Autonomic responses to familiar faces without autonomic responses to familiar voices: Evidence for voice-specific Capgras delusion

Michael B. Lewis and Sarah Sherwood

Cardiff University, UK

Hamdy Moselhy

All Saints Hospital, Birmingham, UK

Hadyn D. Ellis

Cardiff University, UK

Introduction: Patients with Capgras delusion believe that certain individuals have been replaced by duplicates. Unlike normal people, these patients also show reduced autonomic responses to familiar faces, indicating the possibility that it is the covert processes of recognition that are impaired (Ellis, et al., 1997). It has been suggested that such patients would show normal autonomic responses to voices. An auditory parallel of this typical delusion, therefore, is theoretically possible. That is, a delusion whereby mis-recognition of the voice produces the delusional belief of duplication. Such a delusion would only occur in situations where the person is recognised by voice only; and so, even where it does exist, it would often escape diagnosis. Method: We present here a case, H.L., of what appears to be the Capgras delusion for voices in a sighted person. This case was investigated using standard skin conductance tests for face and voice recognition. Results: Consistent with this diagnosis, H.L. displays normal autonomic responses for faces but reduced autonomic responses for famous voices. Discussion: H.L. represents a previously unreported form of Capgras delusion and, further, shows dissociation between autonomic responses to faces and voices. Implications for cognitive models of person recognition are discussed.

In the bizarre condition of the Capgras delusion, sufferers believe that others, often close relatives, have been replaced by impostors (Capgras & Reboul-Lachaux, 1923; Christodoulou, 1977; Enoch & Trethowan, 1991). It has been
considered that Capgras delusion represents a dysfunction of the processes of face or person recognition; and, thus can be understood in terms of the sophisticated cognitive models that have been developed to explain face recognition. One such cognitive neuropsychiatric explanation for the delusion was offered by Ellis and Young (1990), based on the principle that Capgras delusion is the mirror image of the neurological disorder, prosopagnosia (Bodamer, 1947). This is predicated on Bauer’s (1984) suggestion that there are two dissociable routes by which face recognition normally takes place. One of these routes involves the overt recognition of a face. It is this route that is disrupted in prosopagnosia leaving the person unable to access information relating to the face. The second route is the covert route and conveys something about the emotional significance of the person encountered. This route is believed to be responsible for the autonomic responses observed with faces. One example of this autonomic response is the skin conductance response (SCR) to faces (Tranel, Fowles, & Damasio, 1985): Peoples’ skin conductance responses are larger when they see a familiar compared with unfamiliar face. If, as it may be for some prosopagnosics, the covert route is intact then one would expect to see the normal differential SCR pattern for famous and nonfamous faces. This, indeed, is what has often been observed (see Bauer, 1984, 1986).

Ellis and Young (1990) applied the principle of the dual-route account of face processing to understanding the Capgras delusion. Their hypothesis was that the Capgras delusion represents a double dissociation with prosopagnosia, so that it is the covert route that is disrupted and the overt route that is intact. In such a situation, the person would be able to report to whom a face belongs but would not experience the feeling of emotionality or familiarity normally associated with that person. For highly familiar people the difference between the feeling of familiarity that they would expect and the feeling they get would need to be resolved. One simple, if irrational, resolution of this state could be the belief that the person is an impostor—that is, looks like that person but is not him or her.

If Ellis and Young’s (1990) hypothesis is correct then those with the Capgras delusion should also show a lack of covert recognition in terms of their SCRs to faces. Specifically, they should show no differential responses for familiar compared with unfamiliar faces. This prediction has been tested by Ellis, Lewis, Moselhy, and Young (2000), Ellis, Young, Quayle, and de Pauw (1997b), and Hirstein and Ramachandran (1997). All these studies found that patients with Capgras delusion produced no differential SCRs to familiar and unfamiliar faces.

Breen, Caine, and Coltheart (2000) suggest two possible cognitive interpretations of the Ellis and Young (1990) account. In one form, there is a cognitive bifurcation prior to the face recognition system, that is, before the face recognition units (FRUs) in the Bruce and Young (1986) model. Such an account would lead to unnecessary duplication. The second, preferred, form is that the two cognitive routes bifurcate after the FRUs but before the multimodal person identification nodes (i.e., before the person identity nodes, or PINs in the
Bruce and Young model). After the FRUs the pathways bifurcate to the PINs and an affective response module that will mediate the SCR.

