CellPress
OPEN ACCESS

Trends in
Neurosciences

Opinion

Metacognition in the listening brain

Jonas Obleser[1,2,*]

How do you know you have heard right? Metacognition, the ability to assess and monitor one's own cognitive state, is key to understanding human communication in complex environments. However, the foundational role of metacognition in hearing and communication is only beginning to be explored, and the neuroscience behind it is an emerging field: how does confidence express in neural dynamics of the listening brain? What is known about auditory metaperceptual alterations as a hallmark phenomenon in psychosis, dementia, or hearing loss? Building on Bayesian ideas of auditory perception and auditory neuroscience, 'meta-listening' offers a framework for more comprehensive research into how metacognition in humans and non-humans shapes the listening brain.

Highlights

Metacognition is key to understanding and modelling human auditory perception and communication, particularly in complex environments.

The proposed 'meta-listening' framework is based on Bayesian models of perception, exploring how the precision of neural encoding might drive auditory metacognition.

The roles of alpha (~8–12 Hz) oscillatory activity and cortical excitation-to-inhibition balance in shaping metacognitive outcomes in the auditory sensory modality are highlighted.

Subjective confidence judgements in hearing often align with objective performance measures, but there are significant instances of dissociation, particularly in older adults and in individuals with auditory hallucinations.

Current limits in explaining listening behaviour

Imagine yourself on a busy train platform. You hear your name being called out. How confident are you that you have heard correctly? Should you turn and look for who is calling? This question – 'Did I hear this right?' – is a question of **metacognition** (see Glossary), that is, the knowledge of one's own cognitive processes [1]. As listeners navigate a noisy world, perceiving, deciding, and behaving in the face of this perceptual uncertainty calls on their metacognitive or metaperceptive abilities.

A ubiquitous phenomenon is human perceivers' ability to report their confidence in a perceptual decision [2]. Psychologists have recognised this early on [3] and began using confidence rating scales (e.g., 0–3) still in use today. More generally, judging or monitoring one's own hearing ('I did not understand this', 'this is too effortful to listen to') represents a metapercept, as it results from the listener 'reading out' their internal perceptual state.

Under a Bayesian model of perception (Box 1), confidence aligns closely with perceptual precision: Confidence should be proportional to the precision (i.e., inverse of the variance) of the posterior probability for that percept. Analogously, in artificial neural networks (Box 2), confidence or precision is sometimes expressed as the entropy across possible candidates, with low entropy expressing high precision or high confidence [4]. More formally, confidence can be computationally defined as a conditional probability: namely, the (subjective) probability that a perceptual decision has been correct, given this decision and the according evidence [5].

Such metaperceptual judgements have immediate as well as more remote consequences for behaviour. In other words, it does make a difference whether you do, or do not, 'trust your ears'. Audition is a particularly fleeting, time-dependent sensory modality, and a listener's confidence in what they have just perceived is integral to their ensuing behaviour. For instance, when a person with hearing loss lacks confidence in her perceptual abilities, she may ask more often for a sentence being heard to be repeated. Another, more confident or overconfident listener might be less inclined to seek such repetition.

Overconfidence has been identified as a potentially adaptive hallmark of human decision-making and might be evolutionary adaptive [2]. In auditory perception, overconfidence can be observed in

[1]Department of Psychology, University of Lübeck, 23562 Lübeck, Germany
[2]Center of Brain, Behavior, and Metabolism, University of Lübeck, 23562 Lübeck, Germany

*Correspondence:
jonas.obleser@uni-luebeck.de
(J. Obleser).

https://doi.org/10.1016/j.tins.2024.12.007
© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Box 1. Auditory percepts as a Bayesian compromise

Within a Bayesian framework of perception, a listener's prior assumptions (prior), the ascribed likelihood of the incoming sensory evidence under these assumptions (likelihood or sensory evidence), and the resulting *a posteriori* percept are not fixed values. They all reflect form probability distributions (Figure I). These distributions of belief extend over a continuum of possible true states ('unknown parameters') in the environment, such as sound frequency or a prey's location in space. The height of the posterior distributions at any point represents the listener's belief in that value being the true value generating the percept. This is easily seen for one-dimensional phenomena, such as the exact tonal frequency or the azimuth location of prey in space. Applying this concept to more acoustically complex and multidimensional percepts as a human voice requires some extrapolation, but the underlying principle remains.

In Bayesian terms, a 'flat prior' renders a wide range of possible true states to be of almost equal probability. Conversely, a 'strong' or precise *a priori* belief is characterised by a narrow and tall prior distribution, where a small range of possible true states carries high probability or strong belief, while other potential true states are deemed much less likely or nearly impossible.

Relevant to auditory perception, the mathematics of Bayesian perception dictate that (i) the resulting *a posteriori* percept will always be more precise than either the prior assumption or the sensory evidence alone (middle panel of Figure I); and (ii) higher precision in either of these will translate directly into a stronger impact on the resulting percept (Figure I).

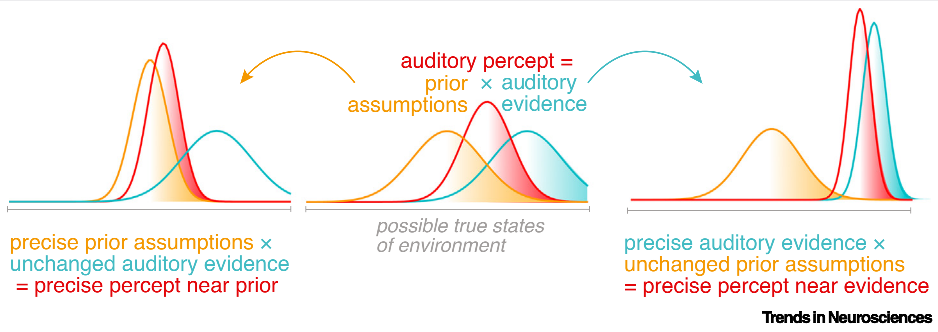


Figure I. Auditory perception as a Bayesian compromise. Note how precise priors (i.e., strong assumptions; orange) and precise sensory evidence (e.g., high-acuity neural encoding in the auditory periphery; cyan) can both have a strong impact on the resulting auditory percept (i.e., the mode of red distributions) and its precision (i.e., the inverse of the red distribution's variance). This degree of precision can inform the listener's subsequent metacognitive confidence judgement: both the left and right scenario should result in a more precise posterior and therefore higher confidence than the middle one.

some individuals in response to false alarms, that is, they react to illusory sensory stimuli with unwarranted high degrees of certainty. Similar behaviour has recently been demonstrated in rodents exhibiting hallucination-like percepts [6]. Thus, an individual with healthy hearing but certain perceptual dispositions might tend to 'hear voices' where there are none or have auditory percepts that their peers do not share (see 'How does metacognition relate to aberrant auditory perception?' later).

