exclusively on simulation of covert recognition as obtained in behavioural tasks (e.g., face-name relearning, face-name interference, or priming). However, given the variety of behavioural and physiological methods that have been used so far, the question arises whether all these measures of covert face recognition reflect the same phenomenon. In this article, we make the case that SCR and behavioural measures of covert recognition, as well as overt face recognition, are mediated by the function of common components (face recognition units in ventral occipitotemporal cortex). Further downstream in processing, however, SCR and behavioural covert recognition reflect different phenomena (Figure 1).

We argue that the published findings on covert recognition as assessed by behavioural measures (face-name relearning, interference, and RT priming) can indeed be explained by damage to a single system: the system which supports overt face recognition. One argument for this proposal is that, across patients, there appears to be a systematic relation between the severity of perceptual disorders and the presence or absence of behavioural covert recognition. Specifically, whereas covert recognition can be observed in virtually all patients with prosopagnosia who show preserved face perception (i.e., the ability to perform tasks other than identification of faces), covert recognition has been absent in many (though not all) patients in whom face perception was found to be severely impaired (Newcombe et al., 1989; Sergent and Villemure, 1989; Etcoff et al., 1991; Schweinberger, 1992; Young and Ellis, 1989). More important, cases of preserved overt and impaired behavioural covert recognition do not seem to have been reported (see also Sperber and Spinnler, 2003, this issue). Although covert recognition may reflect a complete disconnection of an
intact system of face recognition from the processes that signal recognition to an awareness system (DeHaan et al., 1992), this idea is weakened by striking findings that overt recognition of familiar faces can sometimes be provoked in patients with prosopagnosia, particularly when multiple exemplars of faces are shown and the patient is told that all belong to the same category (Sergent and Poncet, 1990; DeHaan et al., 1991; Morrison et al., 2001). Overall, we suggest that behavioural covert recognition is mediated via the same mechanisms as overt face recognition, and our model of this mechanism is specified in detail below.

In contrast to behavioural covert recognition, SCRs may be caused by a
different mechanism, and have been related to an affective response to familiar faces (Breen et al., 2000; Ellis and Lewis, 2001). However, it should be noted that such a response seems to be neither a core component of overt face recognition nor of behaviourally assessed covert recognition. At a neuroanatomical level, there is evidence for a key role of the amygdala in mediating SCR responses to faces (Williams et al., 2001), but the amygdala is unlikely to be involved in the “core system” for face identity analysis (Haxby et al., 2000).

Sperber and Spinnler (2003, this issue) claim that cases of spared overt and impaired covert face recognition have yet to be reported, but here they missed at least two lines of evidence for double dissociations between overt recognition and covert SCR recognition: First, Tranel and coworkers (1995) described four patients with bilateral ventromedial frontal damage, who were unimpaired in overt recognition of familiar faces but failed to generate the usual larger SCR for familiar faces. This pattern was in strong contrast to findings in three patients with prosopagnosia resulting from occipitotemporal lesions, who all showed covert SCR recognition in terms of larger SCRs for familiar faces (Tranel et al., 1995). Second, Ellis and coworkers argued that patients with Capgras delusion exhibit a pattern of impairment that is a mirror image of what is seen in prosopagnosia. Capgras delusion often lacks a clear aetiology, and may occur in the context of psychiatric conditions, or secondary to structural or toxic brain damage. Capgras patients remain able overtly to identify a familiar person’s face. However, they believe that highly familiar people were replaced by impostors, doubles or aliens, and they often hold this belief with extreme conviction (Ellis, 1997). This could indicate that these patients, though unimpaired in cognitive aspects, may lack an appropriate arousal response that usually accompanies a familiar face’s recognition. Ellis and Young (1990) proposed that Capgras syndrome was a result of damage to an affective route to face recognition, perhaps along the dorsal visual-limbic pathway proposed by Bauer (1984). This generated the prediction that Capgras patients would fail to produce an SCR response to familiar faces, a prediction which was subsequently confirmed (Ellis et al., 1997; Hirstein and Ramachandran, 1997). Interestingly, although patients with Capgras delusion lack an autonomic response to familiar faces, they may show preserved priming effects in RTs, which presumably reflect a preserved cognitive route to recognition (Ellis et al., 2000). In sum, covert recognition as assessed by SCR responses should therefore be distinguished from behaviourally expressed covert recognition.1

