OPINION

Control of mental activities by internal models in the cerebellum

Masao Ito

Abstract | The intricate neuronal circuitry of the cerebellum is thought to encode internal models that reproduce the dynamic properties of body parts. These models are essential for controlling the movement of these body parts: they allow the brain to precisely control the movement without the need for sensory feedback. It is thought that the cerebellum might also encode internal models that reproduce the essential properties of mental representations in the cerebral cortex. This hypothesis suggests a possible mechanism by which intuition and implicit thought might function and explains some of the symptoms that are exhibited by psychiatric patients. This article examines the conceptual bases and experimental evidence for this hypothesis.

In its broadest sense, the term 'internal model' applies to any neural representation of the external world[1]. However, in this article I consider the specific concept of internal models in the cerebellum, which has developed from modern theories of movement control[2–7] (BOX 1). According to this concept, the cerebellum forms (through a learning process) an internal model that reproduces either the dynamics of a body part (in the case of a forward model (FIG. 1a)) or the inverse of those dynamics (in the case of an inverse model (FIG. 1b)). This internal model is formed and adjusted as a movement is repeated. Ultimately the internal model helps the brain to perform the movement precisely, without the need to refer to feedback from the moving body part. This explains how we become able to move skillfully after repeated practice. The concept of the control of movement by internal models is supported by computational modelling and experimental neuroscience, and has been realized in the creation of a robot that can learn skilled movements through the use of similar principles[8].

The possible involvement of the cerebellum in non-motor mental functions such as language had been suggested occasionally in the past[9], but it only became the subject of systematic consideration at the beginning of the 1990s. Dow and colleagues[10,11] emphasized that the lateral part of the cerebellar hemispheres, which was the most recent part to evolve, developed in parallel with the cerebral association cortex, which generates mental activities. Others[12] revealed the existence of significant interconnections between the cerebellar hemispheres and both the prefrontal (Brodmann's area 9) and the parietal (area 5) association cortices, by recording evoked potentials in monkeys. Clinical symptoms arising from cerebellar lesions have also suggested the involvement of the cerebellum in mental functions. Schmahmann[13] suggested that the cerebellum regulates the speed, consistency and appropriateness of cognitive processes, and interpreted impairment of these aspects in cerebellar patients as an indication that the patients were unable to precisely control their thoughts. It was at approximately this time that I first formulated the internal-model hypothesis for mental activities[14–17], based on an analogy between the control of the movement of body parts and the manipulation of mental representations. The hypothesis requires the concerted activity of a large-scale network that involves the prefrontal cortex, the temporo–parietal cortex and the cerebellar hemispheres (FIG. 1c,d).

Since these initial efforts nearly two decades ago, our understanding of the shared features of motor control and mental manipulation has advanced and given us insight into the mechanisms of implicit, intuitive and creative[18,19] thoughts and even social interactions[20]. Furthermore, a range of relevant neuroscientific data that support the hypothesis has recently accumulated. Anatomical pathways that interconnect the cerebral cortex and the cerebellum have also been re-analysed[21–24]. These findings have been reviewed and interpreted elsewhere and therefore do not form the focus of this article[25]. In addition, an increasing amount of neuroimaging data has demonstrated co-activation of the cerebellum with the prefrontal cortex and the temporo–parietal cortex in various types of mental tasks[26–39], as predicted by the original hypothesis. There have also been a number of clinical observations that have shown that cerebellar lesions cause mental disorders[40–43] and that various mental disorders accompany cerebellar dysfunction[44–48].

The aim of this Perspective is to elaborate on previous descriptions of the internal-model hypothesis for the control of mental activities, focusing on these more recent findings. I discuss our knowledge of the components that comprise the internal-model control system and show how these can be combined to form a putative thought system. Finally, I examine the hypothesis in light of the recent wealth of neuroimaging and clinicopathological data and discuss the future directions of the field. Based on the conceptual and experimental evidence described above, in this article I aim to establish that internal-model control through cerebro–cerebellar connectivity is a unique component of human intelligence. In this manner, I intend to pave the way for new approaches to understanding the neuroscience of mental mechanisms.

Components of internal-model control

When we think, we manipulate something in our mind. The initial assumption that the internal-model hypothesis makes is that this thought process has neural substrates in the brain, to which control-system principles (BOX 1) apply. The prefrontal

Box 1 | **Control-system scheme for movement**

Imagine a setup in which a computer can control a robot arm in response to inputs that it receives from an operator. According to control theories, during this process the controller (the computer) receives instruction from an instructor (the operator) and accordingly generates command signals that then drive a controlled object (the robot arm) to behave as instructed (FIG. 1a). The controller can be (but is not necessarily) fed with the output of the controlled object (external feedback). This example is a good illustration of the processes that are at work in the control of diverse types of human movement, such as reflexes and compound movements (for example, locomotion and saccades)[52]. In voluntary movement, for example, the motor cortex acts as a controller and a body part (which includes the lower motor centres in the spinal cord) constitutes a controlled object. Instruction signals come from the premotor and supplementary motor cortices and the anterior cingulate gyrus[18].

To form an adaptive control system — that is, a control system that can learn[105] — another element is required: an internal model. Two types of internal model have been proposed. A 'forward' model reproduces the dynamics of a controlled object, whereas an 'inverse' model reproduces a reciprocal of those dynamics. A forward model provides an internal feedback that can replace the external feedback from the controlled object (FIG. 2B). An inverse model, by contrast, provides a controller that does not receive feedback (a feed-forward controller), which can replace the original controller (FIG. 2C). These two types of internal model might operate in combination[6].

cortex is assumed to have the role of the controller, whereas neurons in the temporo–parietal cortex encode the controlled object. The cerebellar hemispheres are assumed to provide the internal model (which can be forward or inverse) (FIG. 1c,d). The instructor, the sensory/perception system (FIG. 2B,C) and the attentional system (see below) are additional components. In the following subsections, recent evidence supporting these functional assignments is described.

The prefrontal cortex as the controller. The prefrontal cortex is a collection of interconnected neocortical areas in the frontal lobe (encompassing everything from the rostral end of the frontal lobe to the premotor areas). It is believed to perform executive functions that are important for the conscious control of thought and action in accordance with internal goals[49]. Of the three major regions of the prefrontal cortex (dorsolateral, medial and orbitofrontal), it is the dorsolateral prefrontal cortex that controls cognitive functions such as abstract reasoning and problem solving. The orbitofrontal and related medial regions are associated with the control of affective and motivational functions[50]. All three regions seem to work together as an interactive functional system.

The symptoms that a lesion of the prefrontal cortex induces in humans — that is, impulsive, inappropriate or disorganized behaviour — are indicative of its executive functions. More concretely, the prefrontal cortex is crucial in situations in which the links between sensory inputs, thoughts and actions are either relatively weakly established or rapidly changing[49]. Typically, playing games such as the Stroop Task[36], the Wisconsin Card-Sorting Task (WCST)[35] or the Tower of London Task[28] induces activation in the prefrontal cortex in neuroimaging studies, and the ability to complete these tasks is sensitive to damage of the prefrontal cortex (see below).

What are the neural mechanisms that the prefrontal cortex uses when it manipulates a controlled object? It is thought that a set of neurons in the prefrontal cortex generates and maintains patterns of activity that represent goals and the means by which to achieve them[49]. These neurons provide signals that influence activity throughout much of the brain and affect many neural processes, including those that are responsible for sensation and/or perception, response execution, memory retrieval and emotional evaluation. It seems that these prefrontal signals collectively constitute command signals that manipulate activity in the neurons that encode the controlled object.

The prefrontal cortex is also involved with working memory, which briefly holds information online and underlies conscious thought processes[51]. Working memory can be tested in monkeys using delay tasks, in which a stimulus is briefly presented and then removed from view for a brief period (10–20 seconds). The monkey is then required to distinguish the stimulus from distractors. The working-memory system proposed by Baddeley and Hitch[52] consists of a central executive-control system, which is located in the dorsolateral prefrontal cortex, and two slave systems: a visuospatial sketchpad, which is responsible for temporal storage of visuospatial material[53], and a phonological loop, which temporarily stores phonology-based material[54]. The working-memory system functions only when it is driven by attention. It is assumed to have various roles, including mental manipulation of material held in the slave systems[55]. These roles are therefore compatible with the assignment of the controller role to the prefrontal cortex.

Mental models in the temporo–parietal cortex. What is the controlled object that is manipulated by the prefrontal cortex when we think? A 'mental model' is a psychological substrate of a real or imaginary situation. This concept of mental models was proposed by Craik[56] and elaborated upon by Johnson-Laird[57] and others. According to their definition, a mental model is a small-scale model of reality that is used to reason, to explain current events and to anticipate future events. Mental models can be constructed from perception, imagination or the comprehension of discourse. Mental models underlie our ability to imagine visual scenes, but they can also be abstract and represent situations that cannot be visualized. The concept of a mental model is a psychological notion, but it finds an objective counterpart in our brain[58]. The internal-model hypothesis of mental processes assumes that the prefrontal controller constructs and manipulates the neural substrates of a mental model, with the mental model acting as a controlled object.

