Chapter 13

Hallucinations in schizophrenia: theory and findings

Summary
Hallucinations are a mysterious psychological phenomenon. In this paper, current theory and research on the neurocognitive basis of hallucination are reviewed. The phenomenology of hallucinations in schizophrenia is discussed, followed by an overview of functional brain imaging studies of cortical areas involved in hallucination. Theories addressing the putative cognitive mechanisms underlying hallucination are critically reviewed. Hypotheses concerning the role of inner speech, speech perception, reality monitoring, mental imagery, verbal self-monitoring and top-down perceptual expectations are discussed in light of the published evidence. Proposals for integration of various cognitive models are described and future directions for research are outlined.

Hallucinations are an intriguing psychological phenomenon. The brain perceives something: a sound, a voice, an image. However, there is no corresponding source in the world outside.

Where do hallucinations come from? In non-western cultures, the answer usually is: from gods and ghosts (Al-Issa, 1995). This is also the case for ancient Greece literature, such as the Illias. In our culture, hallucinations are mostly associated with the use of stimulants, and with medical and psychiatric conditions. A hallucination can be defined as a perceptual experience in the absence of sensory stimulation. In order to distinguish this from mental imagery and dreaming (cf. Slade & Bentall, 1988; Aleman & De Haan, 1998), it is instructive to add to this definition that hallucinations are not under voluntary control of the individual (contrary to mental imagery), and occur in a wakeful state (contrary to dreaming). Hallucinations may occur in a wide range of circumstances. For example, Brasic (1998) lists more than 40 medical and psychiatric conditions in which hallucinations may occur. This paper will be restricted to hallucinations in schizophrenia, on which cognitive research has concentrated over the past few decades (David, 1999). After a description of the phenomenology of hallucinations, neuroimaging studies of patterns of cerebral activation associated with hallucination will be reviewed, and cognitive theories of hallucination will be discussed.

**Phenomenology**

Hallucinations may occur in any sensory modality: auditory, visual, somatosensory, gustatory and olfactory. In schizophrenia, auditory hallucinations are by far most frequent, 65% of patients with schizophrenia has suffered at least once from auditory hallucinations (Slade & Bentall, 1988). Visual hallucinations are less frequent, some 20% of patients. Less than 5% of patients reports hallucinations in the other modalities.

Auditory hallucinations may differ considerably in their phenomenology. They may consist of simple sounds, such as “tapping on the scalp” or ringing of deathbells. In other cases, music is heard. Predominant, however, are verbal hallucinations or “hearing voices”. A well-known classification of these are the hallucinations designated “first-rank” symptoms of schizophrenia by Kurt Schneider (Schneider, 1962). He distinguished three types of hallucination: 1) the patient hears ongoing commentary on his behavior, 2) the patient hears voices talking about him in the third person, and 3) the patient hears his own thoughts spoken aloud.
There is significant inter-individual variability between patients in formal characteristics of hallucinations (Junginger & Frame, 1985). This concerns the frequency (which varies from almost never to continuously), location (inside or outside the head), clearness (which varies from unclear and not understandable to very clear) and loudness (which varies from very soft, almost inaudible, to very loud, screaming). Despite this variability, a study conducted with 54 hallucinating patients (Junginger & Frame, 1985) revealed that a majority of patients can understand the voices clearly, with a volume comparable to normal conversation. With regard to the location, a small majority reported that this was “outside the head”. Nayani & David (1996) replicated these results in a study with 100 patients.

In a large number of cases, the “voices” heard by patients with schizophrenia are experienced as hostile. For example, the voices may continuously criticize the patient’s behavior, or command the patient to behave against his will. Nayani & David (1996) observed that the most frequent expressions in auditory-verbal hallucinations concerned abusive language. Such hallucinations are a stressful experience which may severely impair the patient’s ability to function normally in daily life.

Visual hallucinations may also tend to be of bizarre content. An example is the case described by Silbersweig et al. (1995). This patient saw moving, colorful scenes, with rolling heads without body, and heard these heads speak to him, giving him instructions.

Are hallucinations pathologic?