1 We acknowledge that this argument would be strengthened if data were available in which covert recognition was assessed using a variety of methods. Unfortunately, physiological (SCR) and behavioural measurements were rarely undertaken in the same patient. Two exceptions are a study by Etcoff et al. (1991) who investigated a patient with prosopagnosia in the context of impaired perception, in whom covert recognition could not be demonstrated with either method, and a study by DeHaan et al. (1992) who demonstrated preserved covert recognition both in SCRs and behavioural priming in LF, a patient with prosopagnosia but relatively preserved perception. While these two case studies provide no compelling reason to assume different mechanisms for SCR vs. behaviourally covert recognition, one needs to bear in mind that the association of deficits does not allow one to draw strong conclusions.
2. HOW MANY DIFFERENT FORMS OF PROSOPAGNOSIA?

Faces are powerful social stimuli which provide information for person identification, but also for inferring a person’s sex, age, mood, and (via facial movements) for the comprehension of spoken language. While there are many different aspects of face perception that can be disrupted by neurological conditions (Campbell et al., 1986; Calder et al., 1996), prosopagnosia only refers to an impairment to perceive facial identity. Following Bodamer (1947), prosopagnosia is defined as a specific inability to recognize faces, with largely preserved object recognition. It should not reflect a more general problem with memory for people, nor should it be due to compromised basic sensory processing. Henri Hécaen, (1981) was one of the first who strongly argued for several functional sub-types of prosopagnosia. He distinguished three variants, (1) a deficit in perception which as a result of the loss of face-specific detectors may be limited to faces but is otherwise analogous to apperceptive agnosia, (2) a deficit in the connection between perceptual and mnestic processing, i.e., a face-specific form of associative agnosia, and (3) a metamorphopsic disturbance in which perception of faces appears grossly distorted. The latter was also described in Bodamer’s (1947) original paper, but is of relatively minor importance in contemporary discussions.

The reader may note that some additional variants of prosopagnosia have been reported in the literature. However, we believe that in view of the above definition, the term prosopagnosia is inappropriate for two of these variants: “Amnestic associative prosopagnosia” (Damasio et al., 1990) was described as a face agnosia which is accompanied by an inability to recognize people by other signals such as the voice or the name. This is in fact a modality-independent memory problem for people (see also the patient reported by Hanley et al. (1989), as well as patient 13 reported by Neuner and Schweinberger (2000)). Carney and Temple (1993) reported a patient who had a deficit in naming faces, a condition they called “prosopanomia”. However, it was not assessed whether or not the patient could name people when provided with other sources of information (e.g., voice, semantic description), and it is possible that this patient may have been suffering from anomia for personal names (Hittmair-Delazer et al., 1994; Semenza et al., 1995).

The distinction between apperceptive and associative prosopagnosia continues to be frequently used (DeRenzi et al., 1991), in broad analogy to Lissauer’s (1890) classification of agnosia. The integrity of face perception or structural encoding has been most typically assessed with the simultaneous matching of one or more exemplars of a face across different viewpoints of lighting conditions (Benton and Van Allen, 1968). This test is difficult perceptually but, due to the simultaneous presentation of the faces, involvement of memory is thought to be minimal. While one group of patients (those said to have apperceptive prosopagnosia) fail this test, a second group does remarkably well, and is said to suffer from associative prosopagnosia, in which the inability

---

2 One has to consider that face recognition requires the identification of a specific exemplar within a homogeneous category, whereas object recognition typically only requires basic level categorization of an object. While many patients with prosopagnosia fail on tests that require the recognition of individual exemplars of objects such as cars, glasses etc., some patients appear to exhibit face-selective deficits even when this factor is considered (Farah et al., 1994).
to recognize familiar faces cannot be attributed to a deficit in encoding an appropriate facial representation. This classification has its merits, but it largely ignores the possibility that different aspects of the physical face stimulus may code for different types of information (such as identity, sex, age, expression, etc.). It may therefore be possible to define sub-types of apperceptive prosopagnosia, an idea that has found rather little attention so far.

Nevertheless, there clearly is some relation between covert recognition and perceptual impairment. Specifically, covert recognition could be demonstrated in virtually all patients with associative prosopagnosia but only in some patients with apperceptive prosopagnosia (Sergent and Villemure, 1989; Etcoff et al., 1991; Schweinberger, 1992; Barton et al., 2001). This suggests that covert recognition requires at least some residual capacity to encode facial representations.

