PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Opinion piece

Cite this article: Thakkar KN, Mathalon DH, Ford JM. 2021 Reconciling competing mechanisms posited to underlie auditory verbal hallucinations. *Phil. Trans. R. Soc. B* **376**: 20190702.
http://dx.doi.org/10.1098/rstb.2019.0702

Accepted: 24 August 2020

One contribution of 16 to a theme issue 'Offline perception: voluntary and spontaneous perceptual experiences without matching external stimulation'.

Subject Areas:
neuroscience, cognition, behaviour

Keywords:
hallucinations, psychosis, perception, imagination, agency

Author for correspondence:
Judith M. Ford
e-mail: judith.ford@ucsf.edu

THE ROYAL SOCIETY PUBLISHING

Reconciling competing mechanisms posited to underlie auditory verbal hallucinations

Katharine N. Thakkar[1,2], Daniel H. Mathalon[3,4] and Judith M. Ford[3,4]

[1]Department of Psychology, and [2]Department of Psychiatry and Behavioral Medicine, Michigan State University, East Lansing, MI, USA
[3]Department of Psychiatry and Behavioral Sciences, University of California, San Francisco (UCSF), San Francisco, CA, USA
[4]Mental Health Service, Veterans Affairs San Francisco Healthcare System, San Francisco, CA, USA

JMF, 0000-0002-6500-6548

Perception is not the passive registration of incoming sensory data. Rather, it involves some analysis by synthesis, based on past experiences and context. One adaptive consequence of this arrangement is imagination—the ability to richly simulate sensory experiences, interrogate and manipulate those simulations, in service of action and decision making. In this paper, we will discuss one possible cost of this adaptation, namely hallucinations—perceptions without sensory stimulation, which characterize serious mental illnesses like schizophrenia, but which also occur in neurological illnesses, and—crucially for the present piece—are common also in the non-treatment-seeking population. We will draw upon a framework for imagination that distinguishes voluntary from non-voluntary experiences and explore the extent to which the varieties and features of hallucinations map onto this distinction, with a focus on auditory-verbal hallucinations (AVHs)—colloquially, hearing voices. We will propose that sense of agency for the act of imagining is key to meaningfully dissecting different forms and features of AVHs, and we will outline the neural, cognitive and phenomenological sequelae of this sense. We will conclude that a compelling unifying framework for action, perception and belief—predictive processing—can incorporate observations regarding sense of agency, imagination and hallucination.

This article is part of the theme issue 'Offline perception: voluntary and spontaneous perceptual experiences without matching external stimulation'.

1. Introduction

Auditory-verbal hallucinations (AVHs), also referred to as 'voices,' are vivid perceptions of speech that occur in the absence of any corresponding external stimulus but seem very real to the voice hearer. As Jones [1, p. 566] so aptly summarized the experience, voices '…encapsulate a diverse phenomenological experience, which may involve single and/or multiple voices, who may be known and/or unknown, speaking sequentially and/or simultaneously, in the first, second and/or third person, and which may give commands, comments, insults or encouragement'. That these voices can be experienced as indistinguishable from externally generated percepts and accordingly generate a range of responses and affective reactions—from distress to comfort—has piqued the interest of clinicians, philosophers and cognitive scientists alike. Experiences like AVHs are perhaps the clearest evidence for the inextricable influences of internal contents of the mind on the construction of conscious perception [2].

AVHs are experienced by the majority of people with schizophrenia spectrum disorders. These disorders are characterized most notably by psychosis (i.e. false perceptions and beliefs; disorganized thought and language). They are accompanied by significant distress, functional impairments and decreased

quality of life. In recent years, it has become increasingly apparent that voice hearing seems to be experienced across the wellness spectrum, with a small but significant percentage of the general population hearing voices [3] without seeking treatment or receiving a diagnosis of a psychotic disorder. We refer to these individuals as non-treatment-seeking voice hearers and argue, as others have [4], that understanding the distinctions and overlaps between these groups can provide us significant leverage on understanding the nature and mechanisms of AVHs in populations most troubled by them. We acknowledge the limitations of this approach—namely, that these are not monolithic categories. Naturally, there is variability within these groups and the movement between them. However, based on the distinctions described below, we nevertheless argue that such categorization forms the basis of a fruitful line of inquiry.

Both treatment and non-treatment-seeking voice hearers experience voices occurring spontaneously. However, compared to the treatment-seeking group, non-treatment-seeking voice hearers often attribute voices to a divine being and consider them protective and less bothersome [5]. Importantly, non-treatment-seeking voice hearers feel a sense of agency over their voices that permits them some control over their onset and offset [6]. This may help them to frame the experiences in a positive light [5] and protect them from the self-destructive negative, ruminative thoughts that could render the voices hostile and threatening, and send them to the clinic seeking help [7].

Indeed, dialogic interaction, which can foster a sense of control over one's voices [8], is a route to recovery that may be facilitated by the virtual reality-based Avatar therapy through which the voice is embodied, interacted with and ultimately disregarded [9]. This, and the experience of some non-treatment-seeking voice hearers who learn to channel their voices through practices that regulate their interoceptive awareness [10–12], points toward the importance of learning and volition in the experience of hallucinations, to which we will return later.

How then can we begin to get some theoretical and empirical leverage on AVHs? It would seem that the constructive nature of perception that engenders experienced reality, as well as agency over these perceptual inferences, might be fruitful avenues of inquiry. Indeed, they form the basis of the two most prominent mechanistic models of AVHs.

2. Aetiological models

Several aetiological models of AVHs have been put forward. In the following section, we review the two most prominent models, which on the surface appear to be in direct conflict. We highlight recent ideas from Corlett and colleagues that challenge that notion [13,14], instead considering how these models might be reconciled and integrated within one unified framework.

