COGNITION

Cognition 85 (2002) 189–196

www.elsevier.com/locate/cognit

Discussion

A defense of the subordinate-level expertise account for the N170 component

Bruno Rossion[a,b,*], Tim Curran[c], Isabel Gauthier[d]

[a]*Department of Cognitive and Linguistic Science, Brown University, Providence, RI, USA*
[b]*Cognitive Neuroscience Unit, University of Louvain, Louvain-la-Neuve, Belgium*
[c]*Department of Psychology, University of Colorado, Boulder, CO, USA*
[d]*Department of Psychology, Vanderbilt University, Nashville, TN, USA*

Abstract

A recent paper in this journal reports two event-related potential (ERP) experiments interpreted as supporting the domain specificity of the visual mechanisms implicated in processing faces (*Cognition 83* (2002) 1). The authors argue that because a large neurophysiological response to faces (N170) is less influenced by the task than the response to objects, and because the response for human faces extends to ape faces (for which we are not expert), we should reject the hypothesis that the face-sensitivity reflected by the N170 can be accounted for by the subordinate-level expertise model of object recognition (*Nature Neuroscience 3* (2000) 764). In this commentary, we question this conclusion based on some of our own ERP work on expert object recognition as well as the work of others. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: N170; Face processing; Expertise; Subordinate-level categorization

---

The N170 is the first posterior negative deflection following the visual presentation of a picture of a face, peaking at occipito-temporal sites at around 170 ms (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Bötzel, Schulze, & Stodieck, 1995). Bentin et al. (1996) characterized the response properties of the N170 using a target detection task in which various pictures of faces and other objects (e.g. flowers, cars) were presented and subjects were monitored for the appearance of butterflies (target) in the sequence. In that early study, the N170 response was small or described as absent for non-face stimuli as the negative deflection did not cross the zero baseline level of the EEG. Since then, several event-related potential (ERP) studies have described large and clear negative N170 components following the presentation of faces as well as non-face objects (e.g. Curran,

---

* Corresponding author. Department of Cognitive and Linguistic Science, Brown University, 190 Thayer Street, Box 1978, Providence, RI 02912, USA. Tel.: +1-401-863-1678; fax: +1-401-863-2255.
E-mail address: bruno_rossion@brown.edu (B. Rossion).

PII: S0010-0277(02)00101-4

Tanaka, & Weiskopf, in press; De Haan, Pascalis, & Johnson, 2002; Eimer, 2000a,b; Rossion, Gauthier, Goffaux, Tarr, & Crommelinck, 2002; Rossion et al., 2000; Tanaka & Curran, 2001; Taylor, McCarthy, Saliba, & Degiovanni, 1999). N170 responses are recorded for objects (e.g. houses, chairs, cars, birds, dogs, shoes, or novel objects such as Greebles or 2D "blob" shapes) often at latencies comparable to faces but always at lower amplitudes. N170 responses of large amplitude are also recorded in response to visually-presented words (Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999).

Alternative theoretical interpretations exist for explaining N170 differences between faces and other objects. Carmel and Bentin (2002) (C&B) suggest that their "results demonstrate the domain specificity of the visual mechanism implicated in processing faces, a mechanism which is not influenced by either task or expertise" (C&B, abstract). This "domain-specific" account can be contrasted with the "subordinate-level expertise" account (Rossion et al., 2002; Tanaka & Curran, 2001). First, the N170 can vary widely in amplitude for different object categories (e.g. chairs vs. cars; Rossion et al., 2000), suggesting that it could be generally related to object categorization (for a similar argument in fMRI, see Haxby et al., 2001) and that the processes reflected by the N170 are not uniquely engaged by faces (see Tovée, 1998). Second, the N170 increases in amplitude when the same object is categorized at the subordinate vs. the basic level (e.g. beagle vs. dog; Tanaka, Luu, Weisbrod, & Kiefer, 1999), and when dog and bird experts – who recognize their objects of expertise automatically at the subordinate level – are presented with their objects of expertise (dogs/birds; Tanaka & Curran, 2001). Together with other recent studies (e.g. Curran et al., in press; De Haan et al., 2002; Rossion et al., 2002) this suggests that the N170 reflects the activation of many different sub-processes (and areas) in the occipito-temporal pathway, and can be used as a tool to study the parameters affecting entry-level object categorization, including faces. Among these parameters, the level of categorization and degree of visual expertise seem critical in the increase of amplitude of the N170, and in explaining at least part of the difference between faces and objects (see Tarr & Gauthier, 2000).