Regarding the neuroanatomical substrate of prosopagnosia, bilateral occipitotemporal lesions are found in a majority of patients with prosopagnosia, but there is a significant number of cases with well-documented unilateral right lesions. It has been claimed (Damasio et al., 1990) that unilateral RH lesions need to be more extensive to cause prosopagnosia, and that the RH variant of prosopagnosia corresponds to an apperceptive deficit. While some correlation may exist, the fact that some patients with unilateral RH lesions exhibit no (or only minor) perceptual impairment (DeRenzi et al., 1994; Schweinberger et al., 1995) makes it difficult at present to establish with certainty a link between neuroanatomical and functional sub-types.

Finally, while prosopagnosia is usually a result of acquired brain damage, an increasing number of cases have been reported with developmental prosopagnosia, a condition which per definition is present from early childhood, in the absence of neurological history. A familial factor has been reported for some cases, possibly pointing to a genetic contribution (DeHaan, 1999). Covert recognition usually has been found to be absent in developmental prosopagnosia (Campbell and DeHaan, 1994; Bentin et al., 1999), and it has been argued that this is because the condition precludes the establishment of face memories (Barton et al., 2001). It is noteworthy that a recent fMRI study has found an absence of the normal face-selective responses in the fusiform gyrus and the inferior occipital gyrus in one developmental and two childhood prosopagnosics, suggesting that these two areas are essential parts of the neural network subserving face recognition (Hadjikhani and DeGelder, 2002).

To summarize, one needs to distinguish at least two forms of prosopagnosia: apperceptive and associative. Importantly, there is a relationship of these variants to covert recognition, in that covert recognition appears to require a degree of successful encoding of facial representations. This would not seem to support the idea that covert recognition is mediated by a neural system that is completely independent from the neural system for overt face recognition.

3. HOW CAN COVERT RECOGNITION BE EXPLAINED IN TERMS OF FUNCTIONAL MODELS OF FACE RECOGNITION?

Models of normal face recognition have typically been developed on the basis of evidence from neuropsychological, as well as experimental studies. For
example, one of the most influential models (Bruce and Young, 1986) makes explicit a dissociation between processes used to compute identity and facial speech. This separation of processes was incorporated into the model on the basis of neuropsychological double-dissociations (Campbell et al., 1986). Similarly, an implemented development of part of this model, first published by Burton et al. (1990; see Burton et al., 1999 for a recent update), was influenced in part by evidence from patterns of identity processing which were present, and not present, in the neuropsychological literature of the time. In turn, these models have been used to account for further neuropsychological phenomena, which were not explicitly built in to their design, and the cross talk between modelling and neuropsychology has been fruitful (Breen et al., 2000; Burton et al., 1991; Farah et al., 1993; O’Reilly and Farah, 1999; Young and Burton, 1999).

One of the contributions of modelling approaches to understanding neuropsychological phenomena is that they force one to ask what is necessary for a working system to comprise. This is well-illustrated by functional accounts of covert recognition in prosopagnosia. As we described above, some accounts of covert recognition propose a secondary, “unconscious” face processing mechanism, which remains intact while an explicit system is damaged (Bauer, 1984). More recent accounts propose damage not to the face recognition system itself, but to the processes which signal awareness (DeHaan et al., 1992). These two classes of explanation cannot be rejected on the basis of current evidence. The first remains popular, in different variants, as we will describe below. The second remains possible, but is somewhat unsatisfactory, since it appeals to processes (those which signal awareness) which are not at all understood. However, a third class of explanation has attracted considerable attention. Using rather different types of model, Burton et al. (1991) and Farah et al. (1993) both demonstrated that it is possible to implement a single-route model of face recognition which can be damaged in such a way that overt recognition is compromised, while covert recognition can be observed due to residual activation in the damaged system. Both models propose that the behavioural tests used to measure covert recognition (face-name relearning, associative priming) are more sensitive than measures of overt recognition, and so it is possible to observe their effects in the absence of overt recognition. One of the main appeals of implemented accounts such as these is parsimony. In short, modellers tend to resist addition of extra routes (for example for face recognition) unless they are forced to do so by the evidence. Perhaps the major contribution of these models is that they demonstrated that it is not necessary to propose a second route of face processing on the basis of behavioural measures of covert recognition. We will return to this point when discussing the “affective route” below.