(a) Altered efference copies

One influential theory of AVHs posits that they represent a failure in self-monitoring, such that voices inside the head (i.e. verbal thoughts and/or subvocal speech) are not identified as self-generated and are thus erroneously misattributed to an outside source and subjectively experienced accordingly [15]. This self-monitoring deficit is argued to arise owing to compromised efference copy signals—'copies' of motor signals that are sent to sensory processing regions [16]. This efference copy deposits a prediction, or 'corollary discharge,' of the expected sensory consequences of the action in sensory brain regions. Importantly, these 'actions' can take the form of thinking, which is considered by some to be our most complex motor act [17]. These efference copies act to attenuate self-generated sensations and may provide a self-tag, or a substrate for agency attribution: 'I infer that I am the author of actions that resulted in this sensation.' Sufficient deviation from the predicted sensory consequences of actions invites the inference that another agent was involved. These signals have been studied in song birds, non-human primates, and in humans during active vocalization and passive listening to recorded vocalizations [18–23]. Auditory cortical responses are reduced during active vocalization compared to passive listening, and this is considered evidence of the suppressive action of an efference copy/corollary discharge mechanism [18–23]. In humans, electroencephalogram (EEG)-based event-related potentials have been used to assay this mechanism [24], using an early negative (N100) component of the event-related potential that is generated in the auditory cortex [25] and is suppressed during vocalization compared to passive listening [20–22,26].

The self-monitoring theory of AVH posits that without effective feed-forward efference copy signals modulating sensory cortex, no sensory suppression occurs, and one's thoughts (couched in the language of our inner speech) are not experienced as self-generated, but rather, are attributed to an external entity. While appealing and intuitive, there are few reports of strong relationships between AVHs and the N100 assay of corollary discharge. One exception here is the finding that treatment-seeking individuals who experienced AVHs failed to show the loss of N100 suppression that occurs in healthy individuals and non-hallucinating patients when their vocalizations were pitch-shifted in real time [23]. In other words, hallucinating patients appeared to have a less precise corollary discharge and thus failed to show the typically enhanced auditory cortical response to prediction-violating pitch-shifted vocalizations. One study of young people with psychosis risk syndrome found that N100 suppression was related to odd beliefs, akin to delusions in psychosis [27].

While suppression of the *amplitude* N100 during vocalization has not proved to be very sensitive to AVH, analysing the N100 data in the time-frequency domain has been more fruitful. N100 amplitude is an admixture of power and phase resetting in the theta band (4–7 Hz), and it can be decomposed into those elements. Failure to suppress phase resetting has proved to be more sensitive to psychotic symptoms, specifically delusions, and more sensitive to the diagnosis of schizophrenia, than is failure to suppress N100 amplitude [28]. Using similar analytic tools, Ford *et al.* [29] reported a relationship between AVH severity and phase consistency preceding speech onset. Because it was highly correlated with subsequent suppression of N100 amplitude, this precisely timed signal preceding speech was posited to reflect the efference copy being sent from speech production areas, heralding the arrival of the imminent sound. Furthermore, coherence between frontal and temporal lobes in the theta band, which was greater during speaking than listening, was reduced in people with schizophrenia, especially in those who heard voices [30]. These studies suggest that the additional precision afforded by time-frequency analysis of EEG signals during the

motor act of speaking can be used to study the pathophysiology of psychosis and psychotic symptoms.

Although not specifically related to AVH, it is important to note that evidence for alterations in efference copy signals in schizophrenia also comes from other sensorimotor domains [31–33], and may underwrite symptoms that can be tied back to a faulty sense of agency more generally.

We suggest that the agency deficit related to psychotic symptoms may be instantiated by a faulty efference copy being sent from cortex to cerebellum [34], via the pons. N100 suppression may reveal the comparison between the expected and observed sensations executed in the cerebellum, with the outcome being relayed through the thalamus to the auditory cortex. Important details about where in the cerebellum the efference copy might land, where the calculation might be made and where in the cerebellum the input to thalamus might be launched remain to be determined. As more is learned about the precise area of the cerebellum involved in the rapid processing of expected outcomes of actions, these regions might then be the target of focused neurostimulation to boost their function.

(b) Associative learning models and overly precise priors

Perception is an inferential process [35]. According to predictive coding accounts, the brain is cast as a prediction machine that is charged with making an inference about the cause of our current sensation [36], an inference that is optimized by prior knowledge about probable candidates [36]. Predictions of sensory input (i.e. top-down expectations) are sent to lower brain areas, and these predictions are compared to actual sensory input. Discrepancies in the form of prediction errors are carried forward, and this prediction error may be used to update the model of the world (i.e. new predictions are formed). Formalized within a Bayesian framework, sensory information (likelihood) is combined with predictions (priors), prediction errors are computed, and the most likely cause of sensory data (posterior) becomes the percept. Now, this prior is associated with a certain precision (strictly, the inverse variance of the distribution of possible values the data could take) and this precision can be thought of as reliability. If priors are more precise (i.e. reliable) than sensory inputs, they will dominate inference and prediction errors will be ignored [37–40]. Alternatively, relatively more precise prediction errors will dominate priors and drive belief updating (changing one's priors for subsequent inference). The powerful contribution of expectation to perception spurred the idea that hallucinations might arise when prior predictions exert an inordinate influence over perceptual inferences, creating percepts with no corresponding external stimuli at all [2,41].

Indeed, in healthy volunteers who have undergone a training period that establishes an association between two stimuli, perceptual experiences of one stimulus (i.e. a tone) can occur in the absence of sensory input, conditional on the presentation of another stimulus (i.e. a visual stimulus) [42], akin to a conditioned reflex [43,44]. More recently, visual-auditory conditioning has been employed to demonstrate that treatment-seeking voice hearers are significantly more susceptible to this effect than patients without hallucinations and controls [45]. Powers *et al*. [46] recently showed that these conditioned hallucinations are mediated by strong prior beliefs, that those priors are stronger in people who hallucinate (in a manner that correlates with hallucinations), and that people with a diagnosed psychotic illness are less likely to update those prior beliefs in light of new evidence. Critically, the neural circuit underlying these conditioned phenomena—including superior temporal gyrus and insula—largely overlapped with the circuit engaged when patients report hearing voices in the scanner [46,47]. These studies underline the role of learning and, more specifically, a bias towards learned top-down information in the genesis of AVHs. Further support for this so-called strong prior account of hallucinations comes from findings that prior knowledge of a visual scene confers an advantage in recognizing a degraded version of that image [48] and that patients at risk for psychosis—and, by extension, voice hearing—were particularly susceptible to this advantage, and its magnitude correlated with hallucination-like percepts. Similarly, there is a version of this effect in audition; voice-hearing participants appear to have an enhanced expectation for speech in degraded auditory stimuli even when not explicitly instructed [49]. That is, speech is perhaps the most salient biological signal for our species, and the auditory system of AVH-prone individuals may be predisposed to inferring speech. Indeed, tonotopic brain mapping suggests that this is true [50].