In light of this ubiquity and relevance of metacognition to hearing and human behaviour in complex environments, it is somewhat surprising that auditory neuroscience has not taken up and explored the role of metacognition to equal extent as some other domains of cognition. Metacognitive research and its elaborate models of confidence have often relied on visual phenomena to hone their predictions [5,7,8] (but see [9,10]). Meanwhile, the growing literature on neural mechanisms of audition, speech, and communication holds many hints to a lurking, often underexplored role of metacognition. This article aims to highlight these developments and promote work towards a testable framework of 'meta-listening'.

What is meta-listening, and why does it matter?

Metacognition is a field of intense debate, and definitions of metacognition vary across the literature. For the current article, the following definition, intended to be broadly in line with current

Glossary

Bayesian models of perception: percepts are well captured by a multiplicative process or 'perceptual compromise' that balances the perceiver's prior belief or expectation with how probable the incoming sensory evidence is under this prior belief (i.e., the likelihood).

Electrocorticography (ECoG): the recording of electric brain activity directly from inside the skull, usually with grids of electrodes placed temporarily on or under the dura mater in severe cases of epilepsy before neurosurgery.

Electroencephalography (EEG): a non-invasive technique for recording electric brain activity from electrodes placed on the scalp.

Metacognition: a family of cognitive processes acting upon cognition itself, for example, monitoring and evaluating one's own perceptions, judgements, emotions, or behaviours.

Metacognitive efficiency: an agent when making metacognitive decisions can make more or less efficient use of internal perceptual decisions. If both decisions are expressed in terms of sensitivity measures d' and meta d', the so-called m-ratio (meta-d' divided by d') expresses metacognitive efficiency.

Oscillation: a rhythmic signal with a defined time scale (frequency), amplitude, and phase. The phase expresses the position of peaks (and troughs) relative to some reference. Oscillations occur in the context of auditory perception in neural signals or in time series of behavioural reports.

Box 2. Metacognition for the listening machine

Artificial neural networks (ANNs) are now nearly omnipresent, including in auditory and verbal communication settings (e.g., in cars and smartphones, and not least in assistive systems in critical sectors like healthcare and security). This calls for systematically testing the validity of different models of confidence in emulating behavioural and neural processes of human cognition [98].

Amidst the fast progress in 'machine listening' and the increasingly ubiquitous role of auditory or language-based human–machine interactions [99], we lack a good model of how forms of 'joint' metacognition between a listening machine and a human could be achieved. How could rudimentary metacognitive abilities for machine listening be implemented?

The advent of computationally powerful and effective, but mostly 'black-box' generative transformer architectures as in ChatGPT [100] highlight the need for transparent, 'explainable', and therefore safe-to-use algorithms [101,102]. Interesting new leads in this respect are endeavours to 'open the box' and scrutinise the layer-specific representations in language-predictive artificial neural networks by a joint analysis with humans' neural representations and communication behaviour [103–105] .

Why is this relevant for future work on metacognition and confidence in listening? 'Conventional' artificial-intelligence functionalities in hearing-aids, chatbots, and car navigation systems can all be considered to rest on 'first-order' perceptual judgements, utilising and maximising the network's prediction accuracy. Very few, if any, state-of-the-art machine-learning approaches to speech recognition or natural language processing have thus far been aimed at explicitly implementing levels of second-order judgement and metacognitive efficiency in these artificial agents. The question remains how accurately the machine can diagnose its own (in-)accuracy (but see in vision, e.g., [106]).

consensus, will be used. Readers are referred to recent reviews for more detailed discussions, for example, [5,7,11].

As a working definition, 'meta-listening' summarises those neural and psychological mechanisms that allow a listener (i) to be aware of some aspect or feature of their auditory perception; (ii) to assess or 'read out' and judge this perceptual aspect; and accordingly (iii) to use this information about their percept (i.e., not only the percept itself) to report on it and to act on it.

How could such metaperceptual read-outs be neurobiologically implemented (Figure 1, Key figure)? Can the neurobiology of auditory metaperception help us predict when a listener will be subjectively performing well (i.e., experiencing 'listening success')? This explanatory gap matters: metacognitive abilities might prove relevant to communication satisfaction, to compliance in hearing rehabilitation, and to help-seeking behaviour more generally. Also, the increasingly recognised value of hearing loss as a key, potentially modifiable risk factor for two disorders involving metacognitive monitoring, namely dementia [12,13] and psychosis [14], call for a more comprehensive understanding of metacognition in audition.

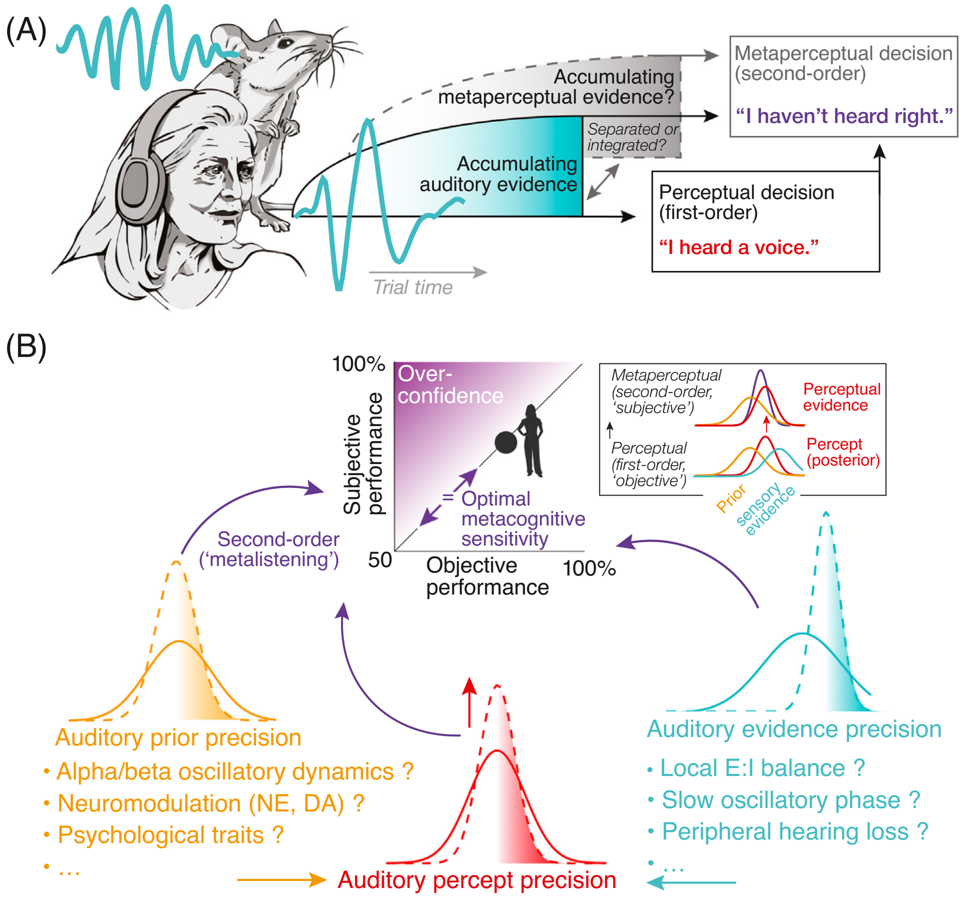
Dissociations of objective and subjective listening success

Metaperceptual and perceptual processes can be differentiated using concepts from signal detection theory [15]. This allows us to distinguish 'type I' or first-order, perceptual decisions from 'type II' or second-order, metaperceptual decisions [15,16]. First-order perception involves a decision criterion or bias (e.g., a listener's level of evidence required to perceive a sensory signal) and sensitivity (e.g., how segregated the representation of the sensory signal is from sensory noise) [17].