The experiments reported by C&B aimed to challenge the subordinate-level expertise view by showing that the N170 is immune to task factors and the level of expertise for faces. The first experiment reported by C&B included two tasks: (1) a 'car-monitoring task', in which participants were shown photographs of four equally probable categories (human faces, birds, cars and furniture items), and instructed to detect cars; and (2) an animacy decision task, in which the same four types of stimuli were presented and participants categorized each object as animate or inanimate. In both tasks, faces elicited a larger N170 component than cars (the pre-determined target in the first task), which was in turn larger than for birds and furniture. The amplitude difference between faces and cars was especially reduced in the car-monitoring task, being only significant at the mastoid sites. Interestingly, the larger N170 for cars compared to other non-face categories was also found in the animacy detection task, when cars were not the target of interest. However, the larger N170 for cars in the animacy detection task was found only in participants who performed the car detection task first (the order of task was counterbalanced across subjects). Moreover, cars were the only stimuli to present an amplitude modulation due to the task, with the N170 to faces immune to task effects. According to C&B, these results support the domain-specificity hypothesis because the N170 to faces is immune to strategic manipulations, whereas the N170 to other categories can be influenced by task or attentional factors, suggesting that it does not represent the same processes.

These results are interesting, particularly the fact that participants who performed the car-monitoring task first somewhat ‘carried out’ the ERP effect to the second task. This demonstrates how task factors can be important in modulating the response of relatively early visual potentials, even implicitly. However, our interpretation is that these results do not truly challenge the “subordinate-level expertise” model. This first experiment shows that attending to cars increases the N170 response for cars: this is consistent with the extensive literature showing increases of amplitude for early visual potentials when objects are attended vs. unattended (for a review, see e.g. Luck, Woodman, & Vogel, 2000). C&B argue that the N170 response to faces is immune to attentional modulation or task factors, but unfortunately they did not include an explicit manipulation of the attentional load on faces (no behavioral data were reported either).[1] In fact, in a previous ERP study, when the attention load to centrally-presented faces was explicitly manipulated, attention enhanced the amplitude of the N170 to faces but did not affect the N170 amplitude to chairs (Eimer, 2000b). This suggests that attention can modulate the N170 to both faces (Eimer, 2000b) and other objects (C&B), making this issue a moot point in terms of comparing the domain-specific view to the subordinate-expertise level alternative. Furthermore, differentiating face and object processing by task factors is complicated by several considerations. Faces, contrary to most common objects, are categorized automatically at the individual level (Tanaka, 2001). As a result, instructions to categorize faces at different levels (e.g. “detection” vs. “individualization”) may not necessarily influence face processing and its neural correlates (Tarr & Gauthier, 2000). The primacy of the individual level, leading to smaller effects of task manipulations, is also observed for non-face objects in cases of visual expertise (Gauthier, Skudlarski, Gore, & Anderson, 2000; Gauthier, Tarr et al., 2000; Tanaka & Taylor, 1991). This has at least two important consequences. First, if faces call automatically for attention (Cauquil, Edmonds, & Taylor, 2000) and/or automatized expertise-related processes (Tanaka, 2001), it is expected that the N170 to faces will show little modulation by task factors (see Gauthier, Tarr et al., 2000, for relevant evidence in a fMRI study). Second, comparing faces and objects in passive presentation paradigms or oddball tasks (when the objects compared require no response or have to be ignored) may not be adequate because faces or objects of expertise will automatically recruit additional processes as compared to categories for which one is not expert. Therefore, attentional effects (or the lack thereof) due to task instructions need to be interpreted in the context of attentional differences due to one’s acquired processing biases for a category.