It is worth reviewing the Burton et al. (1991) and Farah et al. (1993) models briefly here, as they are relevant to the papers by Bobes et al., and by Sperber and Spinnler in this issue. The Burton et al. proposal is essentially a disconnection hypothesis, which posits damage to links between FRUs (units coding visual aspects of faces) and PINs (units coding people, accessible through any recognition route, and signalling familiarity when highly active). In this
way, recognition is possible through other routes (for example name recognition routes reach PINs as normal), but activation of (intact) face representations produces only small activation at PINs. This small activation is insufficient to signal familiarity, but is sufficient to modulate sensitive measures of covert recognition such as priming. The Farah et al. model is rather different in detail. In this model, recognition routes converge on information held about individuals (rather than at a prior PIN stage). Furthermore, familiarity is apparently signalled separately for each recognition route (names, faces and so forth). Their proposal is that the visual representations of faces themselves are damaged, by disruption of the computational processes allowing such representations to form. In this way, only poorly-formed visual representations are possible and these, in turn, pass on information which is impoverished to the extent that covert recognition is possible.

How do these models differ in their predictions? Although the mechanisms proposed are rather different in each, they are both intended to capture the same set of phenomena, those associated with covert recognition in associative prosopagnosia. Detailed comparisons can be found in Young and Burton (1999) and O’Reilly et al. (1999). However, it is worth asking whether they can be differentiated on the basis of the two papers presenting new data in this issue. It seems to us that the case of Emma (Sperber and Spinnler, 2003) is equally explicable in terms of either model. Both models have points of convergence between face and name recognition, and though these are different (modality-free PINs in one, stored semantic knowledge in the other) it is relatively simple to imagine systematic degradation of either convergence point leading to the pattern described for this patient. However, in the case of the study by Bobes et al., (2003, this issue) there seem to be differential predictions. The task set for subjects, and the patient, is rather a hard one. Subjects have to match previously unfamiliar faces, and do so sequentially. In fact, unfamiliar face matching is surprisingly difficult, even when faces are presented simultaneously, and subjects are under no time pressure (Bruce et al., 1999). Subjects are excellent at matching two identical images, but even very small changes in viewing angle, lighting or capture device, lead to quite significant detriments. In order to carry out this task sequentially, it seems necessary to form rather good visual representations. Furthermore, this is required for both an explicit (behavioural) match, and a covert (ERP) measure. It seems to us, then, that an account of associative prosopagnosia relying on damage to the visual representations of faces is unlikely to provide an explanation of the patient described in this paper. Instead the account by Burton et al. seems preferable, as it assumes intact visual representations, with damage occurring downstream, in connections to higher level knowledge.

Note that if this were a familiar face matching task, subjects would have further strategies available to them. For example, if they could label the first of a pair “George Bush” then they might compare this with a label generated for the second. This clearly would not be an option for patient FE. Although we can only speculate in the absence of further data, it is possible that something like a labelling strategy accounts for his failure in the behavioural part of this study. Perhaps it is necessary to form some person-level label in order to make explicit
matches, rather than relying on low level visual information. Whether this suggestion has any value remains to be seen, however, the major point is that the study appears to discriminate between two versions of the “residual activation” account which have attracted attention in the literature.

We now turn to a novel issue, not commonly discussed in the covert recognition literature, but which we predict will become important for future studies. As we have already discussed, a key distinction in prosopagnosia is between associative and apperceptive deficits. We are particularly interested in the associative form here, because it is here that covert recognition is most clearly established. However, it is worth briefly pointing out that the more perceptual aspects of face processing are certainly not unitary. It is therefore perhaps surprising that the focus of research in this area has been on identification only. Most models of face processing assume that there are a number of (to some extent) independent processing routes for computing different aspects of the face. The most common are identity, expression, facial speech and sex (Bruce and Young, 1986). Nevertheless, with very few possible exceptions (DeGelder et al., 2000), models of covert recognition focus almost exclusively on the computation of identity, rather than any of these other types of information which are available in face images. In most reports of associative prosopagnosia, patients are reported to have preserved ability to perceive all of these other variables, and this is provided as evidence that the deficit is not apperceptive.