In the neocortex, sensory information is first analysed in sensory areas and then integrated to form a perception somewhere in the broad temporo–parietal association areas. For example, this integration is reflected in the neuronal activity that is induced in an inferotemporal association area when a monkey views an object[59,60]. After a certain degree of integration has occurred, sensory information is transformed into a mental representation. Such a mental representation of a visual pattern has been revealed in neurons located in the anterior part of the inferotemporal area[61]. In binocular rivalry, a different visual stimulus is presented to each eye. The subject does not perceive a combination of the two stimuli; instead, perception alternates between the two stimuli. In this case, the activity of the neurons in the inferotemporal area reflects what is being perceived, not what is being sensed through the retina. This was ingeniously demonstrated in a monkey that had been trained to indicate on which side of its body a stimulus was perceived, by pulling a lever on the same side as the stimulus[61]. During performance of the task, responses to rival stimuli were recorded from neurons

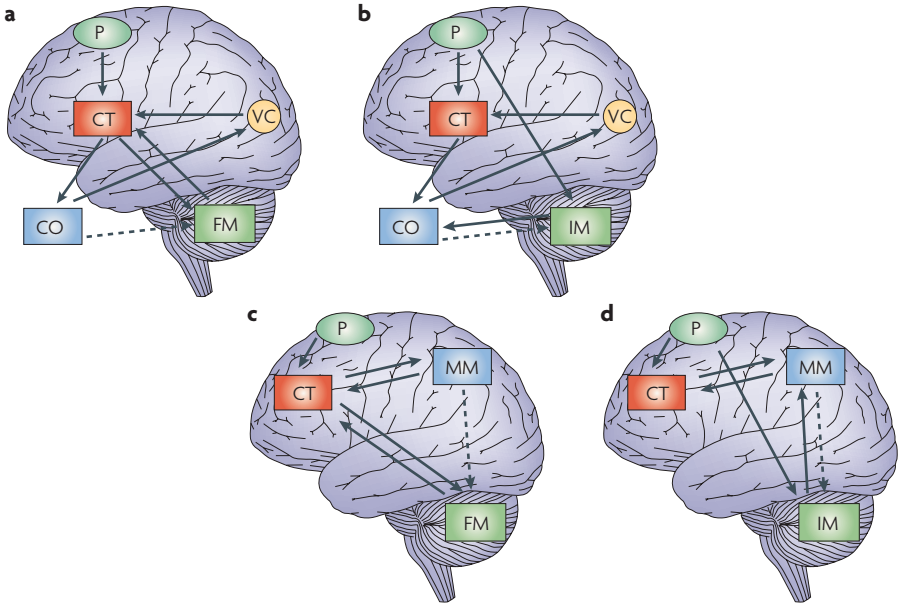

Figure 1 | **Internal-model control systems for voluntary movement and mental activity. a**,**b** | 'Forward' (**a**) and 'inverse' (**b**) model control systems for movement. According to the instruction given by the instructor (P) in the premotor cortex, the supplementary cortex or the anterior cingulate gyrus, the controller (CT) in the motor cortex sends command signals to the controlled object (CO; a body part or a lower motor centre). The visual cortex (VC) mediates feedback from the body part to the motor cortex. The dashed arrow indicates that the body part is copied into an internal model in the cerebellum (either a forward model (FM) or an inverse model (IM)). In the forward-model control system, control of the CO by the CT can be precisely performed by referring to the internal feedback from the forward model. In the inverse-model control system, feedback control by the CT is replaced by the inverse model itself. **c**,**d** | Forward- and inverse-model control systems for mental activities. In response to an instructor located in brain tissues that include the anterior cingulate gyrus, the controller in the prefrontal cortex initially controls a mental model (MM) that is expressed in the temporo–parietal cortex. The dashed arrow shows that the mental model is copied to a forward model or an inverse model in the cerebellum.

in the inferotemporal area. This experiment indicated that the inferotemporal-area neurons lie in a territory of the brain that is linked with conscious awareness. Binocular rivalry is also being studied in human subjects using magnetoencephalography (MEG). It has preliminarily been reported[62] that, coincident with the switching of perception that is reported by the subject, MEG power is suppressed over the parietal and occipital cortices ahead of the behavioural response.

Various types of mental representations and complex aggregates of them would constitute mental models. A relatively simple example of a mental model is visual imagery[63], which consists of activity in visual association areas that is either directly perceived or internally generated without any external visual support. Various types of visual imagery (even including the fantasies, hypnagogic imagery and hallucinations that occur in unusual situations) provide mental models that can be manipulated in a variety of thought processes, such as spatial reasoning, problem solving using analogical representation, or exploring the novel or emergent properties of objects that are implied by images[64]. In these thought processes, visual images can be manipulated by scanning, zooming, rotating, transforming or synthesizing them. These mental images are thought to be encoded in a neuronal circuit in the temporo–parietal cortex (in a column, in a set of columns[61] or more globally in regions that include the primary sensory and motor areas[65,66]). Episodic or semantic memories that are stored in the medial temporal lobe[67] might be examples of mental models that can be retrieved to conscious awareness. Visual long-term memory stored in the temporal cortex is retrieved under the top-down control of the prefrontal cortex[68]. Various conscious visuo–auditory experiences, including memories of past events, can be retrieved by electrical stimulation of the temporo–parietal cortex of a human subject[69]. This suggests that mental models are encoded in neuronal circuits at or close to the sites that were stimulated.

Recent neuroimaging and lesion studies (reviewed below) have demonstrated activation in the temporo–parital cortex during various types of thought tasks. The activity revealed by these neuroimaging studies presumably represents changes in the internal states and/or dynamic response characteristics of the neuronal circuits that encode mental models. Such changes are thought to be induced by prefrontal command signals. In these recent studies, mental models often seem to contain information on rules of computation or transformation (see below).

In order for the temporo–parietal cortex to have a role as a controlled object in the internal-model control system, it must receive command signals from, and give feedback to, the prefrontal cortex (FIG. 1c,d). The existence of projections from the prefrontal cortex to the posterior parietal and temporal cortices has previously been demonstrated[70], and these projections were recently reconfirmed in a prosimian primate[24]. It has also been established that somatosensory, auditory and visual pathways converge onto the prefrontal cortex through the temporo–parietal cortex[70].

Internal models in the cerebellum. The cerebellum is composed of numerous modules called microcomplexes, each of which is a unit learning machine that is structured uniformly with some minor regional variations (FIG. 3). The intricate cerebellar neuronal circuit provides a mechanism by which an internal model for a specific function can form in each microcomplex (BOX 2). The general idea is that the input–output relationship of a microcomplex is adaptively modified by the activity of climbing fibres. Climbing fibres convey signals representing errors and induce long-term depression (LTD) in conjunctively activated parallel-fibre–Purkinje-cell synapses[71,72]. This error learning is the basis of the cerebellum's capacity to form and update internal models.

Error signals can be derived by comparing the outputs of a controlled object with those of its forward model, when both are responding to common command signals (FIG. 2Ba,b). The inferior olive is the sole source of climbing fibres and could be the site of this comparison. This would require the inferior olive to receive projections both from cerebellar nuclei and from the temporo–parietal cortex. Indeed, an inhibitory projection from cerebellar nuclei to the inferior olive does exist (FIG. 2B), but there is no evidence for a direct projection from the temporo–parietal cortex. Nevertheless, there are indirect projections to the inferior olive from the parietal cortex