The signal-detection framework can be applied to second-order perception as well: a listener's metacognitive bias reflects how much perceptual evidence they require to feel confident in a given percept. Their metacognitive sensitivity indicates how well they distinguish correctly perceived signals from false alarms. Following from this, **metacognitive efficiency** is sometimes quantified by normalising an individual's metacognitive sensitivity on their first-order sensitivity (*m*-ratio) [8,18].

Key figure

Auditory perception, the integration of perceptual priors with auditory evidence, and its relation to metacognitive assessments of own perception ('meta-listening')



Trends in Neurosciences

Figure 1. (A) In human and non-human audition, the spectro-temporal or spatial accumulation of sensory evidence (shown in cyan colour) allows perceptual decisions (e.g., detecting a voice; localising prey; or discriminating conspecific vocalisations from background noise; 'first-order' percepts, shown in red). Also, both humans and non-human animals show evidence of a separable, metacognitive level of monitoring and reporting about this perceptual process (here termed 'meta-listening'; purple), for example, statements of confidence in the perceptual decisions just taken. There is debate what are the most important sources of evidence for this second-order, meta-perceptual judgement (sketched here as a separate stream of metaperceptual evidence accumulation for illustration only; see also purple arrows in (B). For detail and debate see, for example, [19,97]). (B) A common framework conceptualising any auditory percept (red) as a multiplicative result of prior expectations (orange) and the probability of incoming acoustic evidence under these assumptions ('likelihood', cyan), and their respective precision (=1/variance). Various neurobiological and psychological factors likely modulate these precisions and shape auditory perception. See main text for references. Meta-perceptual processes are directly influenced by the precision of the auditory percept itself (dashed distributions highlight the possible variation in precision) and allow for 'subjective', second-order reports or experiences of listening performance. Middle panel and inset: this second-order Bayesian process of 'meta-listening' can align or misalign with 'objective', first-order listening performance (i.e., exemplary individual shown). Measures of metacognitive sensitivity aim to quantify this agreement (e.g., [8,16]). See text for sources of overconfidence in listeners.

There is ongoing debate and not enough evidence yet on whether the first-order, perceptual evidence accumulation and its second-order, metacognitive evidence accumulation are separate processes (as implied schematically in Figure 1A). Readers are referred to [19] for an in-depth review of the nuanced suggestions of how confidence signals might be generated psychologically, in parallel or building on first-order perceptual decisions. This debate notwithstanding and as the next section will show, the listener's decision outcomes at a first-order perceptual level and a second-order metacognitive level can clearly dissociate.

Ideally, subjective confidence judgements in hearing will align with objective measures, such as neuroaudiological thresholds or auditory task performance (e.g., gap detection, speech in noise comprehension). An optimally efficient listener's subjective assessment would closely match their objective performance. Indeed, subjective ratings of speech intelligibility strongly correlate with objective measures of number of words comprehended ($r > 0.9$) [20,21].

However, such congruence of objective and subjective measures does not apply universally. For instance, the Speech, Spatial, and Qualities of Hearing (SSQ) questionnaire [22] is commonly used to measure hearing-aid benefit, but it correlates only poorly with objective auditory neural responses or audiometric tests [23,24]. A large online study of over 1100 healthy adults (age range 18–74) recently drove home this dissociation [25]: the older the listeners, the demonstrably worse they performed in an objective digits-in-noise test. However, the same listeners showed no discernible age-related change in their self-rated listening difficulties as per the SSQ.

These findings align with theories about individual differences in handling adverse listening situations [25] and highlight the dissociation between objective and subjective listening assessments. This dissociation is also a telling example of reduced metacognitive efficiency, as introduced earlier in this section. The sources of such reduced metacognitive efficiency in listening deserve future studies to explore more deeply their neurophysiological underpinnings, and we will turn to candidate neural mechanisms next.

Neural precision shapes auditory metacognition

Figure 1A illustrates the sensory and perceptual cascade of listening, starting from when the sensory epithelium of a human (or other non-human animal) receives environmental input, over the accumulation of sufficient sensory evidence to elicit a percept, perceptual decisions, and subsequent behavioural output. Also shown is the key idea of a 'metaperceptual' cascade that builds on top of (or in parallel to; for discussion see [19]) the neural and psychological perceptual cascade. As an outcome of such metaperceptual evidence accumulation, metacognitive statements like 'hearing my name must have been a false alarm' become possible.

Both the perceptual and metaperceptual cascades of Figure 1A must be implemented neurophysiologically at various levels of the neural processing hierarchy: from the auditory peripheral system to auditory cortex; in auditory–thalamic–striatal loops; and on to brain-wide, domain-general functional networks that closely interface with domain-specific auditory processing [26–28].

Intuitively, all these neural encoding stages carry various degrees of noise: neural encoding can turn out more or less precise, and neural precision can become undermined, for example, by the neurobiology of aging [29–31] (see discussion later). However, neural precision can also vary rapidly in a state-like fashion depending on the neuromodulatory dynamics of the neural circuit as a whole. For example, general anaesthesia can degrade frequency tuning in auditory cortex [32]; excitatory feedback from auditory cortex can shape precision in inferior colliculus [33]; or stimulation of layer-6 auditory cortical neurons can affect excitation–inhibition balance in a corticothalamic loop, where the precise stimulation timing differentially promotes tone detection versus tone discrimination [34]).

This flexibility and malleability of the precision of neural encodings is key when we aim to capture auditory perception and auditory metacognition in terms of **Bayesian models of perception** (see Box 1), where precision of distributions is a main feature that shapes perceptual outcomes. As illustrated in Figure 1, we can assume that the resulting metacognitive assessment of one's own auditory perception will vary depending on (i) precision of the prior; (ii) precision of the likelihood or sensory evidence; and (iii) the resulting precision of the posterior. Lastly, (iv) post-decisional factors [7] not shown explicitly here will matter (those post-decisional factors relate closely to the 'indirect drivers of meta-listening' discussed later).

The takeaway here for auditory studies of metacognition is the following: a Bayesian formalisation of the listener's percepts and their neural encodings can help us pinpoint sources of perceptual (im-)precision and predict a listener's metacognitive outcome; for example, their perceptual confidence or their experienced listening effort.