In the absence of a direct cause, such as use of stimulants, or a medical condition such as a brain-tumor, hallucinations are usually taken to imply some form of mental illness. However, a number of studies has demonstrated that a substantial part of individuals from the normal population (varying from 5-25%) reports hallucinatory experiences (Aleman et al. 2001; Morrison et al. 1999; Young et al. 1986; Barrett & Etheridge, 1992). For example, a British study among 203 college students reported that 13% of the respondents answered “certainly applies to me” to the item “In the past I have had the experience of hearing a person’s voice and then found that no one was there” (Young et al. 1986). Such hallucinatory events have not only been reported by college students but have been corroborated by large epidemiological studies (Tien, 1991). On the basis of these studies, it can be concluded that hallucinatory experiences form a continuum with normal psychological processes (cf. Slade & Bentall,
An important distinction with hallucinations in schizophrenia is that, in individuals from the normal population, hallucinatory events rarely are experienced as unpleasant, emotional threatening or hostile (Barrett & Caylor, 1998). In much cases a person only hears his or her name, while no one is around.

**Brain activity during hallucinations**

Which brain areas are involved in experiencing a hallucination? Researchers have tried to answer this question with the use of modern functional neuroimaging techniques, such as PET and fMRI (for a recent review, see Weiss & Heckers, 1999). In the first study (McGuire et al. 1993), 13 patients were scanned with PET in an episode of their illness in which they experienced hallucinations. They were scanned again on a second occasion, when the hallucinations were absent. Compared to the second measurement, hallucination-related activity was observed in language-related areas, especially Broca's area (involved in speech production). Although to a lesser extent, activity was also found in the anterior cingulate (involved in attentional processes), and in the left temporal cortex (a.o. auditory perception and memory processes). In a comparable design, Suzucki et al. (1993) observed an increase in regional bloodflow in the left temporal lobe (auditory association cortex) in five hallucinating patients. Silbersweig et al. (1995) reported activation of subcortical structures, the parahippocampal gyrus and the middle temporal gyrus in five patients during auditory hallucinations. One of their patients also hallucinated in the visual modality. For these hallucinations, activation was observed in visual areas (lingual, fusiform and occipital gyri).

Lennox et al. (1999) imaged a hallucinating patient with fMRI. This patient hallucinated with consequent intervals: approximately for 26 seconds he heard a "voice", followed by a comparable period in which hallucinations were absent. The patient indicated with a key press when he heard the "voice". This was an ideal condition for a controlled fMRI study, in which a within-subject comparison could be made between hallucinatory periods and hallucination-free periods. The results revealed strong activity in the right middle temporal gyrus. In the same way, using the "button-pressing method", Dierks et al. (1999) managed to scan three patients with fMRI. They observed activity in Broca's area, in the temporal gyri, and in the primary auditory cortex (Heschls gyrus). Studies by David et al. (1996) and Woodruff et al. (1997) also indicate that primary sensory areas may be involved in the experience of hallucination. These
investigators found that the primary auditory cortex is less responsive to auditory stimuli during hallucinations compared to absence of hallucinations. This is an indication that the primary auditory cortex is actively involved in hallucinations, and therefore no “resources” are left for additional processing of auditory stimuli. For visual hallucinations, an identical finding has been reported (Howard et al., 1995). In the most recent study, Shergill et al. (2000) used a novel fMRI method to measure brain activity during hallucinations in 6 patients. In this “random sampling” method, a large number of individual scans is acquired at unpredictable intervals in each subject while they are intermittently hallucinating. Immediately after each scan, subjects report whether they had been hallucinating at that instant. Neural activity is then compared for the scans when patients were and were not experiencing hallucinations. The results revealed a distributed network of cortical and subcortical activity associated with auditory hallucinations: inferior frontal/insula, temporal cortex bilaterally, right thalamus and inferior colliculus, and left hippocampus and parahippocampal cortex.

To summarize, neuroimaging studies reveal a distributed network of cortical and subcortical areas involved in the experience of hallucinations. Although the exact role of these areas is not clear yet, it could be hypothesized that hallucinations are triggered by activity in subcortical and frontal areas, which in turn project to modality-specific association cortex, thereby leading to a conscious perceptual experience. With respect to auditory hallucinations, some studies observe activity in language-production areas during auditory hallucinations, some studies observe activity in the primary auditory cortex, but all studies report activity in the temporal lobe, more specifically in the middle or superior gyri. For visual hallucinations, activity is observed in secondary visual cortex.