The second ERP experiment reported by C&B aimed at testing directly the role of visual

---

[1] In their second experiment, comparing human and ape faces, C&B asked their subjects to respond to human faces. They then compared the N170 responses to human faces between experiments, showing no modulation of the N170 amplitude. However, the comparison of the N170 – which presents a high degree of variability even within participants – across the two experiments is complicated by several confounding factors. The proportions of targets differed between experiments (cars were 25% in the first experiment, whereas human faces were 50% in the second), the similarity of distractors differed between experiments, and all subjects completed Experiment 1 before Experiment 2 in the same session (confounding the order of experiments).

expertise in shaping the N170 response. Interestingly, a recent study (Tanaka & Curran, 2001) revealed that the N170 amplitude is larger for bird and dog experts than novices for these categories. Another recent study (Rossion et al., 2002) demonstrated that picture-plane inversion of artificial non-face objects of expertise (i.e. Greebles) leads to a delay of the N170, similarly to what is usually found only for face inversion (De Haan et al., 2002; Eimer, 2000a; Rossion et al., 1999, 2000; Sagiv & Bentin, 2001). However, C&B's experiment differed notably from these studies in that it included a manipulation of the level of expertise *within* the face domain, by comparing the N170 to facial configurations for which we are experts (human faces) to facial configurations for which we are supposedly not experts (ape faces). The reasoning was that if expertise accounts for face specificity, the N170 elicited by human faces should be clearly distinguished from that elicited by ape faces. In contrast, the N170 should be similar for these two categories if the N170 response depends merely on the presentation of a face configuration. In this study, participants performed a "human face-monitoring task", pressing a key for human (but not for ape) faces. The choice of this task is surprising because it gives a different status to the two types of stimuli compared within the experiment, but justified by the goal of comparing the effect of explicit attention to a (human) face on the N170 in this experiment to the N170 obtained in the previous experiment (when faces were not the center of attention).

The N170 elicited by ape faces in human participants was as conspicuous as that elicited by human faces, or even larger (although not significantly), and significantly delayed (about 10 ms). According to C&B, this result suggests that the perceptual mechanism associated with the N170 is selectively engaged by human faces and by other stimuli that preserve perceptual characteristics of the human face, independently of expertise. This interpretation is questionable for several reasons. First, C&B's study does not manipulate the effect of expertise (or does it only to the extent that it is confounded with differences between categories). There is no independent measure of subjects' ability to discriminate the different stimuli, or the extent to which they are processed in a holistic manner (see Gauthier, Skudlarski et al., 2000; Gauthier & Tarr, in press for correlations between measures of expertise and holistic processing as well as with neurophysiological responses in face-selective areas). Expertise claims have been made previously by comparing groups of experts and novices on the same stimuli and/or empirically using well-established behavioral criteria to define the level expertise (Gauthier et al., 1998; Gauthier, Skudlarski et al., 2000; Gauthier & Tarr, 1997, in press; Gauthier, Tarr et al., 2000; Rossion et al., 2002; Tanaka & Curran, 2001; Tanaka & Taylor, 1991). We believe that direct demonstrations that the N170 is influenced by expertise (Rossion et al., 2002; Tanaka & Curran, 2001) outweigh the single null result reported by C&B. Second, contrary to what is advanced by C&B in their paper, behavioral experiments comparing human and monkey/ape faces do support the expertise hypothesis. For instance, the face inversion effect in animals and humans is found only for faces for which subjects have developed an expertise, providing that faces of the same ape species are grouped together (see Parr, Dove, & Hopkins, 1998). Finally, it is worth noting that C&B *do* actually find a significant N170 latency difference between human and ape faces (a 10 ms delay for ape faces) that would make sense in terms of the subordinate-level expertise model. This point can be better illustrated by referring to another ERP study also using monkey faces which was coincidentally published the same month as C&B's paper (De Haan et al., 2002). In this study, the N170 was recorded to the presentation of human and monkey faces, presented both upright and inverted, and in both human adults and infants. In adults, the N170 to upright monkey faces was of larger amplitude and longer latency than for human faces. These results are consistent with C&B who found a similar (though non-significant) amplitude difference as well as a consistent significant (10 ms) latency difference. However, the addition of inverted faces allowed De Haan et al. to draw a different conclusion: when human faces are presented upside-down, they also elicit an amplitude increase and delay (10 ms): the electrophysiological response for inverted faces completely overlaps with monkey faces. This inversion effect for human faces is well documented (e.g. Eimer, 2000a; Rossion et al., 1999, 2000; Sagiv & Bentin, 2001; see Rossion & Gauthier, 2002 for a review). As would be predicted by the expertise model, inversion did not affect the N170 to monkey faces in the De Haan et al. (2002) study. In sum, the N170 amplitude was different for upright faces than other stimuli (monkey faces upright and upside-down, and human faces upside-down), and inversion increased the amplitude and latency only for human faces. Supported further by ERP results from 6-month-old infants, which showed that a possible infant analog of the N170 was larger for human faces than objects but showed no inversion effect, De Haan et al. (2002, p. 206) concluded that "experience drives the specialization of the N170". Similarly, C&B's results of a latency difference between monkey and human faces at the level of the N170[2] can be interpreted as indicating that our expertise with human faces does not generalize completely to inverted faces, or to ape faces.