Candidate neural dynamics of auditory metacognition

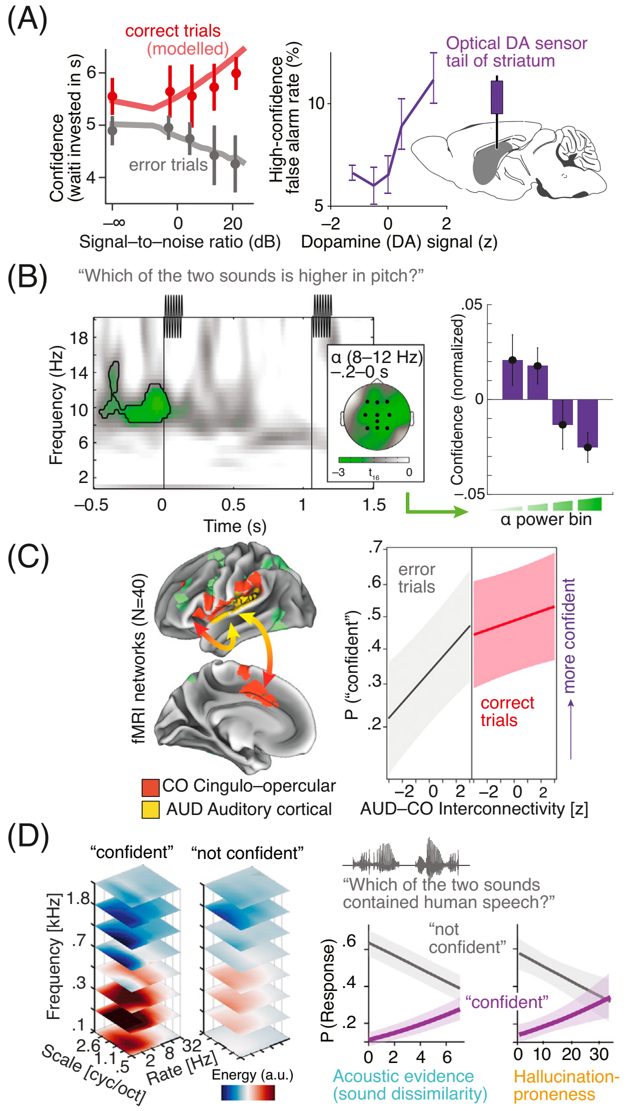
The recent literature in human and non-human auditory neuroscience has implicated various neural influences on metacognitive outcomes. I here aim to highlight the potential role of alpha (~8–12 Hz) oscillatory activity on the one hand and the effects of cortical desynchronisation on the other.

Alpha oscillatory power, dominating the **electroencephalography** (EEG) spectrum, modulates during listening tasks mainly in a posterior network of parietal and superior temporal cortex [35]. These modulations primarily reflect a listener's behavioural goals [36,37] and expectations [38,39] rather than features of the acoustic signal [40]. The generation of alpha **oscillations** has been linked to the neuromodulatory locus coeruleus–noradrenergic (LC–NE) system, which in turn affects the dopaminergic system [41–43].

A helpful conjecture is that alpha and beta (~18–25 Hz) oscillations set a regulatory milieu for neural transmission and encoding [44–46]. Supporting this assumption, alpha/beta oscillations are primarily efferent in nature and exhibit long-range connectivity [26]; they are relatively resistant to entrainment by external sensory rhythms [45]; and they have been associated with inward-oriented states of mind [47]. As such, alpha/beta oscillations are likely reflecting how auditory cortex populations represent prior predictions about expected auditory stimuli. An **electrocorticography** study using a Bayesian modelling framework and random tone sequences supported this idea [48,49].

In mice, a dopaminergic signal in the striatum scaled with animals' high detection confidence despite no acoustic evidence (i.e., in false-alarm trials), suggesting a dopamine-related increase in prior precision [6] (Figure 2A). This is reminiscent of human listeners' high-confidence decisions in a recent auditory forced-choice task, where confidence in a trial covaried with preceding alpha oscillatory power (Figure 2B; [50]). When listeners judged which of two (physically identical) tones had higher pitch, their confidence scaled linearly with pre-tone alpha power. This aligns well with findings from various sensory modalities that link alpha oscillatory dynamics to metacognitive report (e.g., self-rated attention, awareness, or visual confidence [51,52]). The results discussed here are broadly in line with the notion that catecholaminergic neuromodulation more generally, such as dopaminergic signalling in the striatum, affects prior precision (for review see [53]).

Unfortunately, however, these results by themselves do not allow us to disambiguate whether alpha oscillatory power changes metacognitive judgements by modulating prior precision, or



Trends in Neurosciences

Figure 2. How metacognitive abilities surface in neural and behavioural signatures of listening. (A) Mice, when detecting tones under varying signal-to-noise ratios (SNR), invest varying amounts of time waiting for a reward. This time investment is a proxy for confidence. A hallmark of metacognitive sensitivity is the interaction of SNR and accuracy: correct trials at good SNR elicit the highest confidence. High-confidence false detections covary with dopaminergic activity in the tail of the mouse striatum, underlining dopamine's role in perceptual and metaperceptual auditory decisions. Adapted with permission from [6]. (B) Lower centro-parietal alpha (8–12 Hz) oscillatory power in the human EEG prior to an auditory stimulus covaries with higher perceptual confidence reports: low alpha power may unspecifically promote neural excitability, boosting both perceptual and metaperceptual evidence. Adapted with permission from [50]. (C) The interconnectivity of human auditory (AUD, yellow) and cingulo-opercular (CO, red) networks in task-related haemodyamics predicts listeners' confidence. This effect is driven especially by incorrect trials; a result paralleling the high-confidence false alarms reported in mice in panel (A). In correct trials (70%), confidence reports are relatively higher. Adapted with permission from [75]. (D) When N=168 human listeners judge synthesised complex (i.e., non-speech) sounds, a reverse-correlation approach shows that trait-like hallucination-proneness (orange) and sensory evidence (i.e., more speech-typical frequency and modulation content; more acoustic distance between two sounds in a pair; cyan) elicit more 'confident' responses. Adapted with permission from [92]. Abbreviations: a.u., arbitrary unit; cyc/oct, cycles per octave; DA, dopamine; EEG, electroencephalography; fMRI, functional magnetic resonance imaging.

sensory-evidence precision, or both (Figure 1B; for further elaboration see, e.g., [52,54]). The identical-tones paradigm discussed earlier speaks to a mechanism via precise sensory-evidence encoding: trials with lower pre-stimulus alpha activity [50] or with higher pre-stimulus cortical desynchronisation [55] also showed stronger phase concentration in auditory cortical responses after stimulus onset.

We might conclude that states of low alpha oscillatory power boost neural excitability in sensory cortices and thus allow for increased perceptual confidence [50,52]. If this holds true, however, the often described oscillatory, excitability-dependent fluctuations in stimulus detectability [56–58] may be conflated with changes in metacognitive assessment, with higher confidence resulting for stimuli presented in high-excitability phases. To discern whether it is the first-order detection or the second-order confidence that adhere to oscillatory rhythmicity, researchers would need to also estimate behavioural confidence time courses. In fact, the underlying oscillatory properties of metacognition over time are underexplored [59] (particularly in audition; Ref. [60] hints at an oscillatory dissociation of bias from sensitivity). Numerous studies have identified neural oscillatory signatures of confidence or metacognitive efficiency as discussed earlier (e.g., [50,52,61,62]) – however, the potential waxing and waning over time of an internal confidence quantity has hardly been studied.