3. Reconciling strong priors and weak corollary discharges

On the face of it, the two aforementioned theories of AVHs are at direct odds. Corollary discharge accounts of AVHs posit a *reduced* influence of expectations on perception, where associative learning accounts posit an *enhanced* influence of expectations on perception. 'Expectations' in the corollary discharge account comprise the predicted low-level sensory experience on the basis of a set of planned movement dynamics and are established in the milliseconds preceding an action and are dynamically updated as one moves, speaks and, possibly, thinks [29]. In the associative learning account, 'expectations' can be more complex cognitive representations (i.e. higher-level beliefs) and are typically established as a result of experiences acquired over a longer timescale—a lifetime, even. Perhaps because they operate on such different timescales and at different levels of a putative information processing hierarchy one may imagine that these different 'expectations' play different, but perhaps, coordinated roles in the experience of AVHs.

Indeed, Corlett and colleagues have suggested that a reduced influence of corollary discharge signals on perception can engender a stronger reliance on expectations derived from past experiences [13,14]. Evidence supporting such a claim comes from Bayesian models of multi-sensory integration, which have established that the prioritization or weighting of streams of information during inference depends on their relative reliability or precision [51–54]. Audiovisual integration of sensory cues is critically dependent upon the relative precisions of the information to be integrated [53]—the more precise (less variable) source of information is weighted more highly. They build the following argument assuming similar integration of external and self-related sources of information in overall inference. When corollary discharge fails, the weighting of self-related information decreases. In doing so, it results in a shift towards reliance on prior expectations about the external world—by contrast to priors about the self, inviting

the inference that, while I am thinking, the sound is emanating from an agent outside of myself[1]. Thus, weak and strong prior effects can coexist; and failures of corollary discharge can, in theory, give rise to the strong prior expectation that a voice will be heard—as has been observed [49]. This proposed account reconciles evidence for strong priors with impaired corollary discharge signalling in schizophrenia, but also for mixed evidence regarding the influence of history and knowledge on, for example, visual processing in schizophrenia, arguing that the uncertainty created by weaker priors in lower levels of a processing hierarchy are mitigated by stronger priors at higher levels [48,55–57].

Now we turn again to the distinction between treatment and non-treatment-seeking voice hearers and consider this distinction in terms of the reconciliation between the strong priors and weak corollary discharge account described above. Non-treatment-seeking voice hearers have a ready explanatory framework and have spent time cultivating volitional control over perception. In considering what this may tell us about the mechanisms of AVH in non-treatment-seeking voice-hearers, we invoke the notion of imagination within the predictive coding framework. A system that is equipped with the machinery to generate a prediction of the world in order to infer the causes of external stimuli can thus enable perception of the external world, but also self-generate perceptual states without external stimulation—that is, to imagine [58,59]. In the case of AVHs in non-treatment-seeking voice hearers, we assume that voices are experienced by weighting (perhaps volitionally) prior expectations (regarding some imagined event) more relative to the sensory evidence (that the event is not occurring). In this way, they can better anticipate AVHs and can understand them through their preferred explanatory lens [12]. This would be consistent with the sensitivity of non-treatment-seeking voice hearers to environmental volatility—the rate of change of environmental contingencies: they are able to update their beliefs rapidly and appropriately through learning, unlike the treatment-seeking groups [46]. If one can move between possible priors with alacrity, reconfiguring one's expectations smoothly and accurately, one would probably have strong attentional control over learning and belief updating, and by extension, imagination.

4. Predictions

We have an account of treatment-seeking and non-treatment-seeking voice hearing then, which evolves as follows.

In *treatment-seeking voice hearers*, the auditory cortex may be overly active and distracting in the absence of external stimulation. This activity may arise from multiple sources, including noise resulting from an altered neurochemical excitatory/inhibitory balance [60] or a dysfunctional efference copy failing to appropriately suppress sensory responses to the consequences of one's own actions/thoughts. As Leptourgos & Corlett note [14], this activity may be reconciled and understood by the voice hearer, perhaps in association cortices, by the allocentric top-down belief that someone is speaking [13,14]. Faulty efference copy signals result in a failure to predict and thus modulate the sensory response to not only the consequences of overt actions, but probably also the sensory consequences of one's own bodily signals. Indeed, there is ample evidence that bodily signals (e.g. heartbeats) are integrated with exteroceptive signals and influence perception [61], and auditory processing is modulated by the cardiac signal [62].

In *non-treatment-seeking voice hearers*, the auditory cortex may also be excessively active, perhaps resulting from a top-down 'listening attitude'—an openness to hearing speech when there is none [63] and strong priors in the form of imagination that allow for the generation of a vivid perceptual state. For example, there are non-treatment-seeking voice hearers who self-identify as clairaudient (that is, they perceive their voices as a cherished meta-physical gift [5]). It appears they engage in an imaginative practice that imposes strong priors on their perception [10], thereby overcoming the prediction error (engaged by the fact that no one is actually talking). This would imply that sensory attenuation, and by extension efference copy signalling, is intact and perhaps even under volitional control in non-treatment-seeking voice hearers.

In considering the latter proposition regarding the volitional nature of sensory attenuation in non-treatment-seeking voice hearers, we would highlight that two aspects of efference copy signalling that highlight its *selective* influence on sensory processing. First, there are situations, particularly in young animals, in which it is advantageous for self-generated movement not to generate an efference copy signal (or for that signal to be ignored) and for a self-generated movement to be processed as if it were generated externally [64,65]. This permits the animal important information about how it controls its own body—information that would typically be suppressed via an efference copy signal. Second, corollary discharge signals do not always suppress the sensory consequences of one's own action—in fact, sensory signals are sometimes enhanced by the efference copy signal, for example, when learning new language or song [66]. This is to say that the lack of a suppressing efference copy signal permits a system access to information about self-generated movements, speech and thoughts that serve a particular purpose for the individual and which may be relevant for the experience of AVHs figure 1.