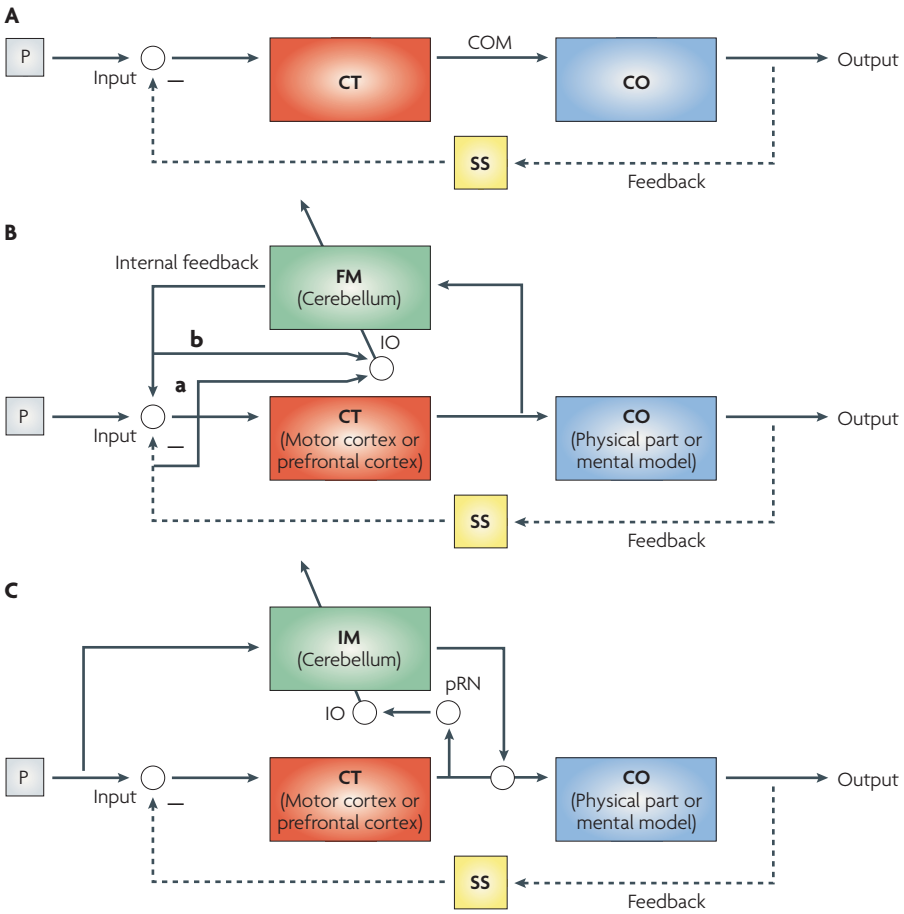

Figure 2 | **Block diagrams for internal-model control. A** | The basic structure of a control system, consisting of a controller (CT) that manipulates a controlled object (CO), an instructor (P) that gives an instruction to the controller, and a sensory system (SS) that mediates feedback (indicated by –) to the controller. This external feedback is represented using a dashed line as it can be spared when internal models are functioning. Circles indicate junctions at which signals converge or are relayed. **B** | A forward-model (FM) control system. In this system a forward model mimics the dynamic properties of the CO. Error signals are derived by comparing the outputs of the CO (monitored by the SS; **a**) with those of the forward model (**b**) in the inferior olive (IO), and are sent into the forward model to modify it. **C** | An inverse-model (IM) control system. An inverse model mimics the reciprocal of the dynamic properties of the controlled object. For learning in the inverse model, feedback errors are derived from the command signals (COM) that are generated by the CT. The arrows that obliquely cross the forward-model and inverse-model boxes represent the pathway signals that tune the dynamics of the forward model or the inverse dynamics of the inverse model. pRN, parvocellular red nucleus.

through midbrain structures. For example, the parvocellular red nucleus (pRN) receives inputs from the posterior parietal cortex[73,74] and in turn projects excitatory signals to the inferior olive[75]. A study on cats demonstrated that electrical stimulation of the parietal cortex induces slow impulse activities in climbing fibres[76].

Feedback-error learning[4,8] is another mechanism that has been suggested to contribute to learning in inverse models (FIG. 2C). If the inverse model does not work perfectly, there will be discrepancies between the output of the controlled object and the signals given by the instructor (FIG. 2C). These discrepancies, called feedback errors, will drive the motor cortex; hence, the command signals (COM in FIG. 2A) generated by the motor cortex reflect feedback errors. These command signals are sent to the cerebellum through the pRN and the inferior olive (FIG. 2C) and drive LTD-based learning processes in the inverse model (BOX 2).

The involvement of the pRN in motor learning has recently been supported by a study of voluntary motor control[77]. In this experiment, a monkey was trained to perform a reaching task with its right arm with and without the encumbrance of prisms that shifted its gaze to the right. Injection of a depressant reagent, muscimol, into the left pRN did not affect the 'no-prism' gaze–reach calibration, but did impair the learned calibration for the effect of the prisms. Error-signal pathways for the internal-model control of mental activity have yet to be investigated, but because the pRN receives inputs from the supplementary motor cortex, the cingulate gyrus, the prefrontal eye field and the posterior parietal area[73,74], it is possible that the pRN mediates learning in the cerebellum in connection with various non-motor cortical activities.

To provide an internal model for mental activity, the cerebellar hemispheres should have connections with the prefrontal and temporo–parietal cortices in the manner illustrated in FIGS 1,2. An area of the cerebellum encoding a forward model must be reciprocally connected with the prefrontal cortex (FIGS 1c,2B). In fact, the prefrontal cortex (area 46) and the cerebellar hemispheres (crus I and II of the ansiform lobule) in monkeys[21] are connected in this manner. An inverse model, on the other hand, requires there to be parallel connections between the cerebellum and the prefrontal cortex, and these areas should receive a common input from the instructor and should in turn project to a common controlled object (FIGS 1d,2C). This condition also seems to be fulfilled, because both the prefrontal cortex and the cerebellar hemispheres (the latter through the pontine nucleus) receive inputs from the anterior cingulate gyrus[24,78] and project to the temporo–parietal cortex[22–24]. Therefore, anatomical data support either type of internal model for mental activity and do not tell us which type is used *in vivo*.

The question of which type of internal model is represented by a cerebellar microcomplex has been addressed in motor-control cases by recording temporal patterns of unit spike discharges from Purkinje cells through a microelectrode. Purkinje-cell activities in the paraflocculus of monkeys performing ocular following behaviour[79] favour the inverse model, whereas those in the most lateral area of the cerebellar hemispheres of cats performing visuomotor tracking[80] are consistent with the forward model. Recordings from the cerebellar hemispheres of monkeys performing hand/arm movements under force-fields yielded conflicting data[81,82]. Presently, such unit recording cannot be applied to the human cerebellum. Unless we develop methods to analyse the activity of individual Purkinje cells in humans, it will be difficult to differentiate the contributions of forward and inverse models to the control of mental activities.

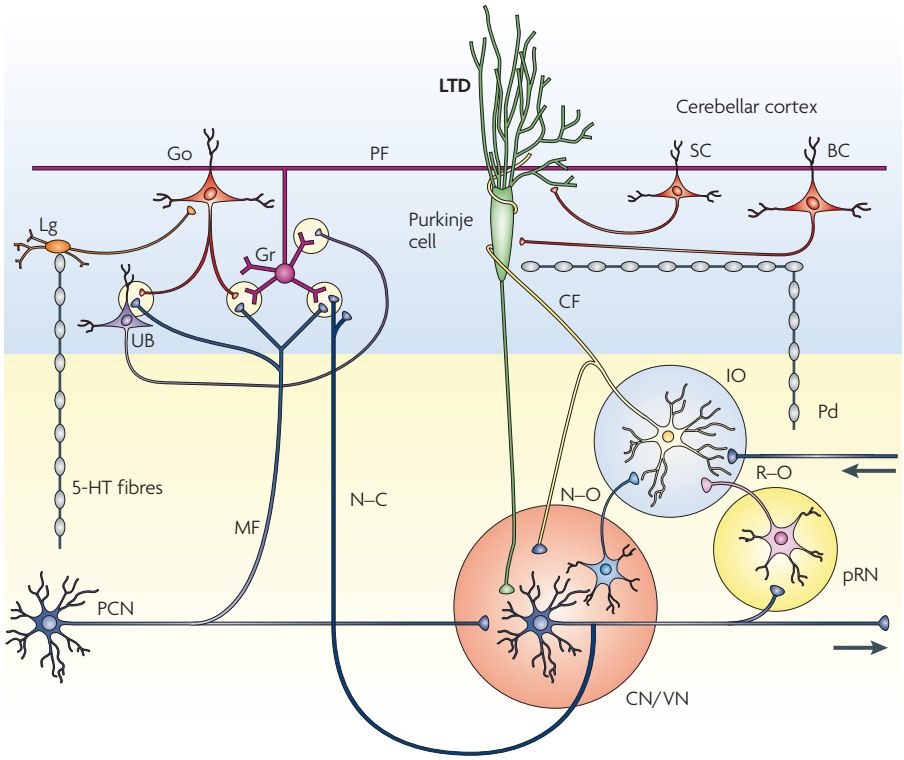

Figure 3 | **The neuronal unit machine of the cerebellum.** The cerebellum contains numerous modules, each of which is made up of uniformly structured neuronal circuits. Representative neuronal connections in this neuronal unit machine are indicated. Long-term depression (LTD) in Purkinje cells is induced by the conjunctive activation of parallel fibres (PF) and climbing fibres (CF). 5-HT, 5-hydroxytryptamine (serotonin); BC, basket cell; CN/VN, cerebellar nuclei/vestibular nuclei; Go, Golgi cell; Gr, granule cell ; IO, inferior olive; Lg, Lugaro cell; MF, mossy fibre; N–C, nucleo–cortical projection; N–O, nucleo–olivary projection; PCN, precerebellar nucleus; Pd, peptidergic fibre; pRN, parvocellular red nucleus; R–O, rubro–olivary projection; SC, stellate cell; UB, unipolar brush cell.

Neural mechanisms of thought

Here I consider how the components reviewed above are combined during a problem-solving thought. A basic assumption that I adopt is that the thought involves both explicit and implicit processes and that the explicit process occurs in the cerebral cortex whereas the implicit process is inherent to the cerebellum (BOX 3).