One criticism offered by Breen et al. (2000) of the work conducted on the SCRs of Capgras patients is the fact that it has not been established whether patients who show no differential SCRs to famous faces reveal normal responses to other stimuli. Typically, repeated tones or deep breaths are employed to establish that a good SCR signal is present (e.g., Ellis et al., 1997b; Tranel et al., 1985). Breen et al. (2000) suggest that an appropriate control task would be the study of the patients SCRs to famous and nonfamous voices. If Capgras delusion is purely a visual phenomena then we would expect patients to demonstrate normal SCRs to voices. This has not been done—probably because, until recently, it has not been known what would be the normal SCRs to familiar and unfamiliar voices. Sherwood, Lewis, and Ellis (2001), however, have established that voices, like faces, evoke differential SCRs determined by whether they are recognised or are unfamiliar. This finding with normal participants means that similar tests can be conducted on patients suffering from delusions; but it also means that it is necessary to update the representation by Breen et al. (2000) of the Ellis and Young (1990) account to include voice processing.

Although the suggestion by Breen et al. (2000) of using a voice task as a control task for studying Capgras delusion may be useful it does make an important assumption. This assumption is that the bifurcation of the two routes of recognition occurs at the level of the FRUs and not the multimodal PINs. All of the evidence described so far would be consistent with either location. If the bifurcation occurs at the FRUs then one would expect that SCRs to voice and face tasks could be differentially affected in patients with different forms of the delusion. If the bifurcation occurs at the level of the PINs, however, then identification of a person from any modality would lead to similar SCRs. In this case, if there is reduced differentiation for faces then there would also be reduced differentiation for voices.

Voices, like faces, have been subject to a degree of psychological research. Some of this research has been aimed at incorporating voice recognition processes into models of person recognition. Ellis, Jones, and Mosdell (1997a), for example, investigated short-term and long-term priming effects between voices and faces. It is considered that the processes of voice recognition occur in parallel to face recognition via voice recognition units (VRUs—equivalent to FRUs), which are connected with the multimodal PINs. This idea can be used to identify more precisely the cognitive location of the bifurcation of the routes to recognition.

Although there is no conclusive evidence for placing the bifurcation at either the FRUs or the PINs, there is some anecdotal evidence for the FRU location (see Figure 1). This evidence comes from Hirstein and Ramachandran’s (1997) patient who would claim that his father, when in front of him, was an impostor but did not suffer delusional mis-identification when interacting with him over the telephone.
There is a corollary to the model described in Figure 1, which concerns the possibility of disruption of the route between the VRUs and the affective response module. If such a disruption occurred, then the patient would experience normal recognition for people that he/she could see. If recognition took place by voice alone, however, then the feeling of familiarity would fail to be evoked. The contrast between overtly recognising a voice and failing to have a feeling of familiarity would have similar consequences to the situation where the visual route to familiarity has been disrupted. The patient would possibly believe that the voice belonged to an impostor. This would be an auditory counterpart to the more typical visual-Capgras delusion. The voice-specific Capgras delusion could possibly be harder to diagnose, however, because of the lesser importance of recognition by voice alone and, hence, the sufferer may not be aware of the delusion to the same extent as someone suffering from the visual-Capgras delusion.

There is an important exception to the lesser importance of voice-specific Capgras delusion and this is with blind patients. There are at least three reports of patients who, although blind, still suffered from Capgras-like delusions that
people (or a pet in one case) had been replaced by duplicates (Reid, Young, & Hellawell, 1993; Rojo, Caballero, Iruela, & Baca, 1991; Signer, Van Ness, & Davis, 1990). The existence of such cases demonstrates that Capgras delusion is not exclusively confined to mis-recognition of faces. Other senses, such as hearing or touch, can be similarly affected. Although these cases demonstrate the possibility of the existence of a voice-specific Capgras delusion, none of them involved recording SCR's to famous and nonfamous voices.

The prediction of the model described in Figure 1 is that it should be possible to find Capgras patients of two types. First, those who have normal SCR's to voices and reduced differential SCR's to faces. Second, those who have reduced differential SCR's for voices and normal SCR's for faces. This second type is less likely to be found, of course, as instances of voice-only recognition are not as common as face-only recognition. Further, in instances where the face and voice can both be used, the face is likely to be the primary source of information.

We now present a possible case of voice-specific Capgras delusion. Initial investigation of this patient focused on the possibility that she was suffering from the visual form of the delusion. It became apparent during an interview, in which she described a recent incident of the delusion, that we might be dealing with a voice-specific form of the delusion. This self-report was followed up by measuring SCR's to faces and voices.

Case H.L.