Dynamic causal modelling analyses that invite the conclusion that neural signals are flowing from expectations downwards or sensations upwards, might speak to the proposed distinctions between treatment and non-treatment-seeking voice hearers. For example, we know from functional connectivity analyses that there is enhanced coupling between thalamus and auditory cortex in patients who hear voices and decreased coupling between thalamus and cerebellum in delusional patients [67]. The cerebellum [68] and thalamus are both loci for efference copy signals [69,70], with the thalamus being involved in specifying prior perceptual predictions [71] and prediction error processing in canonical cortical microcircuits [72]. It may be that the thalamus is a seat of perceptual belief updating [73]—weighting inputs from the association and primary cortices along with efference copies to arrive at the best explanation for the current circumstances. Data from Powers and colleagues [46] seemed to suggest that the cerebellum was likewise correlated with belief updating, with a dearth of cerebellar response to perceptual contingency change. Likewise, Shergill *et al.* [74] found the same during efference copy-driven force perception in patients with positive symptoms. It may be that the cerebellum is important for performing intentional actions (like thinking and imagining) and for processing the impact of those actions on perception [75]. Effective connectivity analyses might facilitate dissection of the relative weightings of different sources of information to different

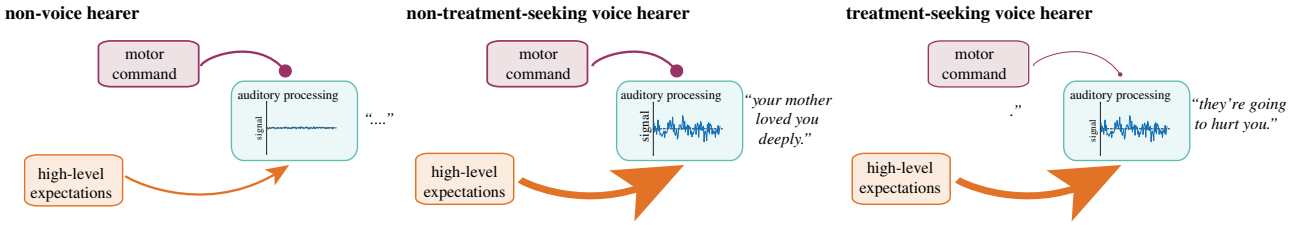

Figure 1. Predictions and prediction errors across the voice-hearing spectrum, healthy non-voice hearers, non-treatment seeking voice hearers and treatment seeking voice hearers (left to right, respectively). In each depiction, three units are considered in this oversimplified schematic: an auditory processing unit (blue), a motor command unit (purple) and a unit that represents high-level beliefs (orange). Accordingly, in all three, signals from the motor command unit represent efference copy signals associated with the content of speech/verbal thought that inhibit the auditory processing unit. Signals from higher-order regions represent higher-order beliefs about the auditory processing signal. The width of the lines indicates the strength of the corresponding signal. In all three, there is no external auditory information, and the net outcome or percept (represented in the text next to the auditory processing unit) is shaped by the interaction among these nodes. In non-voice hearers (left panel), we represent a system in which there is no excessive internal activity in the auditory processing unit, in which an intact corollary discharge signal is appropriately suppressing any sensory consequences of inner speech, and in which higher-order beliefs are optimally exerting their influence on the incoming sensory data—thus, perception is veridical and the observer hears nothing. In non-treatment seeking voice hearers (middle panel), we propose elevated activity in the auditory cortex, in this case, owing to strong expectations that are perhaps volitionally imposed, and in this example related to the spirit of one's deceased mother. In this case, these strong priors outweigh the actual sensory input (of silence) and the observer hears a comforting message. We propose that in these non-treatment-seeking voice hearers, the efference copy signal is intact. Finally, for treatment-seeking voice hearers (right panel), we predict that, in the absence of input, there is greater activity in the auditory cortex that may arise from internal activity and/or reduced efference copy signalling, leading to a failure to suppress the auditory response to internal speech. This activity is reconciled by a strong influence of an expectation for human speech and, in this case for danger, thus leading to the percept of a threatening voice. (Online version in colour.)

perceptual conclusions, in different individuals with and without hallucinations.

More broadly, what does this thesis predict about imagination, perception and hallucination?

(a) Engaging in imagery should change your perceptual experience

Empirical work suggests that this is acutely true. In difficult perceptual detection tasks, prior imagery increases the likelihood of false alarms [76]. Studying how some non-treatment-seeking voice hearers learn to cultivate their voice-hearing experiences suggest that they adopt a listening attitude [63], prepared to hear voices by imagining what it would be like to hear voices [10,11] in internal and external contexts in which they are most receptive, for example, when they are physically still and perceptually reduced [12]. Perhaps then, more perceptually structured/predictable and motorically engaged settings might help treatment-seeking voice hearers. Indeed, listening to and performing music mollifies voice hearing [77].

(b) People with strong imagery should be at higher risk for psychosis and hallucinations in particular

There is a robust body of evidence indicating enhanced vividness of mental imagery in schizophrenia based on both subjective report and experimental data [78–81]. Stronger imagery has sometimes [82–84] but most often not [79,84,85] related to hallucination severity, thus leading to the conclusion that enhanced mental imagery confers vulnerability to hallucinations, but that it is probably not a proximal mechanism. One potential conceptualization of these findings in schizophrenia is that there is a greater ease by which long-term memory representations can be brought into the current contents of consciousness [81]. Indeed, the notion of impaired inhibition of information from long-term memory is present in several mechanistic accounts of hallucinations [86–88], and has been supported by neuroimaging data [89]. These findings bear, perhaps, on predictive coding accounts of hallucinations whereby the contents of long-term memory can be mapped onto the prior distribution.