In the literature on normal face recognition, there is beginning to be a debate about the mechanisms underlying different types of decisions on faces. A key distinction here is whether some decisions are actually supported by independent aspects of the stimuli themselves. So, for example, could it be that some aspects of the physical stimulus in a face photograph code the person’s sex, while others code the expression or identity? There are two types of study which support this view. Firstly, Schyns and coworkers (Gosselin and Schyns, 2001; Schyns et al., 2002) have shown in a series of experiments that different parts of the face, analysed at different spatial scales, are used selectively by subjects to make different decisions (for example sex and expression). Second, studying the physics of images (rather than perceptions of them) Calder et al. (2001) have shown that expression and identity can be separated by simple statistical procedures such as principal components analysis. These types of study begin to suggest that the visual analysis of faces may be rather modular, with specific perceptual processes having become specialised on different physical signals.

This work in normal face recognition opens up the possibility that neuropsychological studies of deficits in face processing might become more fine-grained. Rather than categorising prosopagnosia into apperceptive and associative, it may be possible to find sub-types of apperceptive prosopagnosia, in just the same way that the associative form has become more finely taxonomised (as is the main focus of this paper). A dissociation between, say, perception of expression and perception of sex, would allow very interesting further investigations into the sources of information used for these various types of perception in the normal case. In fact, this proposal has an attractive link to early discussions of prosopagnosia. Sergent (e.g., 1985, 1989) has discussed at
some length the relationship between the physical characteristics of the incoming face image, and the anatomical and cognitive mechanisms which must process it. She has suggested (see for example Sergent, 1989) that neuropsychological deficits such as prosopagnosia may be tied to more fundamental deficits in processing specific spatial frequencies of light which are particularly important in face identification. Although this detailed proposal has not been upheld, the more general idea, that there is a link between physical aspects of the stimulus, and computation of specific information from a face, is one which seems to have considerable value. We hope that this might be explored in further reports of prosopagnosic patients.

Finally, we return to the notion of two separate routes for recognition of identity, one measured with behavioural tests (face learning, priming) and another with physiological measures such as SCR. It seems clear that these two classes of measure are dissociable. However, it seems less clear that such a dissociation requires that we posit the existence of an affective face recognition mechanism. In particular, it is commonly hypothesised that the SCR provides an index of a processing system whose primary purpose is to orient the perceiver to the face emotionally. However, this is not yet proven. In fact, the SCR is rather a blunt tool for measuring response, and appears primarily to provide an index of familiarity-discriminating between known and unknown faces (seen and unseen faces in the case of the paper by Bobes discussed here). The hypothesis that this signals emotional response is an ingenious one put forward, for example, by Ellis and Young (1990) to account for Capgras delusion, and other related neuropsychiatric conditions.

This hypothesis appears to have become widely accepted. For example, describing SCR measures to familiar faces Breen et al. (2000) write “The affective response is, in part, a measure of how familiar something is, but the arousal/orienting response is a more primitive response to stimuli in one’s environment” (p. 66). This statement seems to us to be at least debatable. First, the conflation of arousal and orienting responses seems unwarranted. Arousal is uni-polar (and partly indexed by SCR) whereas orienting response is at least bi-polar. Second, can we really take it for granted that a familiarity mechanism is less “primitive” than the arousal/orienting response? On one view, the familiarity mechanism is logically prior to any response based on knowledge of the stimulus.

The proposal that the SCR measure (and other physiological measures) reflects an independent affective processing route for faces, in fact has some unparsimonious characteristics. Such a route would need to replicate some processes already executed by the more general “cognitive” face perception system. For example, each requires an independent mechanism for deciding whether a face is familiar or not. On a strong two-route hypothesis, such as that put forward by Breen et al. (2000), this requires parallel analysis of output from an FRU stage into PIN-level analysis (is this a familiar person?) and affective analysis (what is my response to this person?). Since it is impossible to compute an appropriate affective response without individuating the person, this model requires considerable replication of function.

An alternative to the currently popular view is that physiological indices of
Covert recognition reflect the fact that there is more than one way to compute the familiarity of a face. There seems to be clear evidence that there can be covert computation of familiarity, in the absence of explicit recognition, and that this can arise in at least two, dissociable ways, i.e., behaviourally and physiologically. One mechanism for computing familiarity (at the level of PINs) operates in a modality-independent manner (Burton et al., 1990, 1999). As suggested by some recent data from SCR responses, there may be a second mechanism that computates familiarity in a modality-specific manner for faces but not for names or voices (Ellis et al., 1999; Lewis et al., 2002). However, it seems to us that this is the limit of the necessary implications of the data. A system to compute affective response needs much more computational power than a system to compute familiarity, because in the former case one needs to identify the stimulus. If there are indeed two routes to familiarity, then one can imagine that a mismatch between these could produce phenomena of the kind associated with delusional misidentification (Ellis and Young, 1990). Furthermore, it is evident that at some stage in the process of person recognition, humans have the facility to compute affective response to a facial stimulus. However, whether such processes are integrated into the face recognition process, to the extent of replicating some existing function, remains to be proved. We anticipate that these are the types of issue which will be the focus of future computational modelling within this field.