The internal-model hypothesis implies that, during repeated trials of a thought that uses a particular mental model, an internal model (either forward or inverse) that mimics the mental model is formed in the cerebellum. A forward model would replace the mental model as the controlled object (FIG. 2B). During manipulation of the forward model, we would not be aware of the content of the thought, as this manipulation would be taking place in the cerebellum (BOX 3). However, we would be aware that we were thinking (that is, we would be aware of the act of the thought), because the controller is in the prefrontal cortex. When the manipulated forward model began to represent a correct solution, it would then be fed back to the prefrontal executive cortex to complete the thought. We might then feel that we have solved the problem by immediate insight without reason — that is, by what has been called 'intuition' (defined by the Oxford English Dictionary as "the immediate apprehension of an object by the mind without the intervention of any reasoning process"). If an inverse model was used, this would replace the prefrontal cortex as the controller (FIG. 2C) and should enable us to proceed entirely implicitly (without sensing the act itself). The internal model of control can thus explain implicit and intuitive aspects of thought.

Let us consider what happens when a novel problem is presented as an instruction. While the working-memory system briefly holds the given problem in the slave storage systems, the central executive of working memory converts the problem into command signals. These command signals act on a mental model to alter its internal states or dynamic characteristics. This manipulation should eventually lead the mental model to generate a solution to the given problem. However, when the problem is novel, existing mental models might not be able to yield the correct solution. The mismatch between the mental-model-generated solution and the given problem (E1 in FIG. 4) would activate the brain's novelty system, as has previously been postulated to occur in neuronal-network models of classic conditioning[83]. The novelty system has been suggested to consist of the hippocampal CA1 area and dopaminergic neurons in the ventral tegmental area (VTA)[84]. The novelty system in turn activates the attentional system, which intensifies the working memory in terms of both magnitude and persistence[83]. Command signals will then strengthen their action on the mental model. When a correct solution is generated, it is fed back to the central executive to complete the thought. These are possible events during the initial explicit phase of problem-solving thought (BOX 3).

If no correct solution is derived during the explicit phase, the highly attentive operation of the working memory will soon cease. However, the prefrontal cortex contains neurons that exhibit sustained activities that are selectively related to a given task and robust against interference from intervening distractions[85]. These neurons might form a non-working-memory executive part of the prefrontal cortex (FIG. 4), which would have the controller role during the second, less attentive phase of the thought. This executive part of the cortex will continue to act on mental models, but intermittently and with a lesser degree of attention. In addition, the thought will be conducted implicitly using a forward model in place of a mental model. While trials continue, error learning (by E2; see FIG. 4) tunes the forward model to closely mimic the mental model. The modified forward model eventually replaces the mental model as the controlled object so that we implicitly continue to manipulate the forward model (FIG. 2B). Similarly, an inverse model might also be tuned by feedback-error learning, to replace the controller function of the executive cortex. Thus, the thought system will proceed implicitly until, as described above, a correct answer is generated seemingly intuitively.

Finally, I will consider a scenario in which there is an 'intuitive leap' in our thoughts. Einstein[86] wrote that discovery begins with an immediate sensory experience (E) and then, in the next step, proceeds with an intuitive leap to an axiom (A). It has been proposed that this intuitive leap arises from interactions between the prefrontal cortex and the cerebellum[19]. Is

Box 2 | **Cerebellar microcomplexes**

In the brain, different types of neurons are interconnected through synapses. The connectionist's view is that our mental, motor and emotional activities are reproduced by a system composed of neural nets with appropriate structural designs. Analyses of neural-net structure are well advanced in the cerebellum. In each functional unit of the cerebellum (called a microcomplex), the connections from a mossy fibre afferent to a granule cell to a Purkinje cell provide a three-layered neural-net structure. This neural-net structure has been modelled by Simple Perceptron, the first man-made learning machine, which consists of three layers of model cells (receptor cells instead of mossy fibres, association cells instead of granule cells and effector cells instead of Purkinje cells)[106,107]. An outside teacher instead of the climbing fibres in the microcomplex modifies the connection efficacy between the association and effector cells according to the success or failure of the machine in pattern recognition. Another model is the adaptive filter[108,109], a filter that self-adjusts its transfer function to give an optimal performance at certain input-wave frequencies. Liquid-state machines[110,111] are another type of model that is composed of a neural network or a similar computational construct. They consist of large collections of units, such as neurons. The name comes from the analogy of a stone falling into a liquid. The motion of the falling stone represents the input signal that is converted into a spatio–temporal output signal (in the analogy this signal is the pattern of liquid displacement (ripples)).

In the cerebellum, climbing fibres originating from the inferior olive convey error signals that drive a long-term depression (LTD)-based learning process at parallel-fibre–Purkinje-cell synapses[71,72]. A microcomplex also receives a third type of afferent that originates in the hypothalamus or in other brainstem structures and contains various amines or neuropeptides[112]. These neuromodulators are presumed to set the activity level of the cerebellar circuit or switch the operational mode of a microcomplex to match the required behaviour. The structure and function of microcomplexes have been rigorously analysed in connection with various reflex movements, such as the vestibulo–ocular reflex[113], the conditioned eye-blink[114] and the nociceptive withdrawal reflex[115].

there a special mechanism that alerts our conscious awareness to the completion of the implicit thought process? Following on from the idea of 'latent inhibition' in classic conditioning[83], I tentatively suggest that errors in manipulating the forward model (E3 in FIG. 4) prevent the output of the mental model from reaching the conscious awareness (FIG. 4). Successful manipulation might lead to diminution of the errors and hence removal of this blockade; the correct solution would then suddenly reach the conscious awareness in the manner of an intuitive leap.

Mental activity in the cerebellum

The previous sections have outlined the components of the internal-model hypothesis for the control of mental activities. This hypothesis predicts that there should be co-activation of the cerebellar hemisphere with the prefrontal and temporo–parietal cortices during the performance of mental tasks (FIG. 1c,d). Here I briefly outline the results of recent neuroimaging studies that demonstrate such co-activation (FIG. 5).

Prediction error. When a subject receives an unexpected painful heat stimulus on the back of their left hand, the hippocampus and the most lateral part of the cerebellum are activated simultaneously with the superior frontal and superior parietal gyri[29]. The thought mechanism illustrated in FIG. 4 explains this observation: the discrepancy between the prediction made through a forward model (no pain) and the actual event (pain) activates the attentional system and also induces error learning in the forward model.

Attention tasks. In one task[30], subjects either pressed a button at a comfortable pace or selectively attended to and responded to targets by pressing a button. Significantly greater cerebellar activation was exhibited during performance of the task that required greater attention. The thought system in FIG. 4 also explains this observation: errors generated when the subjects respond incorrectly to targets activate the attentional system, which in turn activates the forward model.

Language tasks. In a task in which subjects had to generate verbs that were appropriate for given nouns, normal subjects showed fMRI co-activation in the left prefrontal and left parietal cortices and in the right posterolateral portion of the cerebellum[26]. Similarly, in a verbal fluency test in which participants were instructed to silently provide as many different words as they could beginning with a specified letter within one minute, fMRI neuroimaging demonstrated co-activation in the left prefrontal and left dorsolateral cortices and in the right cerebellum[31]. In a selective letter-generation test in which subjects were asked to produce a set of letters from four designated sets of six letters (under certain restrictions), significant activation was exhibited in the mid-dorsolateral prefrontal cortex, the inferior frontal gyrus, the precuneus, the supramarginal gyrus and the cerebellum[32]. These results suggest that internal models of the cerebellum encompass rules of word generation.

Mental calculations. Well-trained abacus experts have the ability to mentally perform calculations on extraordinarily large numbers (often numbers with more than ten digits) with unusual speed and accuracy, even when they do not use an abacus or make any motor responses. A sophisticated visual strategy is thought to underlie this unusual ability. An fMRI study[27] revealed that the cerebellar hemisphere is activated bilaterally in combination with the frontal operculum, the superior precentral sulcus, the posterior parietal cortex including intraparietal sulcus areas, and the anterior cingulate gyrus during such calculations.

Planning and working memory. A PET study revealed that the Tower of London Task activates the cerebellar hemispheres along with the prefrontal, anterior cingulate, premotor, parietal and occipital cortices[28]. These areas

Box 3 | **Two types of thought: explicit and implicit**

When we attempt to solve a novel problem, we initially devote strenuous conscious effort (explicit thought processes) to it. If the results of this effort are unsatisfactory, we continue to think about the problem repeatedly but less attentively. Later still we might recall the problem from time to time but be otherwise unaware of it. Nevertheless, the thought seems to proceed implicitly because a solution often appears suddenly, without obvious conscious effort.

The cerebellum seems to be the site of implicit thought. Various forms of motor learning that involve the cerebellum are indeed executed implicitly. Whereas electrical stimulation of the cerebral cortex evokes conscious experiences in humans[69], stimulation of the cerebellum causes no more than increased alertness or attenuated depression and anxiety[116]. Transcranial magnetic stimulation of the cerebellum does not evoke conscious experiences[117]. These observations support the assumption that thought processes in the cerebellum do not reach conscious awareness. It is not assumed, however, that all events in the cerebral cortex reach conscious awareness; many of them might remain subliminal.