Finally, C&B discuss the different latency and scalp distribution of the N170 to faces and what they refer to as a visual "negative-going" component or N1 for objects. The question of what defines a component and when two ERP deflections can really be considered as different components is beyond the scope of this commentary (see Rugg & Coles, 1995). The N170 is the first negative posterior deflection that can be observed at occipito-temporal sites following the presentation of any 2D shapes, or even strings of letters (Bentin et al., 1999), and should accordingly be called a visual (posterior) N1 component. First, it should be noted that the N1/N170 does not peak earlier for faces than objects in most ERP studies (e.g. Eimer, 1998, 2000a; Rossion et al., 2000). In fact, the N170 can peak earlier for some object categories than for faces, and later for others (see Rossion et al., 2000 for a comparison with several object categories) and the latency variability of this component between object categories could easily be due to low-level visual differences between pictures such as contrast, luminance and spatial frequencies.[3] Second, given the multiple visual extrastriate areas activated in interlocked time courses following the presentations of visual stimuli, it is unlikely that this large component reflects a single processing stage originating from a particular brain structure (see Regan, 1989).[4] It is thus plausible that differences in the number and locations of sources

---

[2] The absence of a significant amplitude increase for ape faces compared to human faces in C&B's study is consistent with the general observation that the increase of amplitude for inverted faces is a less robust effect than the N170 peak delay. For instance, Bentin et al. (1996) also described a significant delay of the N170 to inverted faces, but no significant increase of amplitude.

[3] Note for instance that the N1/N170 delay for pictures of objects relative to faces in C&B's study appears to start at the level of the preceding P1.

[4] A point also emphasized by Sagiv and Bentin (2001).

contributing to the N1/N170 associated with different stimuli and tasks will give rise to some changes in the scalp distribution of this component. Similarly to its amplitude, the topography of the N1/N170 can appear slightly different between faces and objects, but also between different object categories (Rossion et al., 2000). Should we thus give a different label to this component for each type of visual stimulus to which it is evoked? Most ERP studies of visual processing record a clear negative N1/N170 component to pictures of non-face objects, and in our own studies (carried out with at least 64 channels), the topographical distribution of the N1/N170 is highly similar for faces and other object categories (e.g. see Rossion & Gauthier, 2002, Fig. 2). In any case, given the poor spatial resolution of ERPs, it is impossible to know whether ERP components in two conditions, even if identical in topographical distribution, latency and amplitude, represent the activity of the exact same neural network. Similarly, it is impossible to argue conclusively that ERP components in two conditions (here the N1/N170 to faces and objects) reflect the activity of completely independent neural networks because of differences in their scalp distribution. In summary, although the question of whether faces and objects of expertise recruit the activity of functionally overlapping neural substrates is an important question, the scalp distribution differences between faces and objects in novices emphasized[5] by C&B also present a moot point to resolve the debate between domain-specific and expertise accounts.

In conclusion, Carmel and Bentin (2002) offer a multifaceted attack on the subordinate-expertise account of the face-selectivity observed in the N1/N170. Unfortunately, their arguments made on the basis of attentional effects, the human response to ape faces and the scalp distribution of the N1/N170 response for faces and objects do not support the domain-specificity hypothesis. It is worth pointing out that the subordinate-expertise account of face-selectivity remains the only account of the development of this phenomenon, outside of the possibility that face-selectivity is rooted in an innate mechanism tuned to the face geometry. C&B do not challenge the existence of expertise effects on an object-elicited potential that is at the least extremely similar in space and time to the face N170. In this context, a new empirical question arises: is expertise the common cause for face-selectivity and selectivity to objects of expertise, or do we need to invoke two mechanisms, one innate and the other rooted in experience, to explain phenomena that are so similar?

Acknowledgements

This work was supported by the National Science Foundation and by the James S. McDonnell Foundation. The authors thank Martin Eimer and Gary Cottrell for comments on the paper, as well as the members of the Perceptual Expertise Network for many relevant discussions.