(c) Treatment and non-treatment-seeking voice hearers should differ in their volitional control over sensory attenuation

We expect that tasks which assay sensory attenuation—like force-matching, efference copies of eye-movements and speaking/listening signals should be relatively intact in non-treatment-seeking voice hearers. They should retain the ability to distinguish between perception and imagination when they choose to, which could be assayed with reality monitoring tasks. We are arguing that on occasion, non-treatment-seeking voice hearers voluntarily, in a sense, overweight their priors, manifesting as hallucinations. Non-treatment-seeking voice hearers are less likely to possess such control, although perhaps this is something that they can learn via dialogic interaction with their voices, as advocated by the hearing voices movement and hearing voices network meetings.

(d) Imagery should duplicate the hallucinatory experience

A growing body of evidence from brain imaging studies suggests internal voices, or AVH, activate the same speech production and reception areas as do external voices coming from the environment. This evidence comes from symptom capture studies of people with a diagnosis of schizophrenia who hear voices [90] and from non-help seeking people who hear voices [91–93]. There is also evidence that healthy people, who do not hear voices and who do not have a psychiatric diagnosis, also activate those brain areas when instructed to imagine speech in their own voices [94].

Box 1. Some possible ways to reconcile predictive coding and corollary discharge accounts of AVH.

— Test the relationship between EEG-based corollary discharge function and weighting of expectations in both treatment and non-treatment seeking voice hearers.
— Test functional connectivity between cerebellum, thalamus, and auditory cortex using cerebellum seeds derived from the motor act of vocalizing (tapping the corollary discharge mechanism) and seeds derived from conditioned hallucinations.
— Using multi-modal imaging, compare treatment and non-treatment seeking voice hearers' neural-metabolic coupling, as it relates to severity of voice hearing.
— Using EEG assays of corollary discharge function, compare treatment and non-treatment seeking voice hearers.

However, when imaging speech in another person's voice, schizophrenia patients with a predisposition to hear voices failed to activate brain areas involved in monitoring inner speech [94].

5. Conclusion

In summary then, in exploring the constructive nature of perception, of which volitional imagination is a special case, we have cast some light on hallucinations, their underlying mechanisms and possible treatment. This is the beginnings of a sketch, which probably raises more questions than it answers, however, it challenges some long-held convictions about hallucinations, and it signals some new lines of inquiry we hope that we and others can follow. We list some ways to compare the two accounts of AVH (box 1).

Data accessibility. This article has no additional data.

Authors' contributions. We all conceived of, wrote, and edited the paper and agreed on the final version.

Competing interests. We declare we have no competing interests.

Funding. This study was supported by grants from the VA Merit I01CX000497 programme, VA Senior Research Career Award to J.M.F. and the National Institutes of Health (NIH) (grant nos. R01MH076989, R01 MH058262, R01 MH112644).

Acknowledgements. We thank Dr Philip Corlett for his contributions to an earlier draft of this paper.

Endnote

[1]While we believe that conditioning is one route to hallucinations, the conditional stimuli need not be external environmental cues; they could just as readily be changes in variables in the internal milieu, like heart rate, blood pressure, cortical activity or neurotransmitter release.

References

1. Jones SR. 2010 Do we need multiple models of auditory verbal hallucinations? Examining the phenomenological fit of cognitive and neurological models. *Schizophr. Bull.* **36**, 566–575. (doi:10.1093/schbul/sbn129)
2. Powers III AR, Kelley M., Corlett PR. 2016 Hallucinations as top-down effects on perception. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **1**, 393–400. (doi:10.1016/j.bpsc.2016.04.003)
3. Sommer IE, Daalman K, Rietkerk T, Diederen KM, Bakker S, Wijkstra J, Boks MP. 2010 Healthy individuals with auditory verbal hallucinations; who are they? Psychiatric assessments of a selected sample of 103 subjects. *Schizophr. Bull.* **36**, 633–641. (doi:10.1093/schbul/sbn130)
4. Cicchetti D, Rogosch FA. 1996 Equifinality and multifinality in developmental psychopathology. *Dev. Psychopathol* **8**, 597–600. (doi:10.1017/S0954579400007318)
5. Powers III AR, Kelley MS, Corlett PR. 2017 Varieties of voice-hearing: psychics and the psychosis continuum. *Schizophr. Bull.* **43**, 84–98. (doi:10.1093/schbul/sbw133)
6. Swyer A, Powers III AR. 2020 Voluntary control of auditory hallucinations: phenomenology to therapeutic implications. *NPJ Schizophr.* **6**, 19. (doi:10.1038/s41537-020-0106-8)
7. Daalman K, Diederen KM, Hoekema L, van Lutterveld R, Sommer IE. 2016 Five year follow-up of non-psychotic adults with frequent auditory verbal hallucinations: are they still healthy? *Psychol. Med.* **46**, 1897–1907. (doi:10.1017/S0033291716000386)
8. Powers AR, Bien C, Corlett PR. 2018 Hearing their voices: aligning computational psychiatry with the hearing voices movement. *JAMA Psychiatry* **75**, 640. (doi:10.1001/jamapsychiatry.2018.0509)
9. Leff J, Williams G, Huckvale M, Arbuthnot M, Leff AP. 2014 Avatar therapy for persecutory auditory hallucinations: what is it and how does it work? *Psychosis* **6**, 166–176. (doi:10.1080/17522439.2013.773457)
10. Luhrmann TM. 1989 *Persuasions of the witch's craft: ritual magic and witchcraft in present-day England*, viii, 382 pp. Oxford, UK: B. Blackwell.
11. Luhrmann T.M. 2012 *When God talks back: understanding the American evangelical relationship with God*, xxv, 434 pp, 1st edn. New York, NY: Alfred A. Knopf.
12. Asprem E. 2017 Explaning the esoteric imagination. *Aries* **17**, 17–50. (doi:10.1163/15700593-01701002)
13. Corlett PR, Horga G, Fletcher PC, Alderson-Day B, Schmack K, Powers III AR. 2018 Hallucinations and strong priors. *Trends Cogn. Sci.* **23**, 114–127. (doi:10.1016/j.tics.2018.12.001)
14. Leptourgos P, Corlett PR. 2020 Embodied predictions, agency, and psychosis. *Front. Big Data* **3**, 27. (doi:10.3389/fdata.2020.00027)
15. Frith CD. 1987 The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychol. Med.* **17**, 631–648. (doi:10.1017/S0033291700025873)
16. Ford JM, Mathalon DH. 2005 Corollary discharge dysfunction in schizophrenia: can it explain auditory hallucinations? *Int. J. Psychophysiol.* **58**, 179–189. (doi:10.1016/j.ijpsycho.2005.01.014)
17. Hughlings-Jackson J. 1878 On affectations of speech from disease of the brain. *Brain* **1**, 304–330. (doi:10.1093/brain/1.3.304)
18. Crapse TB, Sommer MA. 2008 Corollary discharge circuits in the primate brain. *Curr. Opin. Neurobiol.* **18**, 552–557. (doi:10.1016/j.conb.2008.09.017)
19. Eliades SJ, Wang X. 2008 Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* **453**, 1102–1106. (doi:10.1038/nature06910)
20. Ford JM, Gray M, Faustman WO, Roach BJ, Mathalon DH. 2007 Dissecting corollary discharge dysfunction in schizophrenia. *Psychophysiology* **44**, 522–529. (doi:10.1111/j.1469-8986.2007.00533.x)
21. Wang J, Mathalon DH, Roach BJ, Reilly J, Keedy SK, Sweeney JA, Ford JM. 2014 Action planning and predictive coding when speaking. *Neuroimage* **91**, 91–98. (doi:10.1016/j.neuroimage.2014.01.003)
22. Greenlee JD, Jackson AW, Chen F, Larson CR, Oya H, Kawasaki H, Chen H, Howard III MA. 2011 Human