Cognitive theories
Four approaches can be distinguished in recent cognitive theories regarding the mechanism of hallucination. These approaches focus respectively on 1) “inner speech”, 2) speech perception, 3) reality discrimination, and 4) mental imagery. In contrast to the latter, the first two approaches are confined to auditory-verbal hallucinations.
Inner speech

Most individuals report the experience of “inner speech” (either occasionally or continuously) when they think. Some hallucinating patients indicate that they cannot distinguish well between their inner speech and the “voices” they hear. In addition, subvocal muscle activity has been reported, associated with hallucinations (Green & Kinsbourne, 1990). The “inner-speech” hypothesis of hallucinations holds that some distortion in the production of inner speech leads to the erroneous interpretation that the “inner speech” is of non-self origin. A cognitive neuropsychological study by David & Lucas (1993) was not able to confirm the inner speech hypothesis. On the basis of Baddeley's working memory model (Baddeley & Hitch, 1974; Baddeley, 1986) these authors argued that inner speech is mediated by phonological processes in short-term memory (the “phonological loop” in Baddeley's terms). If hallucinations and inner speech both call on resources in the phonological loop, a dual task will disrupt either process, given the limited nature of processing resources in working memory. David & Lucas (1993) went on to demonstrate that phonological processing during auditory-verbal hallucinations was not affected in a continuously hallucinating patient. The authors suggest that their finding implies that inner speech and auditory-verbal hallucinations are different processes. Moreover, in contrast to what the inner speech hypothesis would predict, Haddock et al. (1996) did not find specific impairments in phonological processing which could underly distortions in inner speech in hallucinating patients as compared to non-hallucinating patients.

McGuire et al. (1993) and Dierks et al. (1999) reported activity of Broca's area during hallucinations, which may be consistent with the inner speech theory. However, other PET and fMRI studies failed to find Broca area involvement (Silbersweig et al., 1995; Lennox et al., 1999).

Speech perception

According to Hoffman (Hoffman et al., 1999) a dysfunction of the speech perception system underlies auditory-verbal hallucinations. In the analysis of every-day sound characteristics, there is an important degree of acoustic ambiguity, due to background noise, and due to the “pasting” of phonemes (also called “blurring”). Syntactical and semantical expectations, based on earlier learnt words, therefore play a crucial role in speech perception. Hoffman’s hypothesis is that hallucinations arise from an impairment in verbal working-memory,
which leads to pronounced linguistic expectations that could generate spontaneous perceptual “outputs”.

Evidence for this hypothesis comes from a study in which hallucinating and non-hallucinating patients were compared on a speech perception task in which the presented speech was so distorted that it was difficult to recognize (Hoffman et al., 1999). As predicted, hallucinating patients performed significantly worse than their non-hallucinating comparison patients. On a measure of verbal working-memory (sentence repetition), the hallucinating group also performed worse, but not on a measure of sustained attention, indicating that the performance differences could hardly be ascribed to attentional deficits.

However, on the basis of his theory Hoffman (1986) predicted that hallucinating patients would have more difficulties in the production and processing of speech than non-hallucinating patients, but subsequent research has failed to confirm this prediction (Slade & Bentall, 1988). Nevertheless, PET and fMRI studies of brain activity during hallucinations are consistent with the speech-perception hypothesis: all studies report activation in temporal auditory-linguistic association areas.

**Reality monitoring**

“Source monitoring” refers to the ability to distinguish between different sources of information, e.g., whether something was read in a newspaper, or whether it was told by a friend (Johnson, Hashtroudi & Lindsay, 1993). Reality discrimination and reality monitoring (Johnson & Raye, 1981) are considered to belong to this category of processes. **Reality discrimination** refers to distinguishing between internally generated information and externally presented information (e.g., imagination and perception), whereas **reality monitoring** refers to memories of whether information was or internal or external origin (e.g., did I imagine it, or did it really occur?). Thus, reality discrimination refers to the “online” distinguishing of external versus internal sources, whereas reality monitoring refers to information that was presented or generated in the past. Reality discrimination measured with a signal detection task was reported by Bentall & Slade (1985), who found that hallucinating patients made significantly more errors than non-hallucinating patients (specifically, the hallucinating patients erroneously indicated that a word had been presented in a burst of white noise). An example of a reality monitoring task is a memory task in which the subject is asked to remember words that have either been said by the experimenter or have
been generated by the subject himself (after indications by the experimenter). Subsequently, the subject is asked to indicate, from a list of words, whether a word was a) previously read by the experimenter, b) generated by the subject himself, or c) whether the word is new. According to the reality monitoring hypothesis (Bentall, 1990), hallucinating patients will more frequently erroneously assign self-generated words to an external source (by indicating that the word was presented by the experimenter). A number of behavioral studies has provided evidence for such a relation between reality monitoring errors and the occurrence of hallucinations (Bentall, Baker & Havers, 1991; Morisson & Haddock, 1997; Brébion et al., 2000). However, the question remains as to how specific disorders in reality monitoring are to hallucinations. Keefe et al. (1999), for example report that patients without hallucinations (but with other positive symptoms) made the same errors as hallucinating patients.