---

[5] Note that the statistical significance of such differences in scalp distribution is not tested in C&B's study and previous studies, and so it is hard to assess their robustness and the size of such effects.

References

Bentin, S., Allison, T., Puce, A., Perez, A., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*, 551–565.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. *Journal of Cognitive Neuroscience*, *11*, 235–260.

Bötzel, K., Schulze, S., & Stodieck, R. G. (1995). Scalp topography and analysis of intracranial sources of face-evoked potentials. *Experimental Brain Research*, *104*, 135–143.

Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: factors influencing distinct processing of faces. *Cognition*, *83*, 1–29.

Cauquil, A. S., Edmonds, G. E., & Taylor, M. J. (2000). Is the face-sensitive N170 the only ERP not affected by selective attention? *NeuroReport*, *11* (10), 2167–2171.

Curran, T., Tanaka, J. W., & Weiskopf, D. M. (in press). An electrophysiological comparison of visual categorization and recognition memory. *Cognitive, Affective, & Behavioral Neuroscience*.

De Haan, M., Pascalis, O., & Johnson, M. J. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience*, *14*, 199–209.

Eimer, M. (1998). Does the face-specific N170 component reflect the activity of a specialized eye processor? *NeuroReport*, *9*, 2945–2948.

Eimer, M. (2000a). Effects of face inversion on the structural encoding and recognition of faces – evidence from event-related brain potentials. *Cognitive Brain Research*, *10*, 145–158.

Eimer, M. (2000b). Attentional modulations of event-related brain potentials sensitive to faces. *Cognitive Neuropsychology*, *17*, 103–116.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*, 191–197.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: exploring the face recognition mechanism. *Vision Research*, *37*, 1673–1682.

Gauthier, I., & Tarr, M. J. (in press). Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*.

Gauthier, I., Williams, P., Tarr, M. J., Tanaka, J. (1998). Training "Greeble" experts: A framework for studying expert object recognition processes. *Vision Research*, *38*, 2401–2428.

Gauthier, I., Tarr, M. J., Moylan, J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (2000). Does visual subordinate-level categorization engage the functionally-defined fusiform face area? *Cognitive Neuropsychology*, *17*, 143–163.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed, and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*, 2425–2430.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, *4*, 432–440.

Parr, L. A., Dove, T., & Hopkins, W. D. (1998). Why faces may be special: evidence of the inversion effect in chimpanzees. *Journal of Cognitive Neuroscience*, *10*, 615–622.

Regan, D. (1989). *Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine*. New York: Elsevier.

Rossion, B., Delvenne, J. F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., & Guérit, J. -M. (1999). Spatio-temporal localization of the face inversion effect: an event-related potentials study. *Biological Psychology*, *50*, 173–189.

Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews*, *1*, 63–75.

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelinck, M. (2002). Expertise training with novel objects leads to left lateralized face-like electrophysiological responses. *Psychological Science*, *13*, 250–257.

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P. A., Bruyer, R., Linotte, S., & Crommelinck, M. (2000). The N170 occipito-temporal component is enhanced and delayed to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *NeuroReport*, *11*, 69–74.

Rugg, M. D., & Coles, M. G. H. (1995). *Electrophysiology of mind. Event-related brain potentials and cognition*, Oxford: Oxford University Press.

Sagiv, N., & Bentin, S. (2001). Structural encoding of human and schematic faces: holistic and part-based processes. *Journal of Cognitive Neuroscience*, *13*, 937–951.

Tanaka, J. W. (2001). The entry point of face recognition: evidence for face expertise. *Journal of Experimental Psychology: General*, *130*, 534–543.

Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science*, *12*, 43–47.

Tanaka, J. W., Luu, P., Weisbrod, M., & Kiefer, M. (1999). Tracking the time course of object categorization using event-related potentials. *NeuroReport*, *10*, 829–835.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: is the basic level in the eye of the beholder? *Cognitive Psychology*, *23*, 457–482.

Tarr, M. J., & Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience*, *3*, 764–769.

Taylor, M. J., McCarthy, G., Saliba, E., & Degiovanni, E. (1999). ERP evidence of developmental changes in processing of faces. *Clinical Neurophysiology*, *110*, 910–915.

Tovée, M. J. (1998). Face processing: getting by with a little help from its friends. *Current Biology*, *8*, R17–R320.