As an excitingly open field for future experimentation, alpha oscillations' role in shaping metalistening will need to be reconciled with the known importance of prefrontal cortex for metacognition (e.g., [61,63,64]). The generators of alpha-band activity critically changing metacognitive state in the listener might be generated in part also in prefrontal or anterior cingulate and insular areas, commensurate with contemporary suggestions of a cortical, layer-specific origin of alpha oscillations [65].

In sum, endogenous alpha-oscillatory power modulations likely impact both sensory-evidence precision and perceptual priors, which is largely commensurate with a more domain-general role of neural oscillations in organising and sharpening neural information transmission [66].

Direct versus indirect neural drivers of metacognition in the listening brain

One more mechanistic distinction deserves emphasis. As we have seen, the neural and neurophysiological factors that could cause or modulate varying auditory percept precision are numerous. The list of ones considered here in depth is certainly not exhaustive. When studying neural factors of metacognition in the auditory system, a helpful lead question can be: which are 'direct' neural driving pathways of metacognition in audition, and which are more 'indirect' neural pathways? Let us look at this distinction in more depth.

A neural process might affect metacognitive outcomes 'directly', that is, it could directly affect the precision of the auditory first-order percept itself. This could happen either by shaping the precision with which the afferent sensory evidence becomes represented; or by shaping efferent signals carrying prior expectations at some level of the auditory hierarchy; or both. The neural mechanisms discussed earlier, especially alpha oscillatory changes, would fall into this class.

By contrast, 'indirect' or amodal neural drivers of metacognition would be those neural dynamics that affect confidence or metacognitive aspects of listening without affecting the auditory percept, its priors, or its sensory evidence *per se*. The locus of such indirect drivers would thus be post-perceptual, and might draw, as suggested but not well-established [19], separate accumulation of metacognitive evidence (see Figure 1A).

By this definition, these 'indirect' neural mechanisms would also be less auditory-specific and more supramodal in nature. Candidate mechanisms to exert such 'indirect' supramodal effects on a perceiver's metacognitive state or trait (e.g., [67,68]) are, for example, the coarse neuromodulatory changes (as expressed, e.g., in pupil-linked arousal and the LC–NE system; capturing well the metacognitive traces of listening effort, e.g., [69,70]). Likewise, interoceptive signals from the cardiovascular system (see, e.g., [71]) or large-scale network changes in excitation–inhibition balance [72–74] and interconnectivity [26,75–77] could affect metacognitive outcomes in audition in such a more indirect, post-perceptual fashion.

One obvious way of separating direct from indirect neural influences when analysing metacognitive measures is by statistically controlling for effects at the auditory–perceptual or sensory–encoding level. (Essentially, this approximates statistically testing for mediation.)

A promising approach in metacognition research has thus been to equate first-level perceptual performance through tight control of stimulus and task features. Such first-order, perceptual matching helps isolate those neural mechanisms that drive differences in second-order, metacognitive outcome (in vision see, e.g., [78,79]). Doing so in audition, it can be demonstrated that the impact of pre-stimulus brain state on metrics of confidence is not being fully mediated through encoding changes of the auditory percept *per se* [50,55,80].

Complicating this picture somewhat, however, this matching does not rule out that auditory cortical areas might be nevertheless involved in post-perceptual, metacognitive processes: for example, in a recent study that carefully matched listeners' first-order perceptual task performance at 70%, the interconnectivity dynamics within an auditory–cingulo–opercular control network predicted listeners' confidence reports (Figure 2C; [75]). This result challenges a strict division of labour between direct drivers (i.e., modality-specific and perceptual-precision-mediated) and indirect drivers (i.e., amodal and less perceptual-precision-bound) of metacognition. Here, more research is clearly needed that will simultaneously quantify the local neural dynamics in auditory cortex and brain-wide networks.

How does metacognition relate to aberrant auditory perception?

As introduced earlier, a number of neurological and neuropsychiatric conditions exhibit close links to the auditory system and can affect metacognition. Important examples are dementia, psychosis, and tinnitus.

Dementia has been characterised in terms of a loss in metacognitive abilities (e.g., [81,82]). Notably, amongst the potentially modifiable risk factors for dementia, hearing loss is currently considered the largest one [12,13]. Sensorineural hearing loss as a common, age-related condition is characterised by a loss of fidelity in sensory encoding of sound [83]: more acoustic energy is needed to transduce sound; the loss in temporal precision introduces spectral and temporal smearing; and the dynamic range of sound that can be neurally encoded is severely reduced.

In Bayesian-perceptual terms, this loss of fidelity can be best thought of as a loss in the precision of sensory evidence (i.e., a marked widening of the likelihood function). Age-related hearing loss should thus decrease the resulting perceptual precision. Strikingly, however, older adults are not only most prone to hearing loss, but they show signs of overconfidence in challenging hearing situations [25,84,85]. This overconfidence has been demonstrated to trace back to an increased reliance on semantic context (as a form of perceptual prior [84]).

Higher confidence in older adults thus showcases the impact of prior precision on resulting confidence: if it were for the age-related loss of sensory precision alone, confidence should decrease, not increase with age (cf. Box 1; see [86] for modelling such seemingly paradoxical overconfidence in vision). It remains speculative, at this stage, whether age-related changes in metacognitive abilities and concomitant changes in meta-listening will allow us to better differentiate trajectories of healthy ageing from their pathological variants.

Auditory verbal hallucinations are another striking example of the delicate balance of prior assumptions with the precision of auditory encoding. Moreover, when listeners experience acoustic events that are not physically present, they sometimes do so with high levels of confidence. In fact, a cognitive hallmark of people with schizophrenia (who also commonly experience auditory verbal hallucinations) is a tendency to 'jump to conclusions' [87]. This includes committing prematurely to perceptual decisions, in tune with a notion of overconfidence. Recent experiments have plausibly linked such premature commitment to a decision and a discarding of disconfirming evidence to reduced precision in sensory evidence (e.g., through ketamine-induced changes in excitation–inhibition balance [80,88]; but see [89]).

Importantly, auditory verbal hallucinations are not a perceptual phenomenon limited to states of psychosis. Tendencies to hear voices or have unusual auditory perceptions occur, in gradations, in about 1 in 10 healthy individuals [90] and are part of the personality dimension of schizotypy [91]. A recent psychophysical modelling approach demonstrated that the degree of individual self-reported, trait-like hallucination-proneness comes with a marked under-weighting of speech-typical sensory evidence (i.e., low acoustic frequencies, slow temporal modulation rates) when judging sounds to be speech or not [92]. All participants did use sensory evidence to inform their confidence judgements (i.e., more acoustically dissimilar sound pairs prompted more confident responses in general). However, more hallucination-prone individuals gave overall more ‘confident’ responses (up to 30%, in an inherently ambiguous and challenging perceptual task) (Figure 2D).