auditory cortical activation during self-vocalization. *PLoS ONE* **6**, e14744. (doi:10.1371/journal.pone.0014744)

23. Heinks-Maldonado TH, Mathalon DH, Houde JF, Gray M, Faustman WO, Ford JM. 2007 Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. *Arch. Gen. Psychiatry* **64**, 286–296. (doi:10.1001/archpsyc.64.3.286)
24. Ford JM, Roach BJ, Mathalon DH. 2010 Assessing corollary discharge in humans using noninvasive neurophysiological methods. *Nat. Protoc.* **5**, 1160–1168. (doi:10.1038/nprot.2010.67)
25. Ford JM, Roach BJ, Palzes VA, Mathalon DH. 2016 Using concurrent EEG and fMRI to probe the state of the brain in schizophrenia. *Neuroimage Clin.* **12**, 429–441. (doi:10.1016/j.nicl.2016.08.009)
26. Ford JM, Mathalon DH, Heinks T, Kalba S, Roth WT. 2001 Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. *Am. J. Psychiatry* **158**, 2069–2071. (doi:10.1176/appi.ajp.158.12.2069)
27. Mathalon DH, Roach BJ, Ferri JM, Loewy RL, Stuart BK, Perez VB, Trujillo TH, Ford JM. 2019 Deficient auditory predictive coding during vocalization in the psychosis risk syndrome and in early illness schizophrenia: the final expanded sample. *Psychol. Med.* **49**, 1897–1904. (doi:10.1017/S0033291718002659)
28. Roach BJ, Ford JM, Loewy RL, Stuart BK, Mathalon DH. In press. Theta phase synchrony is sensitive to corollary discharge abnormalities in early illness schizophrenia but not in the psychosis risk syndrome. *Schizophr. Bull.*
29. Ford JM, Roach BJ, Faustman WO, Mathalon DH. 2007 Synch before you speak: auditory hallucinations in schizophrenia. *Am. J. Psychiatry* **164**, 458–466. (doi:10.1176/appi.ajp.164.3.458)
30. Ford JM, Mathalon DH, Whitfield S, Faustman WO, Roth WT. 2002 Reduced communication between frontal and temporal lobes during talking in schizophrenia. *Biol. Psychiatry* **51**, 485–492. (doi:10.1016/S0006-3223(01)01335-X)
31. Bansal S, Ford JM, Spering M. 2018 The function and failure of sensory predictions. *Ann. NY Acad. Sci.* **1426**, 199–220. (doi:10.1111/nyas.13686)
32. Thakkar KN, Diwadkar VA, Rolfs M. 2017 Oculomotor prediction: a window into the psychotic mind. *Trends Cogn. Sci.* **21**, 344–356. (doi:10.1016/j.tics.2017.02.001)
33. Pynn LK, DeSouza JF. 2013 The function of efference copy signals: implications for symptoms of schizophrenia. *Vision Res.* **76**, 124–133. (doi:10.1016/j.visres.2012.10.019)
34. Pinheiro AP, Schwartze M, Kotz SA.2020 Cerebellar circuitry and auditory verbal hallucinations: an integrative synthesis and perspective. *Neurosci. Biobehav. Rev.* **118**, 485–503. (doi:10.1016/j.neubiorev.2020.08.004)
35. von Helmholtz H. 1878/1971 The facts of perception. In *Selected writings of Herman von Helmholtz* (ed. R Kahl), pp. 366–408. Middletown, CT: Weslyan University Press.
36. Clark A. 2013 Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181–204.
37. Feldman H, Friston KJ. 2010 Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* **4**, 215. (doi:10.3389/fnhum.2010.00215)
38. Friston K, Kiebel S. 2009 Predictive coding under the free-energy principle. *Phil. Trans. R. Soc. B* **364**, 1211–1221. (doi:10.1098/rstb.2008.0300)
39. Friston KJ, Stephan KE. 2007 Free-energy and the brain. *Synthese* **159**, 417–458. (doi:10.1007/s11229-007-9237-y)
40. Teufel C, Subramaniam N, Fletcher PC. 2013 The role of priors in Bayesian models of perception. *Front. Comput. Neurosci.* **7**, 25. (doi:10.3389/fncom.2013.00025)
41. Friston KJ. 2005 Hallucinations and perceptual inference. *Behav. Brain Sci.* **28**, 764–766. (doi:10.1017/S0140525X05290131)
42. Seashore CE. 1895 Measurements of illusions and hallucinations in normal life. *Studies from the Yale Psychological Laboratory* **3**, 1–67.
43. Pavlov IP, Horsley GW. 1928 *Lectures on conditioned reflexes: twenty-five years of objective study of the higher nervous activity (behaviour) of animals*, pp. 120–130. New York, NY: Liverwright Publishing Corporation.
44. Ellson DG. 1941 Hallucinations produced by sensory conditioning. *J. Exp. Psychol.* **28**, 1–20. (doi:10.1037/h0054167)
45. Kot T, Serper M. 2002 Increased susceptibility to auditory conditioning in hallucinating schizophrenic patients: a preliminary investigation. *J. Nerv. Ment. Dis.* **190**, 282–288. (doi:10.1097/00005053-200205000-00002)