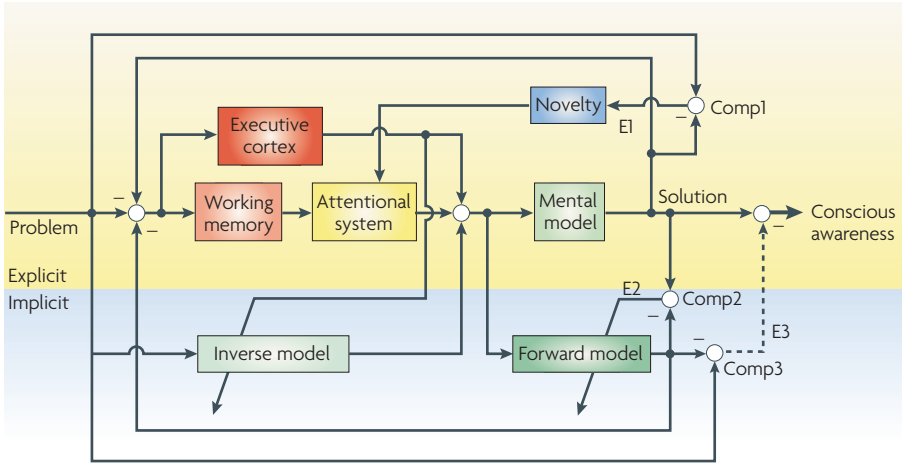

Figure 4 | **Block diagram of a thought system.** The diagram shows a schematic illustration of explicit and implicit thought processes. The working-memory system and attentional system together constitute a controller for a mental model during the attentive phase of explicit thought. The part of the prefrontal cortex that carries out executive functions acts as a controller for less attentive explicit thought. The inverse model provides a feed-forward controller. The novelty system consists of the hippocampal CA1 area and the ventral tegmental area[83]. E1 denotes errors derived from a comparison of the input problem with the output solution from a mental model. E2 denotes errors derived from a comparison of the outputs from the mental model with the outputs from the forward model. E3 denotes errors derived from a comparison of the input problem with the output of a forward model. Comp1 denotes a comparator associated with the novelty system. Comp2 denotes a comparator involving the inferior olive. Comp3 denotes a postulated comparator for E3. Subtraction and repression (in the case of E3) are indicated by a –.

include all of those that are activated in visuospatial working-memory tasks, with the exception of the rostral prefrontal cortex. In a working-memory test, participants were asked to remember six (high load) or one (low load) visually presented letters across a brief delay. An fMRI study revealed that the high-load conditions, relative to the low-load conditions, activated regions in the bilateral superior cerebellar hemispheres and the right cerebellar hemisphere[33].

During participation in a paced auditory serial-addition test (PASAT; another test of working memory), there was co-activation in the prefrontal cortex (areas 9 and 46), the superior parietal gyrus and the cerebellum[34]. However, the activated cerebellar site was in the medial parts of lobule VII, not in the lateral parts of the cerebellar hemisphere as would be expected from the present hypothesis. The activation of lobule VII seems to be related to the speech and auditory activities that are associated with the PASAT paradigm[34]. The reason why the PASAT paradigm failed to activate the lateral part of the cerebellar hemisphere is unclear.

Wisconsin Card Sorting Task and Stroop Task. The WCST and the Stroop Task are classic paradigms used to examine the functions of the prefrontal cortex. A PET study revealed that the WCST co-activates the left or bilateral dorsolateral prefrontal cortices, the bilateral inferior parietal cortices, the left superior occipital gyrus and the left cerebellar hemisphere[35]. A single-photon-emission computed tomography (SPECT) study revealed activation in the cerebellar hemispheres during the Stroop Task[36]. In these tasks, mental models and internal models might have a role in transforming the sorting rule or the colour–name correspondence.

Chess play. Neuroimaging studies of novice chess players revealed activations in the bilateral premotor areas, the parietal cortices and the occipital lobes, as well as a unilateral activation in the left cerebellar hemisphere[37]. The bilateral parietal cortices are activated equally, which is suprising because the right hemisphere has been shown to contribute more to spatial processing[87]. Even more surprising is the paucity of activation in the lateral prefrontal cortex (Brodmann's areas 45 and 46). Similar results were obtained for the Chinese board game Go[38]. The paucity of activation of the prefrontal cortex might be explained by the task that was used for the test not being sufficiently novel to alert the attentional system. It might be that, except in critical situations, players think using a less-attentive executive part of the prefrontal cortex or an inverse model of the cerebellum to that which is assumed in FIG. 4.

Mentation on the future. In one fMRI study[39], subjects were asked to envision an episode in the future or recollect an episode in the past. Both of these acts activated the medial prefrontal cortex (area 10), the posterior cingulate cortex and the medial temporal cortex. However, the task that involved imagining the future differentially activated the left lateral premotor cortex, the left precuneus and the right posterior cerebellum. This might be because a mental model representing past events had already been formed, whereas one representing future events was only newly formed or was reformed from a model for the past. This difference in the work required to form or reform mental models and internal models might cause the observed difference in neuroimages.

Cerebellar dysfunction

A corollary of the internal-model hypothesis is that dysfunction of cerebellar mental areas should lead to mental disorders. If the pathological changes broadly cover both motor and mental areas of the cerebellum, motor disorders are likely to be combined with mental disorders. Recent reports that are reviewed below show that this is indeed the case.

Damage to the cerebellum. Studies have shown that damage to the neocerebellum causes cognitive sequelae that are similar to those that are induced by damage to the dorsolateral prefrontal cortex[88]. A patient with paraneoplastic cerebellar degeneration exhibited selective frontal-lobe executive disturbance, psychomotor slowing and affective changes, despite the absence of apparent extracerebellar involvement[40]. Cerebellar mutism is a syndrome that typically affects children and, in rare cases, young adults who become mute one or two days after surgical tumour operation in the posterior fossa. Studies have shown that a lesion of the cerebellar hemispheres might be the most important factor underlying cerebellar mutism[41]. Another example of the involvement of the cerebellum in mental functions comes from a patient with a large cerebellar infarction whose ability to achieve error-free performance during repeated trials of a noun-to-verb conversion task (see above) was impaired[42]. In another study, patients with right-cerebellar lesions were impaired on an antonym-generation task, but not on other tasks such as noun generation, verb selection or lexical decision[43]. In another language test, patients with cerebellar lesions showed impaired ability to generate lists of words according to a phonemic rule[44]. In

accordance with the neuroimaging studies of language tasks described above, these findings underline the importance of the cerebellum in regulating language function, presumably through its ability to provide internal models that encompass rules of word generation.

Autism. More than 90% of autistic patients examined at autopsy have well-defined cerebellar anatomical abnormalities[89]. Moreover, reductions in Purkinje-cell size[90] and nicotinic-receptor abnormalities[91] have also been reported in the cerebellum of these patients. Autism-like cellular and behavioural phenotypes can be induced in mice by knocking out the gene that encodes Ca^{2+}-dependent activator protein for secretin 2 (CADPS2; also known as CAPS2)[92]; this gene, which is located in susceptibility locus 1 of chromosome 7, is important for cerebellar development[93]. fMRI activation in anatomically defined cerebellar regions has been examined in autistic patients[30]. During an attention task, autistic individuals showed significantly attenuated cerebellar activation. Autism is characterized by a defective 'theory of mind', which leads to a specific loss of the ability to understand the thoughts and motivational processes of other individuals, with consequent actions that are socially aberrant[94]. Poor social interaction skills in autistic patients might arise from impairment of the cerebellum's role in the control of mental activities and social interactions[20]. In particular, the impairment might be caused by the failure to form an internal model that simulates the internal models of other people[95].

Developmental dyslexia. Developmental dyslexia is caused by an impaired ability to divide heard words into discrete segments (such as phonemes), which in turn leads to difficulty in learning spelling–sound associations. It has also been suggested that the underlying deficit is the brain's inability to filter out irrelevant data, such as perceptual noise[96]. The cause of these problems has been pinned down to dysfunction in the frontal and temporal lobes[97], in a magnocellular component of the visual pathway[98], and in the cerebellum[46]. A cerebellar origin of developmental dyslexia was also suggested on the basis that dyslexic children perform less well in certain motor and balancing tasks. Dyslexic participants also show abnormal eye-blink conditioning[46] and poorer performance than control children in tasks in which the speed and accuracy of pointing are both factors[46]. These observations suggest that dyslexic patients have impaired internal models, possibly including those that normally represent rules for analysing heard words and learning spelling–sound associations.