Finally, tinnitus is a poorly understood aberration of auditory function and a source of distress to many. Tinnitus represents an auditory percept, held with high confidence, in the absence of a causative acoustic event. In other words, it is a ‘hallucination-like percept’ [6], for which established non-human animal models exist [93]. It has been put forward that chronic tinnitus reflects an imbalance in combining imprecise sensory evidence (caused by hearing loss, which most often accompanies tinnitus [94]) with an altered weighing of perceptual priors [95]. The perceptual changes in auditory verbal hallucinations highlighted earlier beg the question of whether individuals with tinnitus will also exhibit trait-like metacognitive or metaperceptual changes in the direction of overconfident sound representation.

In sum, the clinically relevant phenomena of hearing loss, auditory hallucinations, and tinnitus can be seen as valuable model systems when understanding metacognition more generally. They help us understand how auditory neurobiology, auditory perception, and metacognitive sensitivity are linked. Hearing and listening also allow us to pinpoint the proposed, multi-faceted sources of metacognitive inefficiency [96]: for example, noisy sensory input (from which confidence is in part derived) can be controlled easily and thus can help to experimentally isolate mechanisms of confidence computation *per se*. Not least, understanding better the metacognition of human listening behaviour will provide viable inroads to improving *in silico* implementations of perception and metacognition (see Box 2).

Concluding remarks and future perspectives

The metacognition of listening and its implementations in the listening brain is a nascent field. As I have aimed to show here, rooting auditory research in a framework of Bayesian perception and drawing from the rich and growing literature of metacognitive modelling permits us to clearly delineate the neural and psychological sources of (im-)precision that shape auditory perception and that affect the metacognitive assessment of our percepts.

Better understanding of the metacognition of listening will also allow us to better target subjective aspects of listening and communication dissatisfaction. It may further allow identifying when and how such subjective experiences deviate from objective markers of neural auditory function (see Outstanding questions). Not least, improving the ability to earlier detect individuals at risk for psychosis or dementia based on changes in auditory perception and altered metacognitive sensitivity is a hopeful prospect.

Outstanding questions

At which stage of auditory processing do first-order and second-order auditory perceptual decisions dissociate? Which are the necessary neural pathways that allow metacognitive sensitivity (i.e., type-II or second-order performance) to be ‘read out’ from a listener’s actual (i.e., type-I or first-order) task performance? This is an unsolved question in vision-based metacognition research as well.

Will knowing the ‘confidence state’ of a system, biological or artificial, allow us to predict its future communication behaviour? This applies to short time scale (e.g., in the next trial) and to longer time scales (e.g., average changes in future response behaviour).

How should *in silico* implementations of listening confidence look? How can biologically inspired neural network architectures that process sound (‘machine listening’) express a degree of rudimentary metacognitive efficiency? Which architectural features of an artificial neural network best match the confident-listening mechanisms identified in humans?

Acknowledgements

Ideas developed and results mentioned in this manuscript have profited greatly from a European Research Council (ERC) Consolidator grant (ERC-CoG-2014 646696 'AUDADAPT'). The author is indebted to many former and current members of his research team 'Auditory Cognition' for comments and reflections on earlier versions of the arguments collected here. Merle Schuckart kindly provided drawings for Figure 1. Not least, the author is grateful for constructive criticism and important suggestions by two reviewing colleagues.

Declaration of interests

The author declares no conflicts of interest.

Declaration of Generative AI and AI-assisted technologies in the writing process

The Generative AI ChatGPT (v4o; a generative pre-trained transformer for next-word prediction) by OpenAI. Inc. had been consulted to obtain alternative, more concise edits for certain single paragraphs at the draft editing stage. The author reviewed and edited all AI-generated suggestions as needed and takes full responsibility for the content of the publication.

References

1. Flavell, J.H. (1979) Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *Am. Psychol.* 34, 906–911
2. Johnson, D.D. and Fowler, J.H. (2011) The evolution of overconfidence. *Nature* 477, 317–320
3. Peirce, C.S. *et al.* (1885) *On small differences of sensation (Memoirs of the National Academy of Sciences, Vol. 3, 5th memoir)*, Government Printing Office
4. Rafiei, F. *et al.* (2024) The neural network RTNet exhibits the signatures of human perceptual decision-making. *Nat. Hum. Behav.* https://doi.org/10.1038/s41562-024-01914-8
5. Mamassian, P. (2016) Visual confidence. *Annu. Rev. Vis. Sci.* 2, 459–481
6. Schmack, K. *et al.* (2021) Striatal dopamine mediates hallucination-like perception in mice. *Science* 372. https://doi.org/10.1126/science.abf4740
7. Fleming, S.M. (2024) Metacognition and confidence: A review and synthesis. *Annu. Rev. Psychol.* 75, 241–268
8. Fleming, S.M. and Lau, H.C. (2014) How to measure metacognition. *Front. Hum. Neurosci.* 8, 443
9. de Gardelle, V. *et al.* (2016) Confidence as a common currency between vision and audition. *PLoS One* 11, e0147901
10. Tang, T. *et al.* (2024) Behavioral and neural measures of confidence using a novel auditory pitch identification task. *PLoS One* 19, e0299784
11. Peters, M.A.K. (2022) Towards characterizing the canonical computations generating phenomenal experience. *Neurosci. Biobehav. Rev.* 142, 104903
12. Livingston, G. *et al.* (2020) Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* 396, 413–446
13. Griffiths, T.D. *et al.* (2020) How can hearing loss cause dementia? *Neuron* 108, 401–412
14. Linszen, M.M. *et al.* (2016) Increased risk of psychosis in patients with hearing impairment: Review and meta-analyses. *Neurosci. Biobehav. Rev.* 62, 1–20
15. Galvin, S.J. *et al.* (2003) Type 2 tasks in the theory of signal detectability: discrimination between correct and incorrect decisions. *Psychon. Bull. Rev.* 10, 843–876
16. Maniscalco, B. and Lau, H. (2012) A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Conscious. Cogn.* 21, 422–430
17. Emmerich, D.S. *et al.* (1972) Response latency, confidence, and ROCs in auditory signal detection. *Percept. Psychophys.* 11, 65–72
18. Rahnev, D. (2023) Measuring metacognition: A comprehensive assessment of current methods. *Preprint (PsyRxiv)* https://doi.org/10.31234/osf.io/waz9h
19. Shekhar, M. and Rahnev, D. (2024) How do humans give confidence? A comprehensive comparison of process models of perceptual metacognition. *J. Exp. Psychol. Gen.* 153, 656–688
20. Obleser, J. *et al.* (2008) Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *J. Neurosci.* 28, 8116–8123
21. Davis, M.H. and Johnsrude, I.S. (2003) Hierarchical processing in spoken language comprehension. *J. Neurosci.* 23, 3423–3431
22. Gatehouse, S. and Noble, W. (2004) The Speech, Spatial, and Qualities of Hearing Scale (SSQ). *Int. J. Audiol.* 43, 85–99
23. Tune, S. *et al.* (2018) Probing the limits of alpha power lateralisation as a neural marker of selective attention in middle-aged and older listeners. *Eur. J. Neurosci.* 48, 2537–2550
24. Herrmann, B. *et al.* (2016) Altered temporal dynamics of neural adaptation in the aging human auditory cortex. *Neurobiol. Aging* 45, 10–22
25. Wostmann, M. *et al.* (2021) Personality captures dissociations of subjective versus objective hearing in noise. *R. Soc. Open Sci.* 8, 210881
26. Alavash, M. *et al.* (2017) Large-scale network dynamics of beta-band oscillations underlie auditory perceptual decision-making. *Netw. Neurosci.* 1, 166–191
27. Alavash, M. *et al.* (2019) Modular reconfiguration of an auditory control brain network supports adaptive listening behavior. *Proc. Natl. Acad. Sci. U. S. A.* 116, 660–669
28. Hartwigsen, G. (2018) Flexible redistribution in cognitive networks. *Trends Cogn. Sci.* 22, 687–698
29. Erb, J. *et al.* (2020) Temporal selectivity declines in the aging human auditory cortex. *eLife* 9. https://doi.org/10.7554/eLife.55300
30. Recanzone, G. (2018) The effects of aging on auditory cortical function. *Hear. Res.* 366, 99–105
31. Ouda, L. *et al.* (2015) Age-related changes in the central auditory system. *Cell Tissue Res.* 361, 337–358
32. Middlebrooks, J.C. *et al.* (2023) High spectral and temporal acuity in primary auditory cortex of awake cats. *J. Assoc. Res. Otolaryngol.* 24, 197–215
33. Blackwell, J.M. *et al.* (2020) Auditory cortex shapes sound responses in the inferior colliculus. *eLife* 9. https://doi.org/10.7554/eLife.51890
34. Guo, W. *et al.* (2017) A corticothalamic circuit for dynamic switching between feature detection and discrimination. *Neuron* 95, 180–194 e185
35. Wostmann, M. *et al.* (2020) Does closing the eyes enhance auditory attention? Eye closure increases attentional alpha-power modulation but not listening performance. *J. Cogn. Neurosci.* 32, 212–225
36. Wostmann, M. *et al.* (2016) Spatiotemporal dynamics of auditory attention synchronize with speech. *Proc. Natl. Acad. Sci. U. S. A.* 113, 3873–3878
37. Wostmann, M. *et al.* (2021) Orienting auditory attention in time: Lateralized alpha power reflects spatio-temporal filtering. *NeuroImage* 228, 117711