4. Can One Relate Functional Processes to Brain Processes as Observed in Neurophysiological or Metabolic Recordings?

Our current knowledge about the neural substrates for face perception in the intact human brain largely derives from studies investigating either metabolic (positron-emission tomography, PET; functional magnetic resonance imaging, fMRI) or electromagnetic (event-related potentials, ERPs; event-related magnetic fields, ERMFs) signal changes associated with face processing. The relative merits of these techniques are reviewed elsewhere (Gazzaniga et al., 2002), and for present purposes we merely wish to emphasize that they have complementary advantages: fMRI and PET measure changes in brain glucose or oxygen metabolism secondary to changes in task-related neural activity; these methods offer a relatively detailed localization of brain activity related to a particular task, but with poor temporal resolution in the order of a few seconds. In contrast, ERPs and ERMFs directly measure changes in task-related neural activity with millisecond time resolution, but - unless intracranial recordings can be used - these methods only offer a limited localization of the brain generators involved. A detailed review of fMRI and PET findings in face perception is beyond the scope of this viewpoint and can be found elsewhere (Haxby et al., 2000). Similarly, we cannot in detail discuss findings from intracranial ERP recordings (Allison et al., 1994a, 1994b, 1999; McCarthy et al., 1999). However, we will comment in more detail on scalp-recorded ERP correlates of face processing as have been studied by Bobes et al. (2003, this issue).
fMRI/PET studies. Haxby and coworkers (2000) identified three core areas involved in face perception. In brief, the inferior occipital gyri are thought to mediate the visual analysis of faces, and project to two regions: The lateral fusiform gyri in ventral occipitotemporal cortex are thought to analyse the invariant aspects of faces, and thus to mediate the recognition of identity. The superior temporal sulcus region is thought analyse changeable aspects of faces, and thus to mediate the perception of eye gaze, expression, or mouth movements (see also Allison et al., 2000). These core areas are supplemented by several additional areas that perform further processing in concert with other systems. With respect to person recognition, the most important of these are anterior and middle temporal regions, which are thought to represent semantic/biographical information about people (Gorno-Tempini et al., 1998), and the left anterior temporal pole which may mediate name retrieval (Grabowski et al., 2001).

ERPs: VPP and N170. ERP research has identified a number of face-specific responses within about 200 ms. These include a vertex-positive potential (VPP, Jeffreys, 1989) and a more recently identified N170 component, an electrically negative wave over occipito-temporal areas approximately 170 ms after the onset of a face. VPP and N170 waves have similar latencies and respond in a similar way to experimental manipulation, such that the possibility remains that both originate from the same source (Itier and Taylor, 2002). The N170 is prominent for faces but absent, or attenuated for visual stimuli other than faces (Bentin et al., 1996). However, it should be noted that an N170 has also been reported for printed words (Bentin et al., 1999), and there seem to be limits to its specificity for faces (Eimer, 1998; Rossion et al., 2000) see also the recent controversy in (Rossion et al., 2000, 2002; Bentin and Carmel, 2002).

The N170 is unaffected by whether or not a face is familiar (Bentin and Deouell, 2000; Rossion et al., 1999; Eimer, 2000). In contrast, this component is larger and delayed in latency for inverted (as compared to upright) faces which are thought to be more difficult to encode into a holistic representation (Rossion et al., 2000). Interestingly, a face-elicited N170 response was found to be absent in a patient with apperceptive prosopagnosia (Eimer and McCarthy, 1999). All these findings suggest that the N170 reflects structural encoding, rather than individual recognition of faces.