Schizophrenia. Schizophrenic patients have been shown to be significantly more impaired in prism-adaptation tests and to have poorer procedural-learning abilities, both of which are thought to involve internal models in the cerebellum[46]. Schizophrenic patients did not differ from healthy controls in terms of after-effects once the prisms were removed, but they did have significantly greater difficulties in reorientating. It is possible that cerebellar dysfunction in these patients affected both their voluntary movements (tested with prism-adaptation tests) and their mental processes, the latter being manifested as schizophrenia symptoms.

Hallucination, a typical symptom of schizophrenia, might arise from the failure of an internal model to provide appropriate internal feedback to the controller, which in turn might therefore not recognize that the experience is not externally produced. Some other abnormalities experienced by schizophrenic patients suggest that there might also be a dysfunction of their self-monitoring through an internal model[48]. Normally we can sense the difference between self-produced and externally produced stimuli, presumably because the effect of a self-produced stimulus is cancelled out by a prediction from an internal model whereas that of an externally produced stimulus is not[99,100]. Schizophrenic patients often fail to show such a sense of difference,

Glossary

Body-part dynamics
The dynamic properties of a body part that are determined by physical factors such as weight, length, centre of gravity and viscosity. The dynamic properties in turn determine the movement of the body part in response to command signals.

Column
A basic functional unit of the cerebral cortex. Each column is approximately 0.5 mm wide and 2 mm high and contains approximately 10,000 neurons. These units operate much like microcircuits in a computer. The human cerebral cortex is thought to have approximately a million columns.

Control system
A term that was originally used to refer to a mechanical or chemical system equipped with a mechanism for manipulating an object or regulating a process. The term now broadly applies to an informational, biological, neural, psychological or social system.

Controlled object
A key part of a control system, a controlled object converts a command into an output action. For example, a muscle converts signals in nerves into a contraction.

Controller
A key part of a control system, a controller converts a given instruction into a command. For example, the brain converts an instructed spatial position of a target into a command, which consists of signals in the nerves that innervate muscles.

Error signals
Signals representing errors in a system. The errors are discrepancies in the performance of a control system from either the instruction (consequence errors) or the prediction by an internal model (internal errors).

Instructor
The part of a control system that supplies an instruction to the controller. The instructor gives a goal towards which a control system should work.

Internal model
A functional dummy of a body part or of a mental representation in the cerebral cortex. Internal models are encoded in the neuronal circuitry of the cerebellum and mimic the essential properties of a body part or mental representation.

Paced auditory serial-addition test
(PASAT). A test that is used to impose a high cognitive load on the working memory. Subjects receive a pseudo-random auditory presentation of a number between 1 and 9 every 3 seconds and are asked to add consecutive numbers and provide the answer to each addition verbally.

Stroop Task
A task in which subjects are instructed to either read words or name the colour in which the words are written. Subjects must selectively attend to one attribute, particularly when naming the colour of a conflict stimulus (for example, the word 'green' displayed in red).

Tower of London Task
A test of planning capability. Typically, starting from an initial condition in which three differently coloured rings are distributed to three poles, the subject is asked to gather all of the rings to one particular pole by moving one at a time and not making a total of more than five moves. A modified version is used in neuroimaging to avoid contamination of the results with activity relating to movements.

Unit learning machine
The cerebellum contains numerous modular units, each of which consists of a uniform set of neuronal circuits that is capable of learning. Each unit learning machine is inserted into a neural control system and carries out the role of an internal model.

Wisconsin Card-Sorting Task
(WCST). In this task, participants are given cards that can be sorted by colour, shape or name, and must deduce the correct sorting criterion. After several consecutive correct sorts, the correct sorting criterion is changed without warning.

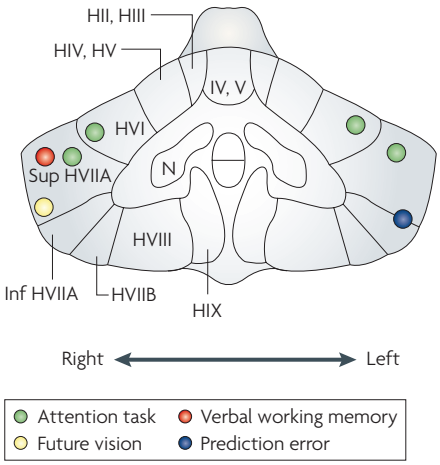

Figure 5 | **Mental activities in the cerebellum.** Neuroimaging studies have provided evidence of a role for the cerebellum in various mental activities. The figure shows a coronal section of a human cerebellum[33] on which the sites of the observed activities are indicated by coloured circles. The tasks that elicited these activities include an attention task[30], a future-vision task[39] and a verbal working-memory task[33]. Activity that was observed during prediction error is also indicated. H, hemisphere area; Inf, inferior cerebellum; N, cerebellar nucleus; Sup, superior cerebellum.

suggesting failed operation of an internal model. These patients also exhibit passivity — the feeling that their will is replaced by that of some other force or agent. This feeling might arise as a result of a lack of awareness of their intended actions, which might be due to impairments in the internal model that predicts a possible outcome of the intended actions.

Future perspectives

The preceding sections have reviewed the evidence that cerebellar internal models control mental activities (this evidence includes the existence of appropriate neural wiring and appropriate mental activity in the cerebellum and the association of some mental disorders with cerebellar dysfunction). As described above, most evidence is consistent with the concept of internal-model control of mental activity. This encourages future investigations into how our brain orchestrates mental activities on the basis of collaborative functions of the cerebral cortex and cerebellum.

However, there are obstacles that hamper our progress towards this goal. A serious insufficiency on the experimental side is that we are unable to directly analyse the signals of individual neurons in the human brain. This makes it difficult to determine whether forward and inverse models co-exist in the same cerebellar areas or are differentially located. As current neuroimaging technology does not allow such a distinction to be made, the idea that there might be both forward and inverse models should be maintained for the time being.

Another gap that needs to be filled concerns the theoretical aspect of this work. The successful development of a computational approach to define mental models will be central to understanding internal-model control, and this understanding could then be applied to the vast field of artificial intelligence.

The hypothesis of internal-model control of mental activities opens up the possibility that implicit mechanisms of thought might be explained. The implicit domain is a major part of our mind, as advocated by Sigmund Freud and supported by recent authors[101–104]. The internal-model control would allow us to scientifically approach this still little-explored and mysterious territory of the mind.

Masao Ito is at the RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan.
e-mail: masao@brain.riken.jp