38. Muller, N. *et al.* (2013) You can't stop the music: reduced auditory alpha power and coupling between auditory and memory regions facilitate the illusory perception of music during noise. *NeuroImage* 79, 383–393
39. Jensen, O. *et al.* (2012) An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends Cogn. Sci.* 16, 200–206
40. Wostmann, M. *et al.* (2017) The human neural alpha response to speech is a proxy of attentional control. *Cereb. Cortex* 27, 3307–3317
41. Dahl, M.J. *et al.* (2020) Noradrenergic responsiveness supports selective attention across the adult lifespan. *J. Neurosci.* 40, 4372–4390
42. Pfeffer, T. *et al.* (2018) Catecholamines alter the intrinsic variability of cortical population activity and perception. *PLoS Biol.* 16, e2003453
43. Ballesteros, J.J. *et al.* (2020) Neural signatures of alpha2-Adrenergic agonist-induced unconsciousness and awakening by antagonist. *eLife* 9. https://doi.org/10.7554/eLife.57670
44. Jensen, O. and Mazaheri, A. (2010) Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front. Hum. Neurosci.* 4, 12
45. Lakatos, P. *et al.* (2016) Global dynamics of selective attention and its lapses in primary auditory cortex. *Nat. Neurosci.* https://doi.org/10.1038/nn.4386
46. Kayser, C. *et al.* (2015) Rhythmic auditory cortex activity at multiple timescales shapes stimulus-response gain and background firing. *J. Neurosci.* 35, 7750–7762
47. Palva, S. and Palva, J.M. (2011) Functional roles of alpha-band phase synchronization in local and large-scale cortical networks. *Front. Psychol.* 2, 204
48. Sedley, W. *et al.* (2016) Neural signatures of perceptual inference. *eLife* 5, e11476
49. Obleser, J. (2016) Tell me something I don't know. *eLife* 5, e15853
50. Wostmann, M. *et al.* (2019) Prestimulus neural alpha power predicts confidence in discriminating identical auditory stimuli. *Eur. J. Neurosci.* 49, 94–105
51. Benwell, C.S.Y. *et al.* (2017) Prestimulus EEG power predicts conscious awareness but not objective visual performance. *eneuro* 4. https://doi.org/10.1523/ENEURO.0182-17.2017
52. Samaha, J. *et al.* (2017) Prestimulus alpha-band power biases visual discrimination confidence, but not accuracy. *Conscious. Cogn.* 54, 47–55
53. Ngo, H.V. *et al.* (2023) Circadian rhythms in auditory hallucinations and psychosis. *Acta Physiol. (Oxf)* 237, e13944
54. Samaha, J. *et al.* (2020) Spontaneous brain oscillations and perceptual decision-making. *Trends Cogn. Sci.* 24, 639–653
55. Waschke, L. *et al.* (2017) States and traits of neural irregularity in the age-varying human brain. *Sci. Rep.* 7, 17381
56. Strauss, A. *et al.* (2015) Alpha phase determines successful lexical decision in noise. *J. Neurosci.* 35, 3256–3262
57. Henry, M.J. and Obleser, J. (2012) Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proc. Natl. Acad. Sci. U. S. A.* 109, 20095–20100
58. Ng, B.S.W. *et al.* (2012) A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. *J. Neurosci.* 32, 12268–12276
59. Trajkovic, J. *et al.* (2023) Two oscillatory correlates of attention control in the alpha-band with distinct consequences on perceptual gain and metacognition. *J. Neurosci.* 43, 3548–3556
60. Ho, H.T. *et al.* (2017) Auditory sensitivity and decision criteria oscillate at different frequencies separately for the two ears. *Curr. Biol.* 27, 3643–3649 e3643
61. Wokke, M.E. *et al.* (2017) Sure I'm sure: prefrontal oscillations support metacognitive monitoring of decision making. *J. Neurosci.* 37, 781–789
62. Faivre, N. *et al.* (2018) Behavioral, modeling, and electrophysiological evidence for supramodality in human metacognition. *J. Neurosci.* 38, 263–277
63. Fleming, S.M. and Dolan, R.J. (2012) The neural basis of metacognitive ability. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 367, 1338–1349
64. De Martino, B. *et al.* (2013) Confidence in value-based choice. *Nat. Neurosci.* 16, 105–110
65. Halgren, M. *et al.* (2019) The generation and propagation of the human alpha rhythm. *Proc. Natl. Acad. Sci. U. S. A.* 116, 23772–23782
66. Kayser, C. *et al.* (2009) Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron* 61, 597–608
67. Lempert, K.M. *et al.* (2015) Relating pupil dilation and metacognitive confidence during auditory decision-making. *PLoS One* 10, e0126588
68. Colizoli, O. *et al.* (2018) Task-evoked pupil responses reflect internal belief states. *Sci. Rep.* 8, 13702
69. Zekveld, A.A. *et al.* (2018) The pupil dilation response to auditory stimuli: current state of knowledge. *Trends Hear.* 22, 2331216518777174
70. Kraus, F. *et al.* (2023) Pupil size sensitivity to listening demand depends on motivational state. *eneuro* 10. https://doi.org/10.1523/ENEURO.0288-23.2023
71. Hauser, T.U. *et al.* (2017) Noradrenaline blockade specifically enhances metacognitive performance. *eLife* 6. https://doi.org/10.7554/eLife.24901
72. Zhang, S. *et al.* (2024) In vivo whole-cortex marker of excitation-inhibition ratio indexes cortical maturation and cognitive ability in youth. *Proc. Natl. Acad. Sci. U. S. A.* 121, e2318641121
73. Atiya, N.A.A. *et al.* (2021) Explaining distortions in metacognition with an attractor network model of decision uncertainty. *PLoS Comput. Biol.* 17, e1009201
74. Harris, K.D. and Thiele, A. (2011) Cortical state and attention. *Nat. Rev. Neurosci.* 12, 509–523
75. Alavash, M. and Obleser, J. (2024) Brain network interconnectivity dynamics explain metacognitive differences in listening behavior. *J. Neurosci.* https://doi.org/10.1523/JNEUROSCI.2322-23.2024
76. Alavash, M. *et al.* (2021) Dynamic large-scale connectivity of intrinsic cortical oscillations supports adaptive listening in challenging conditions. *PLoS Biol.* 19, e3001410
77. Shine, J.M. *et al.* (2019) Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nat. Neurosci.* 22, 289–296
78. Samaha, J. *et al.* (2016) Dissociating perceptual confidence from discrimination accuracy reveals no influence of metacognitive awareness on working memory. *Front. Psychol.* 7, 851
79. Skewes, J. *et al.* (2021) Awareness and confidence in perceptual decision-making. *Brain Multiphys.* 2. https://doi.org/10.1016/j.brain.2021.100030
80. Waschke, L. *et al.* (2019) Local cortical desynchronization and pupil-linked arousal differentially shape brain states for optimal sensory performance. *eLife*, 8 https://doi.org/10.7554/eLife.51501
81. Cosentino, S. and Stern, Y. (2005) Metacognitive theory and assessment in dementia: do we recognize our areas of weakness? *J. Int. Neuropsychol. Soc.* 11, 910–919
82. Vannini, P. *et al.* (2019) Decreased meta-memory is associated with early tauopathy in cognitively unimpaired older adults. *Neuroimage Clin.* 24, 102097
83. Gates, G.A. and Mills, J.H. (2005) Presbycusis. *Lancet* 366, 1111–1120
84. Rogers, C.S. (2017) Semantic priming, not repetition priming, is to blame for false hearing. *Psychon. Bull. Rev.* 24, 1194–1204
85. Rogers, C.S. *et al.* (2012) Frequent false hearing by older adults: the role of age differences in metacognition. *Psychol. Aging* 27, 33–45
86. Rahnev, D. *et al.* (2011) Attention induces conservative subjective biases in visual perception. *Nat. Neurosci.* 14, 1513–1515
87. Evans, S.L. *et al.* (2015) Jumping to conclusions in schizophrenia. *Neuropsychiatr. Dis. Treat.* 11, 1615–1624
88. Salvador, A. *et al.* (2022) Premature commitment to uncertain decisions during human NMDA receptor hypofunction. *Nat. Commun.* 13, 338
89. Toso, A. *et al.* (2024) 40 Hz Steady-state response in human auditory cortex is shaped by gabaergic neuronal inhibition. *J. Neurosci.* 44. https://doi.org/10.1523/JNEUROSCI.2029-23.2024
90. de Leede-Smith, S. and Barkus, E. (2013) A comprehensive review of auditory verbal hallucinations: lifetime