While this is currently the dominant interpretation of the N170, a problem for this account may be that the N170 is seen for line-drawings of faces just as for real faces (e.g., Bentin et al., 2002). In contrast, behavioural recognition rates for real faces drop dramatically if information about texture and light-dark gradients in the image are unavailable (Bruce et al., 1992). It is therefore difficult to relate the N170 to structural encoding of representations for individual recognition of real faces. A possibility is that the N170 reflects the detection of a facial pattern – similar to the face detection stage by which the stimulus is classified as a face, suggested e.g. by Ellis (1983). What seems clear is that the N170 is related to perceptual analysis rather than individual recognition. This component may be generated in the region of the inferior occipital gyri, e.g. the occipitotemporal sulcus (Bentin et al., 1996). A recent study attempted to localize the N170 generators using inverse dipole localisation techniques, and has produced results that were consistent with this proposal (Schweinberger et al., 2002b). The scalp-
recorded N170 therefore probably does not reflect the identical phenomenon as the intracranial N200 (Allison et al., 1994b), a face-specific response from a small cortical patch of fusiform gyrus. Moreover, although the degree to which the N170 ERP response is related to activation of inferior occipital gyri seen in fMRI studies is unclear at present, we think this will be an interesting issue for further research.

**ERPs: N250r.** In the latency range between about 200 and 300 ms, a number of studies have identified a prominent ERP modulation caused by face repetition priming, termed either “visual memory potential”, “early repetition effect”, or N250r (Begleiter et al., 1995; Schweinberger et al., 1995, 2002b). The N250r is an increased negativity maximal over inferior temporal electrodes close to the mastoids, particularly over the right hemisphere, following repeated as compared with unrepeated faces. (The N250r is best seen when using an average reference in combination with large number of electrodes that include inferior temporal locations. In several studies that used mastoids or earlobes as reference electrodes, this component may have been either unnoticed or interpreted as a frontal modulation, just like such a choice of reference electrodes may have led to the interpretation of a temporal N170 negativity as a VPP in some studies).

The N250r is reduced for unfamiliar as compared to familiar faces (Begleiter et al., 1995; Schweinberger et al., 1995; Pfütze et al., 2002), suggesting that it may be related to individual face recognition. The N250r modulation to face repetition survives only a few intervening stimuli and is not seen for long-term repetition priming (Schweinberger et al., 2002a). It currently seems most likely that the N250r reflects a transient activation of facial representations for recognition. However, the N250r is reduced in amplitude for face repetitions that involve different images (Schweinberger et al., 2002b), a finding which is difficult to reconcile with the concept of completely image-independent FRUs. Of particular interest, inverse dipole localisation techniques have implicated the fusiform gyrus as the generator of N250r (Schweinberger et al., 2002b). A number of fMRI studies also identified the fusiform gyrus as a region that is sensitive to face repetition priming (Gauthier et al., 1999; Henson et al., 2000). Again, the degree to which these two phenomena are related will require further investigation.

The extent to which right-lateralized N250r is a face-specific response is also not completely clear. On one hand, there is some indication for a similar ERP modulation for other visual stimuli (Zhang et al., 1997; Rossion et al., 1999), and our own studies have indeed identified a similar N250r also for written names. On the other hand, the name-elicited N250r is not only smaller in amplitude, but also clearly lateralized to left temporal regions (Pfütze et al., 2002; Pickering and Schweinberger, 2002). Thus, it is possible that the N250r reflects repetition-sensitive activation of representations for the recognition of familiar visual stimuli in stimulus-specific cortical modules in ventral temporal (fusiform) regions, but that modules for the recognition of different types of stimuli (e.g., faces, objects, letter strings) may be located very close together (Allison et al., 1994b), albeit with different gradients of lateralization.