46. Powers AR, Mathys C, Corlett PR. 2017 Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science* **357**, 596–600. (doi:10.1126/science.aan3458)
47. Jardri R, Pouchet A, Pins D, Thomas P. 2011 Cortical activations during auditory verbal hallucinations in schizophrenia: a coordinate-based meta-analysis. *Am. J. Psychiatry* **168**, 73–81. (doi:10.1176/appi.ajp.2010.09101522)
48. Teufel C, Subramaniam N, Dobler V, Perez J, Finnemann J, Mehta PR, Goodyer IM, Fletcher PC. 2015 Shift toward prior knowledge confers a perceptual advantage in early psychosis and psychosis-prone healthy individuals. *Proc. Natl Acad. Sci. USA* **112**, 13 401–13 406. (doi:10.1073/pnas.1503916112)
49. Alderson-Day B, Lima CF, Evans S, Krishnan S, Shanmugalingam P, Fernyhough C, Scott SK. 2017 Distinct processing of ambiguous speech in people with non-clinical auditory verbal hallucinations. *Brain* **140**, 2475–2489. (doi:10.1093/brain/awx206)
50. Doucet GE, Luber MJ, Balchandani P, Sommer IE, Frangou S. 2019 Abnormal auditory tonotopy in patients with schizophrenia. *NPJ Schizophr.* **5**, 16. (doi:10.1038/s41537-019-0084-x)
51. Rohe T, Noppeney U. 2018 Reliability-weighted integration of audiovisual signals can be modulated by top-down attention. *eNeuro* **5**, 0315-0317.2018. (doi:10.1523/ENEURO.0315-17.2018)
52. Rohe T, Noppeney U. 2016 Distinct computational principles govern multisensory integration in primary sensory and association cortices. *Curr. Biol.* **26**, 509–514. (doi:10.1016/j.cub.2015.12.056)
53. Rohe T, Noppeney U. 2015 Sensory reliability shapes perceptual inference via two mechanisms. *J. Vis.* **15**, 22. (doi:10.1167/15.5.22)
54. Rohe T, Noppeney U. 2015 Cortical hierarchies perform Bayesian causal inference in multisensory perception. *PLoS Biol.* **13**, e1002073. (doi:10.1371/journal.pbio.1002073)
55. Keane BP, Silverstein SM, Wang Y, Papathomas TV. 2013 Reduced depth inversion illusions in schizophrenia are state-specific and occur for multiple object types and viewing conditions. *J. Abnorm. Psychol.* **122**, 506–512. (doi:10.1037/a0032110)
56. Schneider U, Borsutzky M, Seifert J, Leweke FM, Huber TJ, Rollnik JD, Emrich HM. 2002 Reduced binocular depth inversion in schizophrenic patients. *Schizophr. Res.* **53**, 101–108. (doi:10.1016/S0920-9964(00)00172-9)
57. Schmack K, Rothkirch M, Priller J, Sterzer P. 2017 Enhanced predictive signalling in schizophrenia. *Hum. Brain Mapp.* **38**, 1767–1779. (doi:10.1002/hbm.23480)
58. Kirchhoff MD. 2018 Predictive processing, perceiving and imagining: is to perceive to imagine, or something close to it? *Philos. Stud.* **175**, 751–767. (doi:10.1007/s11098-017-0891-8)
59. Clark A. 2016 *Surfing uncertainty*. Oxford, UK: Oxford University Press.
60. Grace AA. 1991 Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience* **41**, 1–24. (doi:10.1016/0306-4522(91)90196-u)
61. Allen M, Frank D, Schwarzkopf DS, Fardo F, Winston JS, Hauser TU, Rees G. 2016 Unexpected arousal modulates the influence of sensory noise on confidence. *Elife* **5**, e18103. (doi:10.7554/eLife.18103)
62. van Elk M, Lenggenhager B, Heydrich L, Blanke O. 2014 Suppression of the auditory N1-component for heartbeat-related sounds reflects interoceptive predictive coding. *Biol. Psychol.* **99**, 172–182. (doi:10.1016/j.biopsycho.2014.03.004)
63. Hoffman RE. 2010 Revisiting Arieti's 'listening attitude' and hallucinated voices. *Schizophr. Bull.* **36**, 440–442. (doi:10.1093/schbul/sbq025)
64. Tiriac A, Del Rio-Bermudez C, Blumberg MS. 2014 Self-generated movements with "unexpected" sensory consequences. *Curr. Biol.* **24**, 2136–2141. (doi:10.1016/j.cub.2014.07.053)
65. Tiriac A, Sokoloff G, Blumberg MS. 2015 Myoclonic twitching and sleep-dependent plasticity in the developing sensorimotor system. *Curr. Sleep Med. Rep.* **1**, 74–79. (doi:10.1007/s40675-015-0009-9)
66. Schneider DM, Mooney R. 2018 How movement modulates hearing. *Annu. Rev. Neurosci.* **41**, 553–572. (doi:10.1146/annurev-neuro-072116-031215)
67. Ferri J *et al.* 2018 Resting-state thalamic dysconnectivity in schizophrenia and relationships with symptoms. *Psychol. Med.* **48**, 2492–2499. (doi:10.1017/S003329171800003X)