prevalence, correlates and mechanisms in healthy and clinical individuals. *Front. Hum. Neurosci.* 7. https://doi.org/10.3389/fnhum.2013.00367

91. Callaway, D.A. *et al.* (2014) Schizotypal Personality Questionnaire-Brief Revised: psychometric replication and extension. *Pers. Disord.* 5, 32–38
92. Erb, J. *et al.* (2020) Aberrant perceptual judgments on speech-relevant acoustic features in hallucination-prone individuals. *Schizophr. Bull. Open* 1. https://doi.org/10.1093/schizbullopen/sgaa059
93. von der Behrens, W. (2014) Animal models of subjective tinnitus. *Neural Plast.* 2014, 741452
94. Weisz, N. *et al.* (2006) High-frequency tinnitus without hearing loss does not mean absence of deafferentation. *Hear. Res.* 222, 108–114
95. Reisinger, L. *et al.* (2024) Aberrant auditory prediction patterns robustly characterize tinnitus. *eLife* RP99757. https://doi.org/10.7554/eLife.99757.3
96. Shekhar, M. and Rahnev, D. (2021) Sources of metacognitive inefficiency. *Trends Cogn. Sci.* 25, 12–23
97. Balsdon, T. *et al.* (2020) Confidence controls perceptual evidence accumulation. *Nat. Commun.* 11, 1753
98. Kagan, B.J. *et al.* (2024) Toward a nomenclature consensus for diverse intelligent systems: Call for collaboration. *Innovation (Camb)* 5, 100658
99. Rahwan, I. *et al.* (2019) Machine behaviour. *Nature* 568, 477–486
100. openAI (2024) *ChatGPT (GPT-4o)*
101. Barredo Arrieta, A. *et al.* (2020) Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* 58, 82–115
102. Petersen, E. *et al.* (2022) Responsible and regulatory conform machine learning for medicine: A survey of technical challenges and solutions. *IEEE Access* 10, 58375–58418
103. Schmitt, L.M. *et al.* (2021) Predicting speech from a cortical hierarchy of event-based time scales. *Sci. Adv.* 7, eabi6070
104. Kell, A.J.E. *et al.* (2018) A task-optimized neural network replicates human auditory behavior, predicts brain responses, and reveals a cortical processing hierarchy. *Neuron* 98, 630–644 e616
105. Caucheteux, C. *et al.* (2023) Evidence of a predictive coding hierarchy in the human brain listening to speech. *Nat. Hum. Behav.* 7, 430–441
106. Webb, T.W. *et al.* (2021) A task-optimized neural network model of decision confidence. *Proc. Annu. Meet. Cogn. Sci. Soc.* 43, https://escholarship.org/uc/item/8cv4b2qp