H.L. is a 59-year-old mother of three who was born in the West Indies but came to the UK in 1961 where she has worked at various jobs including a long period as a nurse. She has a long history of mental illness. In 1971, she was admitted to hospital for six weeks in a paranoid state. She was treated with Fluphenazine decanoate and was subsequently given a course of seven ECT's. Over the next 17 years she continued to take fluphenazine decanoate as well as chlorpromazine, nitrazepam and procyclidine. In 1988, she stopped taking fluphenazine as it made her tired. H.L. was admitted to hospital in 1989 with a diagnosis of hypomania. She also reported a number of persecutory delusions and failed to recognise her son was outside her door, saying ‘‘You don’t know me, you don’t live here’’.

Between the years 1989 and 2000, H.L. was admitted to hospital six times with a variety of mental disorders. At times she expressed the view that there were two versions of her son. At other times she expressed persecutory beliefs such that her room was bugged and a belief that Trevor McDonald, the newscaster, was the father of her son. She also suffered from migraines and phobia. During this time she was treated with thioridazine, haloperidol, Modecate, Clopixol, as well as the drugs already mentioned.

Clearer evidence of Capgras-like delusions occurred recently. In 1998, she reported that her vicar and some neighbours were impostors. Later the same
year, she said that her son was also an impostor and was admitted to hospital with a diagnosis of schizoaffective disorder. While in hospital, she reported that staff were not who they said they were. Three months prior to testing, H.L. locked her son out of her flat for 11 hours, believing that he was an impostor. Discussion with H.L. (see Appendix) revealed that it was likely that it was her son’s voice and not his face that she was using to (mis-)identify him.

At the time of testing, H.L. had been discharged from hospital and was being monitored by a community-based team. Her medications were Depixol 80 mg/3 weeks and paroxetine 20 mg once daily. She performed perfectly on the shortened versions of the Warrington Recognition Test for Words (25/25) and the Warrington Recognition Test for Faces (25/25). Medical assessment found her cognitive functions to be normal and she was orientated in time, place, and person.

THE EXPERIMENT

In order to assess H.L.’s autonomic responses to faces and voices, her SCRs were recorded using the standardised methodology (Tranel et al., 1985). If H.L. was indeed suffering from the voice-specific form of the Capgras delusion then we would expect her to have normal SCRs to faces and no differentiation to voices.

Method

Controls. Six control subjects were employed. These were all female and of a similar age to H.L. (mean = 62, range = 55–67) and none reported having any history of mental health problems. While these control subjects could not be matched on use of medication, the research conducted by Ellis et al. (1997b) demonstrated that neuroleptic and antidepressant medication do not lead to the differences in differential SCRs to faces found between Capgras patients and psychiatric controls. Although their study only explored SCRs to faces, there is no evidence to suggest, or reason to suppose, that SCRs to voices should be differentially affected by medication.

Stimuli. The stimuli consisted of 60 faces and 40 voices (the difference in number of stimuli is of no consequence as the face and voice task are analysed as two independent experiments). The faces were of 15 famous people and 45 nonfamous people. The images were cut down such that they only contained the face and hair of the individual and they were presented individually in colour on a computer monitor. The voices were of 15 famous people and 25 nonfamous people. The voice segments lasted between 5.5 seconds and 6.5 seconds and were presented via headphones. The speech content of the voices was controlled so that the identity of the individual could not be determined by what was being
said (e.g., famous catchphrases were avoided) but they were also long enough so that they could be recognised at least 50% of the time by controls.

**Autonomic responses.** The study involved the recording of skin conductance responses (SCRs) for each of the presented stimuli. The procedure for this was based on the standardised methodology outlined by Tranel et al. (1985). Two electrodes were attached to the participants’ nondominant hand. A small voltage was applied across these electrodes to obtain a measure for the conductivity of her skin. Each participant’s skin conductance was recorded for the entire presentation of the stimuli (faces or voices). The SCR was determined by the amplitude of any peaks that occurred in a window of 1–5 seconds following the initial presentation of a face or in a window of 1–9 seconds following the presentation of a voice (the difference in the window duration was because of the increased length of time required to recognise a voice rather than a face).

**Procedure**

**The face test.** Following the procedure detailed in Ellis et al. (2000), each face was presented for 2 seconds, with an interval of 20 seconds between stimuli. The first five faces were unfamiliar practice faces and were not analysed. The remaining faces were presented in a random order. SCRs were recorded throughout the presentation of the stimuli. The participants were encouraged to study each face and to remain as still as possible.

Following the SCR recordings, the famous faces were presented to the participant again. The subject was asked whether she knew each person and could give their names or any other information about them. The participant was deemed to have recognised the face if she gave the person’s correct name, gave the name of a character they play, or gave some unique description of them (e.g., “the current President of America”).