68. Ramnani N. 2006 The primate cortico-cerebellar system: anatomy and function. *Nat. Rev. Neurosci.* **7**, 511–522. (doi:10.1038/nrn1953)
69. Bellebaum C, Daum I, Koch B, Schwarz M, Hoffmann KP. 2005 The role of the human thalamus in processing corollary discharge. *Brain* **128**(Pt 5), 1139–1154. (doi:10.1093/brain/awh474)
70. Yao B, Neggers SFW, Rolfs M, Rosler L, Thompson IA, Hopman HJ, Ghermezi L, Kahn RS, Thakkar KN. 2019 Structural thalamofrontal hypoconnectivity is related to oculomotor corollary discharge dysfunction in schizophrenia. *J. Neurosci.* **39**, 2102–2113. (doi:10.1523/JNEUROSCI.1473-18.2019)
71. Horga G, Schatz KC, Abi-Dargham A, Peterson BS. 2014 Deficits in predictive coding underlie hallucinations in schizophrenia. *J. Neurosci.* **34**, 8072–8082. (doi:10.1523/JNEUROSCI.0200-14.2014)
72. Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. 2012 Canonical microcircuits for predictive coding. *Neuron* **76**, 695–711. (doi:10.1016/j.neuron.2012.10.038)
73. Rikhye RV, Wimmer RD, Halassa MM. 2018 Toward an integrative theory of thalamic function. *Annu. Rev. Neurosci.* **41**, 163–183. (doi:10.1146/annurev-neuro-080317-062144)
74. Shergill SS, White TP, Joyce DW, Bays PM, Wolpert DM, Frith CD. 2014 Functional magnetic resonance imaging of impaired sensory prediction in schizophrenia. *JAMA Psychiatry* **71**, 28–35. (doi:10.1001/jamapsychiatry.2013.2974)
75. Lotze M, Montoya P, Erb M, Hulsmann E, Flor H, Klose U, Birbaumer N, Grodd W. 1999 Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *J. Cogn. Neurosci.* **11**, 491–501. (doi:10.1162/089892999563553)
76. Moseley P, Smailes D, Ellison A, Fernyhough C. 2016 The effect of auditory verbal imagery on signal detection in hallucination-prone individuals. *Cognition* **146**, 206–216. (doi:10.1016/j.cognition.2015.09.015)
77. Silverman MJ. 2003 The influence of music on the symptoms of psychosis: a meta-analysis. *J. Music Ther.* **40**, 27–40. (doi:10.1093/jmt/40.1.27)
78. Benson TL, Park S. 2013 Exceptional visuospatial imagery in schizophrenia; implications for madness and creativity. *Front. Hum. Neurosci.* **7**, 756. (doi:10.3389/fnhum.2013.00756)
79. Aleman A, de Haan EH, Kahn RS. 2005 Object versus spatial visual mental imagery in patients with schizophrenia. *J. Psychiatry Neurosci.* **30**, 53–56.
80. Thakkar KN, Park S. 2012 Impaired passive maintenance and spared manipulation of internal representations in patients with schizophrenia. *Schizophr. Bull.* **38**, 787–795. (doi:10.1093/schbul/sbq159)
81. Matthews NL, Collins KP, Thakkar KN, Park S. 2014 Visuospatial imagery and working memory in schizophrenia. *Cogn. Neuropsychiatry* **19**, 17–35. (doi:10.1080/13546805.2013.779577)
82. Bocker KB, Hijman R, Kahn RS, De Haan EH. 2000 Perception, mental imagery and reality discrimination in hallucinating and non-hallucinating schizophrenic patients. *Br. J. Clin. Psychol.* **39**, 397–406. (doi:10.1348/0144665 00163392)
83. Mintz S, Alpert M. 1972 Imagery vividness, reality testing, and schizophrenic hallucinations. *J. Abnorm. Psychol.* **79**, 310–316. (doi:10.1037/h0033209)
84. Sack AT, van de Ven VG, Etschenberg S, Schatz D, Linden DE. 2005 Enhanced vividness of mental imagery as a trait marker of schizophrenia? *Schizophr. Bull.* **31**, 97–104. (doi:10.1093/schbul/sbi011)
85. Oertel V, Rotarska-Jagiela A, van de Ven V, Haenschel C, Grube M, Stangier U, Maurer K, Linden DE. 2009 Mental imagery vividness as a trait marker across the schizophrenia spectrum. *Psychiatry Res.* **167**, 1–11. (doi:10.1016/j.psychres.2007.12.008)
86. George L, Neufeld RW. 1985 Cognition and symptomatology in schizophrenia. *Schizophr. Bull.* **11**, 264–285. (doi:10.1093/schbul/11.2.264)
87. Hemsley DR. 1987 Hallucinations: unintended or unexpected. *Behav. Brain Sci.* **10**, 532–533. (doi:10.1017/S0140525X00023943)
88. Rund BR. 1986 Verbal hallucinations and information processing. *Behav. Brain Sci.* **9**, 531–532. (doi:10.1017/S0140525X00046975)
89. Diederen KM, Neggers SF, Daalman K, Blom JD, Goekoop R, Kahn RS, Sommer IE. 2010 Deactivation of the parahippocampal gyrus preceding auditory hallucinations in schizophrenia. *Am. J. Psychiatry* **167**, 427–435. (doi:10.1176/appi.ajp.2009.09040456)
90. McGuire PK, Shah GM, Murray RM. 1993 Increased blood flow in Broca's area during auditory hallucinations in schizophrenia. *Lancet* **342**, 703–706. (doi:10.1016/0140-6736(93)91707-S)
91. Daalman K, Boks MP, Diederen KM, de Weijer AD, Blom JD, Kahn RS, Sommer IE. 2011 The same or different? A phenomenological comparison of auditory verbal hallucinations in healthy and psychotic individuals. *J. Clin. Psychiatry* **72**, 320–325. (doi:10.4088/JCP.09m05797yel)
92. Diederen KM, Daalman K, de Weijer AD, Neggers SF, van Gastel W, Blom JD, Kahn RS, Sommer IE. 2012 Auditory hallucinations elicit similar brain activation in psychotic and nonpsychotic individuals. *Schizophr. Bull.* **38**, 1074–1082. (doi:10.1093/schbul/sbr033)
93. Diederen KM, van Lutterveld R, Sommer IE. 2012 Neuroimaging of voice hearing in non-psychotic individuals: a mini review. *Front. Hum. Neurosci.* **6**, 111. (doi:10.3389/fnhum.2012.00111)
94. McGuire PK, Silbersweig DA, Wright I, Murray RM, David AS, Frackowiak RS, Frith CD. 1995 Abnormal monitoring of inner speech: a physiological basis for auditory hallucinations. *Lancet* **346**, 596–600. (doi:10.1016/s0140-6736(95)91435-8)