According to Johnson, Hashtroudi & Lindsay (1993), areas in the frontal lobe are crucially involved in reality monitoring. Only a few neuroimaging studies have revealed frontal activity associated with hallucinations, however (McGuiere et al., 1993; en Dierks et al. 1999).

Mental imagery
In the 19th century, Francis Galton wrote that mental imagery exists as a continuum in the population, ranging from a total absence of mental images (subjectively) to imagery of great intensity and vividness, ending in pure hallucination (Galton, 1883). A number of studies investigated the imagery hypothesis (e.g., Roman & Landis, 1945; Mintz & Alpert, 1972; Starker & Jolin, 1982), but with inconsistent results. The fact that none of the studies included adequate behavioral measures may account for this inconsistency. Indeed, Slade & Bentall (1988) have drawn attention to the fact that explaining hallucinatory experiences with a phenomenologically highly similar event – subjectively rated imagery vividness – borders to circularity.

However, it is not easy to think of a method to measure vividness of mental imagery behaviorally. A possible approach could be the one first described by Aleman et al. (1999; cf. Aleman et al., 2000), in which performance is compared on a perception and on an imagery condition of the same behavioral task. According to Johnson and Raye (1981) percepts, which originate from externally presented stimuli, are characterized by more detailed sensory, contextual and semantic information than internally generated images. Evidence that mental images are less rich in perceptual details than ‘real’ percepts and that,
as a consequence, images are more difficult to perform mental operations upon, was recently presented by Kosslyn et al. (1999). The hypothesis that imagery and perception are more alike (and therefore harder to discern from each other) due to increased sensory characteristics of mental images in individuals that experience hallucinations thus predicts that these subjects will show smaller performance differences between a perception and an imagery condition of the same task.

Using this method, Böcker et al. (2000) compared hallucinating and non-hallucinating patients on two measures of auditory and two of visual imagery and perception. No differences were found between both groups when performance on the imagery measures relative to perception performance was compared. However, after performing within-group comparisons, the authors observed more vivid auditory than visual imagery in patients that hallucinated in the auditory modality. Evans et al. (2000) also reported a lack of differences between hallucinating and non-hallucinating patients with schizophrenia on a number of auditory imagery measures. However, these authors did not include perception conditions, nor measures in another non-hallucination modality.

It is interesting to note that it has also been argued that hallucinating patients may suffer from an imagery deficit, rather than a general increase in vividness. For example, Horowitz (1975) hypothesized that hallucinating patients have less vivid mental images, which leads them to attribute occasional vivid images to an external source. However, in both instances of imagery theory, a vivid mental image ultimately gives rise to the hallucinatory experience.

Neuroimaging studies are consistent with activation that would be predicted by the imagery hypothesis: both auditory hallucinations and auditory imagery appear to activate auditory association areas (Dierks et al., 1999; Zatorre et al., 1996). The same holds for visual hallucinations and visual imagery (ffytche et al. 1999; Kosslyn et al. 1999).

**Integrating the various perspectives**

Despite the differences between these four cognitive approaches, there is also some conceptual overlap, which makes the possibility of integration especially attractive. Indeed, it could be argued that two earlier theories, namely the proposals of Frith (1992) and of Grossberg (1999) incorporate elements of more than one approach.
Frith’s theory (Frith & Done, 1988; Frith, 1992) can be seen as an integration of the “inner speech” hypothesis and the reality monitoring hypothesis. According to Frith, hallucinations arise from failures in the monitoring of own intentions during inner speech (sometimes called ‘self-monitoring’ by Frith). As a consequence, the cognitive system does not recognize that inner speech originates from the self, and thus erroneously attributes it to a non-self source. Thus, this approach does not consider the production of inner speech to be impaired, but rather states that auditory hallucinations are derived from defective monitoring of inner speech. Evidence for this hypothesis was recently presented by Johns & McGuire (1999). Hallucinating patients, non-hallucinating patients and normal control subjects were asked to speak presented words out aloud in a microphone. Occasionally, the spoken word was distorted by the experimenter (by modulating the pitch). Participants heard the words in their headphones and were asked to indicate if the source of the heard word: “myself”, “somebody else”, or “unsure”. The hallucinating group made significantly more errors by attributing own (distorted) speech to someone else. Indeed, this study was inspired by the “speech-monitoring” approach, but is clearly also consistent with the speech perception hypothesis of Hoffman (1999). A problem for the verbal self-monitoring theory concern the results reported by Leudar, Thomas & Johnston (1994). These authors investigated whether schizophrenic patients have deficient internal error detection in speech repairs (especially when these occur rapidly, before external acoustic feedback can have come into play). Although patients with schizophrenia showed less internal error detection than controls, consistent with a failure of verbal self-monitoring, there was no difference between patients with and without hallucinations.