**The voice test.** Each voice was presented for between 5.5 seconds and 6.5 seconds, with an interval of 20 seconds between stimuli. The first three voices were practice items (unfamiliar stimuli) and were not analysed. The remaining voices were presented in a random order. SCRs were recorded throughout the presentation of the stimuli. The participants were encouraged to listen carefully to each voice and to remain as still as possible.

Following the SCR recordings, the famous voices were presented to the participant again to gauge their recognition of each item. The same criteria for recognition was used for this task as was employed for the face task.

**Results**

Recognition performance for faces and voices was recorded for two reasons. First, this informs us as to whether H.L. has a serious impairment in either face or voice recognition. Second, the positive recognitions can be selectively
employed to determine which stimuli were recognised during the measurement of SCR. In this way, therefore, it is possible to distinguish between famous but unrecognised stimuli and famous recognised stimuli.

The subjects’ SCR scores were converted into a within-subject standardised ratio based on the established method for comparing the size of differences between subjects (Lykken, 1972; Lykken & Venables, 1971). Such standardised scores are also the scores that have been employed in comparative studies into visual-Capgras delusion (Ellis et al., 1997b, 2000). A ratio was calculated for each subject for her SCRs to faces and a second ratio for SCRs to voices. In this way, it was possible interpret and compare the SCR differential between recognised and unfamiliar stimuli both between the domains (faces and voices) and between the subjects.

The face test. H.L. recognised 14 out of the 15 famous faces. This was more than the controls (mean number recognised = 12.8). She also showed a significant differential, $t(50) = 3.5; p < .01$, in her SCRs for recognised faces and nonfamous faces. The data from each subject were transformed by dividing by the average SCR for recognised items. These standardised data are shown in Figure 2 represented as the ratio of SCRs to recognised and unfamiliar items. The 95% confidence interval for this ratio found from the control subjects was between .166 and .411. H.L.’s SCR ratio occurred within this range (ratio = .391) and so her performance was within the normal range.

The voice test. H.L. recognised only 5 of the 15 famous voices. This compares to between 7 and 13 by the control participants. She did not show a significant differential, $t(27) = 0.527; p > .05$, in her SCRs between the recognised voices and the nonfamous voices. H.L.’s and the controls’ data were standardised in the same manner as described for faces to produce a ratio of SCRs for recognised and unfamiliar voices. These ratios are shown in Figure 2. The 95% confidence interval for this ratio, based on the control subjects, was between .156 and .352. The SCR ratio for H.L. fell outside this region (ratio = 0.804), and so her SCRs to voices cannot be considered to be within the normal range.\(^1\)

Discussion

The SCRs elicited from H.L. indicate that she has normal autonomic responses to faces but reduced differentiation to voices. This pattern of results is what we would expect to find from someone suffering from a voice-specific form of the Capgras delusion. H.L. also showed poor performance in recognising the

---

\(^1\) The voice SCR analysis was also conducted using just the five voices that H.L. recognised. This analysis was conducted in order to determine whether H.L.’s reduced SCR differential was due to recognising a reduced number of faces. This analysis still showed that the ratio of H.L.’s SCRs to the five recognised voices and the unfamiliar voices was significantly less than the control subjects data when calculated on the same items.
identity of voices although she was able to recognise one-third of those she heard. This is consistent with the evidence of visual-Capgras patients who have been found to show significant difficulty in recognising faces (Ellis et al., 1997b). Further, this evidence is consistent with her own report that in the most recent episode of the delusion it was her son’s voice (locked outside her flat at the time) that made her believe he was not who he claimed to be.

One possible weakness in this argument is the fact that H.L. may have generalised deficits in speech perception somehow that produce apparent specific deficits for one stimulus versus another, when the precise parameters of a stimulus (e.g., a person’s voice) are unknown. We can counter this, however, by pointing out that it is equally possible that a similar account of a generalised visual deficit leads to the observed data for the standard visual Capgras delusion (Ellis et al., 1997b). H.L.’s voice recognition skills were reduced but not completely eliminated and so can be seen to be analogous to the reduction in face identification seen with visual Capgras patients.

Despite this possibility, there is an accumulation of evidence that supports the idea that H.L. suffers from a voice-specific form of the Capgras delusion. In this, we believe, she is unique among the patients who have been reportedly tested for autonomic responses—all previous (nonsensory impaired) Capgras patients have shown reduced differentiation in their SCRs to faces whereas H.L. performs normally on this test.