doi:10.1038/nrn2332
Published online 5 March 2008

1. Poon, C.-S. & Merfeld, D. M. Internal model: the state of art. *J. Neural Eng.* **2** (2005).
2. Ito, M. Neurophysiological basis of the cerebellar motor control system. *Int. J. Neurol.* **7**, 162–176 (1970).
3. Ito, M. *The Cerebellum and Neural Control* (Raven, New York, 1984).
4. Kawato, M., Furukawa, K. & Suzuki, R. A hierarchical neural-network model for control and learning of voluntary movement. *Biol. Cybern.* **57**, 169–185 (1987).
5. Miall, R. C., Weir, D. J., Wolpert, D. M. & Stein, J. F. Is the cerebellum a Smith predictor? *J. Motor Behav.* **25**, 203–216 (1993).
6. Wolpert, D. M. & Kawato, M. Multiple paired forward and inverse models for motor control. *Neural Netw.* **11**, 1317–1329 (1998).
7. Wolpert, D. M., Miall, R. C. & Kawato, M. Internal models in the cerebellum. *Trends Cogn. Sci.* **2**, 338–347 (1998).
8. Kawato, M. Cerebellum: models. In *New Encyclopedia of Neuroscience* (ed. Squire, L.) (Elsevier, in the press 2008).
9. Schmahmann, J. D. Disorders of the cerebellum: ataxia, dysmetria of thought, and the cerebellar cognitive affective syndrome. *J. Neuropsychiatry Clin. Neurosci.* **16**, 367–378 (2004).
10. Leiner, H. C., Leiner, A. L. & Dow, R. S. Does the cerebellum contribute to mental skills? *Behav. Neurosci.* **100**, 443–454 (1986).
11. Leiner, H. C., Leiner, A. L. & Dow, R. S. Cognitive and language functions of the human cerebellum. *Trends Neurosci.* **16**, 444–447 (1993).
12. Sasaki, K., Kawaguchi, S., Oka, H., Sakai, M. & Mizuno, N. Electrophysiological studies on the cerebellocerebral projections in monkeys. *Exp. Brain Res.* **24**, 495–507 (1976).
13. Schmahmann, J. D. An emerging concept: the cerebellar contribution to higher function. *Arch. Neurol.* **48**, 1178–1187 (1991).
14. Ito, M. A new physiological concept on cerebellum. *Rev. Neurol. (Paris)* **146**, 564–569 (1990).
15. Ito, M. in *The Principles of Design and Operation of the Brain (Experimental Brain Research)* (eds Eccles, J. C. & Creutzfeldt, O.) 281–292 (Springer, 1990).
16. Ito, M. Movement and thought: identical control mechanisms by the cerebellum. *Trends Neurosci.* **16**, 448–450 (1993).
17. Ito, M.in *Cerebellum and Cognition (International Review of Neurobiology)* (eds Schmahmann, J. D., Bradley, R. J., Adron Harris, R. & Jenner, P.) 475–487 (Academic, 1997).
18. Ito, M. Bases and implications of learning in the cerebellum – adaptive control and internal model mechanism. *Prog. Brain Res.* **148**, 95–109 (2006).
19. Vandervert, L. R., Schlimp, P. H. & Liu, H. How working memory and the cerebellum collaborate to produce creativity and innovation. *Creativity Res. J.* **19**, 1–18 (2007).
20. Wolpert, D. M., Doya, K. & Kawato, M. A unifying computational framework for motor control and social interaction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **358**, 593–602 (2003).
21. Kelly, R. M. & Strick, P. L. Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *J. Neurosci.* **23**, 8432–8444 (2003).
22. Dum, R. P. & Strick, P. L. An unfolded map of the cerebellar dentate nucleus and its projections to the cerebral cortex. *J. Neurophysiol.* **89**, 634–639 (2003).
23. Clower, D. M., Dum, R. P. & Strick, P. L., Basal ganglia and cerebellar inputs to 'AIP'. *Cereb. Cortex* **15**, 913–920 (2005).
24. Fang, P.-C., Stepniewska, I. & Kaas, J. H. Ipsilateral cortical connections of motor, premotor, frontal eye, and posterior parietal fields in a prosimian primate, *Otalemur garnetti*. *J. Comp. Neurol.* **490**, 305–333 (2005).
25. Ramnani, N. The primate cortico-cerebellar system: anatomy and function. *Nature Rev. Neurosci.* **7**, 511–522, (2006).
26. Fiez, J. A., Raicle, M. E., Balota, D. A., Tallal, P. & Petersen, S. E. PET activation of posterior temporal regions during auditory word presentation and verb generation. *Cereb. Cortex* **6**, 1–10 (1996).
27. Hanakawa, T., Honda, M., Okada, T., Fukuyama, H. & Shibasaki, H. Neural correlates underlying mental calculation in abacus experts: a functional magnetic resonance imaging study. *Neuroimage* **19**, 296–307 (2003).
28. Baker, S. C. *et al.* Neural systems engaged by planning: a PET study of the Tower of London task. *Neuropsychologia* **34**, 515–526 (1996).
29. Ploghaus, A. *et al.* Learning about pain: the neural substrate of the prediction error for aversive events. *Proc. Natl Acad. Sci. USA* **97**, 9281–9286 (2000).
30. Allen, G. & Courchesne, E. Differential effects of developmental cerebellar abnormality on cognitive and motor functions in the cerebellum: an fMRI study of autism. *Am. J. Psychiatry* **160**, 262–273 (2003).
31. Schlosser, R. *et al.* Functional magnetic resonance imaging of human brain activity in a verbal fluency task. *J. Neurol. Neurosurg. Psychiatr.* **64**, 492–498 (1998).
32. de Zubicaray, G. I. *et al.* Prefrontal cortex involvement in selective letter generation: a functional magnetic resonance imaging study. *Cortex* **34**, 389–401 (1998).
33. Desmond, J. E., Gabrieli, J. D. E., Wagner, A. D., Ginier, B. L. & Glove, G. H. Lobular patterns of cerebellar activation in verbal working memory and finger-tapping tasks as revealed by functional MRI. *J. Neurosci.* **17**, 9675–9685 (1997).
34. Hayter, A. L., Lamgton, D. W. & Ramnani, N. Cerebellar contributions to working memory. *Neuroimage* **36**, 943–954 (2007).
35. Nagahama, Y. *et al.* Cerebral activation during performance of a card sorting test. *Brain* **119**, 1667–1675 (1996).
36. Narita, H., Odawara, T., Iseki, E., Kosaka, K. & Hirayasu, Y. Psychomotor retardation correlates with frontal hypoperfusion and the Modified Stroop Test in patients with major depression under 60-years-old. *Psychiat. Clin. Neurosci.* **58**, 389–395 (2004).
37. Atherton, M., Zhung, J., Bart, W. M., Hu, X. & He, S. A functional MRI study of high-level cognition. I. The game of chess. *Cogn. Brain Res.* **16**, 26–31 (2003).
38. Chen, X., Zhang, D., Zhang, X., Li, Z. & Meng, X. A functional MRI study of high-level cognition. II. The game of GO. *Cogn. Brain Res.* **16**, 32–37 (2003).
39. Szpunar, K. K., Watson, J. M. & McDermott, K. B. Neural substrates of envisioning the future. *Proc. Natl Acad. Sci. USA* **2**, 642–647 (2007).
40. Collinson, S. L., Anthonisz, B., Courtenay, D. & Winter, C. Frontal executive impairment associated with paraneoplastic cerebellar degeneration: a case study. *Neurocase* **12**, 350–354 (2006).