In accordance with Frith’s speech-monitoring hypothesis, McGuire et al. (1995) found reduced temporal activation during verbal self-monitoring tasks in hallucinating patients. Other neuroimaging studies have implied this region in self-monitoring in healthy subjects (e.g., McGuire et al., 1996). Shergill et al. (2000) also argue that their findings are consistent with Frith’s verbal self-monitoring theory. However, they relate this to the attenuated activation of the supplementary motor area (SMA) during hallucinations. As the SMA has been implicated in the deliberate generation of inner speech, and lesions in this region have been associated with the alien limb syndrome (in which a patient attributes self-generated movements to someone else), the paucity of SMA activation
during hallucinations might be related to a lack of awareness that inner speech has been generated.

A different approach to hallucinations has been described by Grossberg (2000), based on the finding that top-down perceptual expectations can importantly affect the detection of stimuli (Coren, Wards & Enns, 1994). Indeed, such top-down mechanisms play an important role in perception, by modulating, priming and matching incoming bottom-up information (cf. Kosslyn, 1994). Thus, for example, an expected stimulus will be detected better than an unexpected stimulus. The neurophysiological basis of this effect has been well studied (Grossberg 1999), and includes feedback circuits in which a balance is reached between top-down excitation and inhibition (e.g. by feedback into information flow in visual cortex layers via an on-center, off-surround pathway). However, although such top-down expectations can modulate, sensitize, or prime the processing of bottom-up information, they cannot by themselves cause supra-threshold activation of their target cells. Nevertheless, as Grossberg (2000) recently hypothesized, under normal behavioral conditions, a volitional signal can be phasically turned on that can alter this balance to favor top-down excitation, which can create conscious experiences in the absence of bottom-up information. In this way, conscious mental imagery can arise. In addition, Grossberg (2000) proposes a mechanism by which hallucinations in schizophrenia could arise, namely when the phasic volitional signal becomes chronically hyperactive. As a result, top-down sensory volitional signal can be phasically turned on that can alter this balance to favor top-down excitation, which can create conscious experiences in the absence of bottom-up information. In this way, conscious mental imagery can arise. Further details on the possible neurophysiological mechanisms can be found in Grossberg (2000). This theory integrates elements of the imagery hypothesis (which bears on strong top-down processes) and is reminiscent of Hoffman’s statement that “pronounced linguistic expectations can generate perceptual outputs”. Consistent with the perceptual expectations hypothesis are the findings reported by Haddock, Slade & Bentall (1995). They suggested to subjects that on listening to a word repeated over and over (e.g., the word “tress”) they would hear new words. Indeed, the subjects reported hearing more transformations (e.g., stress, dress), but the subjects with hallucinations in addition reported hearing other words (e.g., caressed, Christmas).

An approach similar to the one by Grossberg (2000) has been described by Behrendt (1998), although he is less explicit in proposing a neural mechanism. Behrendt states that hallucinations could arise from “facilitated
formation of cortical associations between representations of expectations and internal symbols” (p. 236). In this way, “mental factors of perception”, rather than sensory information, take over to predominate the context and form of perception, which will in fact be a hallucination. Thus the normal balance between sensory and “mental” factors is distorted. (The “mental factors” of Behrendt refer to top-down mechanisms, whereas sensory factors refer to bottom-up mechanisms; for a more detailed description of theory and empirical evidence regarding top-down influences intrinsic in perception, see Kosslyn & Sussman [1995]).

Conclusion
Most cognitive theorists agree that hallucinations are misattributions of internally generated information to an external source. Different hypotheses have been developed, concerning the role of inner speech, speech perception, reality monitoring, and mental imagery. Probably, the most accurate summary of the current state of affairs was recently advanced by David (1999): “Auditory imagery – that is, a sensory component – is intuitively central to the experience of hallucinations, and recent fMRI studies support this. Either a distortion of the image itself (its prosody, pitch or timbre), its apparent coherence, or ego-alien content, or a defect in the self-monitoring (or a combination of all these) leads to a misattribution of the source. This mislabelling requires more precise cognitive dissection” (p. 101). One component that is not mentioned here by David is a possible perceptual deficit (although he discusses perceptual deficits earlier in his paper). McKay et al. (2000) have provided evidence for higher order perceptual deficits in hallucinating schizophrenic patients. In an attempt to further “cognitive dissection”, we will take on the issue of an imbalance between imagery and perception in the next chapter.
References