**Figure 2.** This graph shows the values of the average SCRs for unfamiliar items as a ratio of the average SCR for recognised items. A ratio of 1, therefore, would represent no difference between recognised and unfamiliar items. Ratios lower than 1 indicate higher SCRs for recognised than unfamiliar items. Data are split according to the seven subjects (H.L. and the six controls) and is also split by the two tasks (Face and Voice). All ratios are considerably lower than 1 except for H.L. on the voice task.
The study of H.L.’s autonomic responses to voices and faces is unique in another way. This study, to our knowledge, is the first to demonstrate a possible dissociation between differential SCRs for faces and differential SCRs for voices. This dissociation has implications for the model of face and voice recognition described in Figure 1. Earlier, we described how there are two possible locations for the bifurcation of the identity pathway into the affective pathway. Bifurcation takes place either: at the FRUs and so, also, at the VRUs as shown in Figure 1; or at the PINs. In the latter case, recognition of a person by either voice or face will lead to similar SCRs. Alternatively, if there is bifurcation prior to the multimodal PINs then the route to the affective module and SCRs can be specifically impaired for voices or faces. H.L. shows normal autonomic responses for faces but abnormal responses for voices and so these results offer some support for the second model: the bifurcation of the identity and affective pathways occurs prior to the multimodal PINs.

The uniqueness of H.L. as a case means, however, that we should be careful of how much we infer from her performance. Our discussion is predicated on the belief that she is a genuine case of voice-specific Capgras delusion. As with any single case study, care must be taken not to infer too much. It is possible that the symptoms she reports and the pattern of SCRs she displays are a result of different dysfunctions. The fact that her data fit the predicted pattern for a voice-specific Capgras patient is reassuring but not necessarily conclusive. With this caveat in mind, H.L. has proved to be potentially highly illuminating for how we understand face and voice recognition and the psychiatric dysfunctions of these processes.

Manuscript received 24 July 2000
Revised manuscript received 16 November 2000

REFERENCES


During the course of this interview, M.L. attempted to find out what H.L. could remember of her most recent Capgras-like episode. This episode was described in her patient notes written soon after the episode. These notes describe an incident where she locked her son out of her flat for 11 hours stating that he was an impostor. It became apparent during the interview that it was possibly her son’s voice that was being mis-recognised rather than his face. This is particularly likely, as he would have been outside the flat where he could not be seen by H.L. except through a spy-hole. This possibility was followed up in later questions while still looking for some visual aspect to the delusion.

This kind of interview can be problematic because the patient may rationalise her behaviour and also it is possible that the questions may lead the patient to say what she thinks she should be saying. Nevertheless, the spontaneous references to her son’s voice and other details of the answers she gave are useful in concluding that the specific episode may have been more due to voice-specific mis-recognition than visual mis-recognition.

M.L. I would like to ask you about the time when you locked T., your son, out of your flat.
H.L. It’s because of the voice I heard. My eyes were blurred at that moment and I was wondering when he went out, whether he was the same person as when he went in.
M.L. Why did you think he wasn’t the same person?
H.L. He didn’t come across as chirpy ... or the person he was.
M.L. Did he look any different?
H.L. There was a mist.

... 
M.L. So when you saw your son, you thought that he was not your son. Is that correct?
H.L. It was the voice as well. Like he wanted to say to me ‘‘Mum, I have had a long day’’. He was in an angry mood as well. He wanted to say to me ‘‘Mum, I’ve had enough of the day’’.
M.L. You thought that he wasn’t the real T?
H.L. mm.
M.L. Was that because of the way he looked or because of the way he sounded?
H.L. A little bit of his voice, and more of his being.
M.L. He sounded different?
H.L. Yes, a little bit.

... 
M.L. I would just like to ask you more about the time you locked T out of the flat. What was it that made you think that that person wasn’t T.
H.L. More the voice. . . . His teeth, his teeth at the bottom used to be a bit wrong, he had them fixed. I kept asking him ‘‘Are you real? Are you real? . . . have you had your teeth done?’’ And he would say ‘‘Mum, don’t ask me those questions’’.

... 
M.L. Do you have any trouble recognising people or think that someone isn’t who they are supposed to be?
H.L. Sometimes when I sit watching TV, even well-known faces, the blurred vision, when I have sat down too long.
M.L. Do they ever sound strange as well as look strange?
H.L. Yes, probably that time you are a bit mixed up in your head. Unexpected of the usual voice, doesn’t sound right.
M.L. And you think that that is not that person because they are sounding strange.
H.L. Yes, you do.

... 
M.L. Do you only have trouble recognising people when this mistiness comes over your eyes?
H.L. Yes, when I have had enough. Especially when my neighbours play their music, you know, you bang your head against the wall—not that I’ve done that. You are complaining about it and nobody tries to help you. It destroys me. It is me that goes into hospital.