**ERPs: N400.** The N400 is a well-known component that is associated with semantic rather than perceptual processing. N400 is a central-parietal negativity
around 400 ms after stimulus onset that was initially demonstrated to reflect the
detection of semantic incongruency in sentences (Kutas and Hillyard, 1980). This ERP component is also seen for the second word of a pair when preceded by a nonassociated word (e.g., dog-table), and is reduced both when the second word is a repetition of the first (table-table) and when it is a semantically associated word (chair-table) (Rugg, 1987). Importantly, an N400-like reduction of central-parietal negativity around 400 ms can also be elicited by faces, when these are preceded by the same or by associated faces. This has been observed both in matching tasks similar to that used by Bobes et al. (2003, this issue), and in priming tasks in which a familiarity decision was made for the second face (Barrett et al., 1988; Barrett and Rugg, 1989; Schweinberger et al., 1995; Schweinberger, 1996; Bobes et al., 2000). It has been further shown that associative priming causes an N400-like modulation with equivalent topography regardless of whether faces or names are used as stimuli (Schweinberger, 1996). This suggests that the N400 effect reflects the facilitation in accessing postperceptual or semantic memory codes for people (Person Identity Nodes or Semantic Information, in our model’s terminology). At the same time it is noteworthy that associative priming did not influence the N170 or the N250r components, suggesting that these earlier components are not sensitive to differences in semantic processing. Most authors agree now that the generators of the N400, at least for language stimuli, are likely to be found in anterior ventral temporal lobe regions (McCarthy et al., 1995; Johnson and Hamm, 2000). fMRI studies on semantic processing of faces (see above) implicate broadly similar regions. We therefore assume that the N400 elicited by faces and common words or pictures probably reflect related but not necessarily exactly the same phenomena. Interestingly, there is some evidence from neuropsychology that semantic knowledge about familiar people can be selectively preserved or spared relative to other semantic knowledge, indicating that it may be subserved by its own neural subsystem (Lyons et al., 2002; Gentileschi et al., 2001). In that respect, a direct comparison between N400 elicited by faces and common objects might be relevant, and does not seem to have been performed as yet. Finally, we note that an ‘N400f’ component, a negativity between 300 and 500 ms that was larger for familiar as compared to unfamiliar faces, has been recently reported (Bentin and Deouell, 2000; Eimer, 2000). It is surprising that only a few ERP studies have directly studied differences between familiar and unfamiliar faces. While both the Bentin and Eimer studies make a step in that direction, it should be noted that in their studies familiar faces differed from unfamiliar ones not only by having an established visual representation or FRU, but also by allowing the access to semantic information. In fact, Bentin and Deouell (2000) relate the N400f to semantic processing involved in familiar face recognition. While these studies do not allow separation of neural correlates of visual recognition from correlates of semantic processing of a familiar face, another recent ERP study (Paller et al., 2000) has begun to address this issue. In this study, faces were experimentally familiarized and the ERP differences between learned and new faces were assessed. Some faces (“unnamed faces”) were learned purely visually whereas others (“named faces”) were supplemented
with biographical and name information during learning. Compared with new faces, named faces elicited more positive ERPs between 300 and 600 ms in the recognition test, both at anterior and posterior regions. Unnamed faces also elicited more positivity than new faces, but only at posterior regions. The posterior portion of this old-new difference was therefore interpreted as a correlate of retrieval of visual facial information, whereas the anterior portion was taken as an indication of retrieval of person-specific semantic/name information. In contrast to the abovementioned studies by Bentin and Deouell and by Eimer, Paller et al. did not find a larger N400f for familiar than unfamiliar faces, but task and/or stimulus differences may account for the discrepancy between these results.

Overall then, the last few years have seen various links emerge between neurophysiological phenomena and cognitive models of face recognition (Burton et al., 1990, 1999), with respect to at least three ERP components: The N170 is related to the detection and/or structural encoding of a face that is likely to be mediated by posterior lateral occipitotemporal cortex areas. The N250r is related to transient changes in activation levels of facial representations that can be elicited by repetition priming and are most likely mediated by ventral temporal areas, in particular the fusiform cortex. Finally, the N400 to faces reflects semantic processing of these stimuli in more anterior ventral temporal areas.

Based on the evidence reviewed above, we would argue that the combination of cognitive modelling and neurophysiology has already provided extremely valuable information with respect to the neural system that mediates the recognition of familiar faces. In our view, one of the biggest challenges for the future will be to study how the brain mediates the learning of new faces (Burton, 1994; Sommer et al., 1991, 1995, 1997). During familiarisation, one’s visual representation moves from being an inflexible, image-specific coding of a face, to a very flexible representation which can be used across a range of transformations in the image. There are no satisfactory explanations for this transformation in current theories of face recognition. Burton et al. (1999) argue that the processes underlying face learning are the same as those underlying repetition priming. While it remains unclear how the brain establishes stable representations of new faces, we would anticipate that this type of issue is one of the major challenges for future research on the neural system for face processing.

Acknowledgements. The authors’ research is supported by grant ref. 17/S14233 from the Biotechnology and Biological Sciences Research Council (U.K.).

REFERENCES


Covert recognition in prosopagnosia