41. Janssen, G. *et al.* Cerebellar mutism syndrome. *Klin. Padiatr.* **210**, 243–247 (1998).
42. Fiez, J. A., Petersen, S. E., Cheney, M. K. & Raichle, M. E. Impaired non-motor learning and error detection associated with cerebellar damage. A single case study. *Brain* **115**, 155–178 (1992).
43. Gebhart, A. L., Petersen, S. E & Thach, W. T. Role of the posterolateral cerebellum in language. *Ann. NY Acad. Sci.* **978**, 318–333 (2002).
44. Leggio, M. G., Silveri, M. C., Petrosini, L. & Molinari, M. Phonological grouping is specifically affected in cerebellar patients; a verbal fluency study. *J. Neurol. Neurosurg. Psychiatr.* **69**, 102–106 (2000).
45. Nicolson, R. I., Daum, I., Schugens, M. M., Fawcett, A. J. & Schulz, A. Eyeblink conditioning indicates cerebellar abnormality in dyslexia. *Exp. Brain Res.* **143**, 42–50 (2002).
46. Stoodley, C. J., Fawcett, A. J., Nicolson, R. I. & Stein, J. F. Balancing and pointing tasks in dyslexic and control adults. *Dyslexia* **12**, 276–288 (2006).
47. Bigelow, N. O. *et al.* Prism adaptation in schizophrenia. *Brain Cogn.* **61**, 235–242 (2006).
48. Blakemore, S.-J., Smith, J., Steel, R., Johnstone, E. C. & Frith, C. D. The perception of self-produced sensory stimuli in patients with auditory hallucinations and passivity experiences: evidence for a breakdown in self-monitoring. *Psychol. Med.* **30**, 1131–1139 (2000).
49. Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
50. Happaney, K., Zelazo, P. D. & Stuss, D. T. Development of orbitofrontal function: current theme and future directions. *Brain Cogn.* **55**, 1–10 (2004).
51. Baddeley, A. Working memory: looking back and looking forward. *Nature Rev. Neurosci.* **4**, 829–839 (2003).
52. Baddeley, A. & Hitch, G. J. in *Recent Advances in Learning and Motivation* (ed. Bower, G. A.) 47–89 (Academic, New York, 1974).
53. Smith, E. E. & Jonides, J. Working memory: a view from neuroimaging. *Cognit. Psychol.* **33**, 5–42 (1997).
54. Stuss, D. T. & Knight, R. T. *Principles of Frontal Lobe Function* (Oxford Univ. Press, New York, 2002).
55. Baddeley, A. D. & Logie, R. H. in *Models of Working Memory* (eds Miyake, A. & Shah, P.) 28–61 (Cambridge Univ. Press, Cambridge, UK, 1999).
56. Craik, K. J. W. *The Nature of Explanation* (Cambridge Univ. Press, Cambridge, UK, 1943).
57. Johnson-Raird, P. N. *Mental Models* (Cambridge Univ. Press, 1983).
58. Held, C., Knauff, M. & Vosgerau, G. *Mental Models and the Mind: Current Developments in Cognitive Psychology, Neuroscience, and Philosophy of Mind* (*Advances in Psychology*) (Elsevier, 2006).
59. Tanaka, K. Inferotemporal cortex and object vision. *Ann. Rev. Neurosci.* **19**, 109–139 (1996).
60. Tsunoda, K., Yamane, Y., Nishizaka, M. & Tanifuji, M. Complex objects are represented in the macaque inferotemporal cortex by the combination of feature columns. *Nature Neurosci.* **4**, 832–838 (2001).
61. Sheinberg., D. L. & Logothetis, N.K. The role of temporal areas in perceptual organization. *Proc. Natl Acad. Sci. USA* **94**, 3408–3413 (1997).
62. Tarok, L. & Srinivasan, R. Transient MEG frequency-tagging response induced by switching stimulus salience during binocular rivalry. (Abstr.) 201.9/BBB20 (Society for Neuroscience Annual Meeting, San Diego, 2007).
63. Kosslyn, S. M., Ganis, G. & Thompson, W. L. Neural foundations of imagery. *Nature Rev. Neurosci.* **2**, 635–642 (2001).
64. Pearson, D., De Beni, R. & Cornoldi, C. in *Imagery, Language and Visuo-Spatial Thinking* (eds Denis, M., Logie, R. H., Cornoldi, C., De Vega, M. & Engelkamp, J.) 1–27 (Psychological Press, 2001).
65. Hesslow, G. Conscious thought as simulation of behaviour and perception. *Trends Cogn. Sci.* **6**, 242–247 (2002).
66. Le Bihan, D. *et al.* Activation of human primary visual cortex during visual recall: a magnetic resonance imaging study. *Proc. Natl Acad. Sci. USA* **90**, 11802–11805 (1993).
67. Miyashita, Y. Cognitive memory: cellular and network machineries and their top-down control. *Science* **306**, 435–440 (2004).
68. Hasegawa, I., Fukushima, T., Ihara, T. & Miyashita, Y. Callosal window between prefrontal cortices: cognitive interaction to retrieve long-term memory. *Science* **281**, 814–818 (1998).
69. Penfield, W. & Perot, P. The brain's record of auditory and visual experience; a final summary and discussion. *Brain* **86**, 595–696 (1963).
70. Fuster, J. M. *The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe* (Raven, New York, 1997).
71. Ito, M. Cerebellar long-term depression: characterization, signal transduction and functional roles. *Physiol. Rev.* **81**, 1143–1195 (2001).
72. Ito, M. Cerebellar circuitry as a neuronal machine. *Prog. Neurobiol.* **78**, 272–303 (2006).
73. Ralston, D. D. Corticorubral synaptic organization in *Macaca fascicularis*: a study utilizing degeneration, anterograde transport of WGA-HRP, and combined immuno-GABA-gold technique and computer-assisted reconstruction. *J. Comp. Neurol.* **350**, 657–673 (1994).
74. Burman, K., Darian-Smith, C. & Darian-Smith, I. Macaque red nucleus: origins of spinal and olivary projections and terminations of cortical inputs. *J. Comp. Neurol.* **423**, 178–196 (2000).
75. Burman, K., Darian-Smith, C. & Darian-Smith, I. Geometry of rubrospinal, rubroolivary, and local circuit neurons in the macaque red nucleus. *J. Comp. Neurol.* **423**, 197–219, (2000).
76. Oka, H., Jinnai, K. & Yamamoto, T. The parieto-rubro-olivary pathway in the cat. *Exp. Brain Res.* **37**, 115–125 (1979).
77. Yttri, E., Smith, A., Reid, E. & Thach, W. T. Inactivation of parvocellular red nucleus impairs prism adaptation to and learning of contralateral gaze-reach shift. (Abstr.) 440.15/K8 (Society for Neuroscience Annual Meeting, Atlanta, 2006).
78. Schmahmann, J. D. & Pandya, D. N. Anatomical investigation of projections to the basis pontis from posterior parietal association cortices in rhesus monkey. *J. Comp. Neurol.* **289**, 53–73 (1989).
79. Shidara, M., Kawano, M., Gomi, H. & Kawato, M. Inverse-dynamics encoding of eye movements by Purkinje cells in the cerebellum. *Nature* **365**, 50–52 (1993).
80. Miles, G. B., Cerminara, N. L. & Marple-Horvat, D. E. Purkinje cells in the lateral cerebellum of the cat encode visual events and target motion during visually guided reaching. *J. Physiol.* **571**, 619–637 (2006).
81. Pasalar, S., Roitman, A. V., Durfee, W. K. & Ebner, T. J. Force field effects on cerebellar Purkinje cell discharge with implications for internal models. *Nature Neurosci.* **9**, 1404–1411 (2006).
82. Yamamoto, K., Kawato, M., Kotosaka, S. & Kitazawa, S. Encoding of movement dynamics by Purkinje cell simple spike activity during fast arm movements under resistive and assistive force fields. *J. Neurophysiol.* **97**, 1588–1599 (2007).
83. Schmajuk, N. A., Lam, Y.-W. & Gray, J. A. Latent inhibition: a neural network approach. *J. Exp. Psychol.* **22**, 321–349 (1996).
84. Lisman, J. E. & Grace, A. A. The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* **46**, 703–713 (2005).
85. Miller, E. K., Erickson, C. A. & Desimone, R. Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J. Neurosci.* **16**, 5154–5167 (1996).
86. Einstein, A. *Letters a Maurice Solvince* (Gauthier-Villars, Paris, 1956).
87. Spriner, S. P. & Deutsch, G. *Left Brain, Right Brain: Perspectives on Cognitive Neuroscience* 5th edn (Freeman, New York, 1998).
88. Diamond, A. Close interrelation of motor development and cognitive development and the cerebellum and prefrontal cortex. *Child Develop.* **71**, 44–56 (2000).
89. Carper, R. A. & Courchesne, E. Inverse correlation between frontal lobe and cerebellum sizes in children with autism. *Brain* **123**, 836–844 (2000).
90. Fetami, S. H. *et al.* Purkinje cell size is reduced in cerebellum of patients with autism. *Cell. Mol. Neurosci.* **22**, 171–175 (2002).
91. Lee, M. *et al.* Nicotinic receptor abnormalities in the cerebellar cortex in autism. *Brain* **125**, 1483–1495 (2002).
92. Sadakata, T. *et al.* Autistic-like phenotypes in *Cadps2*-knockout mice and aberrant CADPS2 splicing in austistic patients. *J. Clin. Invest.* **117**, 931–943 (2007).
93. Sadakata, T. *et al.* Impaired cerebellar development and function in mice lacking CAPS2, a protein involved in neurotrophin release. *J. Neurosci.* **27**, 2472–2482 (2007).
94. Deuel, R. K. Autism: a cognitive developmental riddle. *Pediatr. Neurol.* **26**, 349–357 (2002).
95. Ito, M. Nurturing the brain as an emerging research field involving child neurology. *Brain Dev.* **26**, 429–433 (2004).
96. Sperling, A. J., Lu, Z. L., Manis, F. R. & Seidenberg, M. S. Deficits in perceptual noise exclusion in developmental dyslexia. *Nature Neurosci.* **8**, 862–863 (2005).
97. Galaburda, A. M., Sherman, G. F., Rosen, G. D., Aboitiz, F. & Geschwind, N. Developmental dyslexia: four consecutive patients with cortical anomalies. *Ann. Neurol.* **18**, 222–233 (1985).
98. Stein, J. & Walsh, V. To see but not to read: the magnocellular theory of dyslexia. *Trends Neurosci.* **20**, 147–152 (1997).
99. Blakemore, S.-J., Wolpert, D. M. & Frith, C. D. Central cancellation of self-produced tickle sensation. *Nature Neurosci.* **1**, 635–640 (1998).
100. Blakemore, S. J. & Sirigu, A. Action prediction in the cerebellum and in the parietal lobe. *Exp. Brain Res.* **153**, 239–245 (2003).
101. Solms, M. & Turnbull, O. *The Brain and the Inner World*. (OTHER, 2002).
102. Wegner, D. M. *The Illusion of Conscious Will* (MIT Press, Cambridge, Massachusetts, 2002).
103. Crick, F. & Koch, C. Towards a neurobiological theory of consciousness. *Semin. Neurosci.* **2**, 263–275 (1990).
104. Crick, F., Koch, C., Kreiman, G. & Fried, I. Consciousness and neurosurgery. *Neurosurgery* **55**, 273–282 (2004).
105. Barlow, J. S. *The Cerebellum and Adaptive Control* (Cambridge Univ. Press, New York, 2002).
106. Marr, D. A theory of cerebellar cortex. *J. Physiol.* **202**, 437–470 (1969).
107. Albus, J. S. A theory of cerebellar function. *Math. Biosci.* **10**, 25–61 (1971).
108. Fujita, M. Adaptive filter model of the cerebellum. *Biol. Cybern.* **45**, 195–206 (1982).
109. Mauk, M. D. & Donegan, N. H. A model of Pavlovian eyelid conditioning based on the synaptic organization of the cerebellum. *Learn. Mem.* **4**, 130–158 (1997).
110. Yamazaki, T. & Tanaka, S. Neural modeling of an internal clock. *Neural Comput.* **17**, 1–27 (2005).
111. Yamazaki, T. & Tanaka, S. The cerebellum as a liquid state machine. *Neural Netw.* **20**, 290–297 (2007).
112. Schweighofer, N., Doya, K. & Kuroda, S. Cerebellar aminergic neuromodulation: towards a functional understanding. *Brain Res. Rev.* **44**, 103–116 (2004).
113. Nagao, S., Ito, M. & Karachot, L. Eye field in the cerebellar flocculus of pigmented rabbits determined with local electrical stimulation. *Neurosci. Res.* **3**, 39–51 (1985).
114. Christian, K. M. & Thompson, R. F. Neural substrates of eyeblink conditioning: acquisition and retention. *Learn. Mem.* **10**, 427–455 (2003).
115. Apps, R. & Garwicz, M. Anatomical and physiological foundations of cerebellar information processing. *Nature Rev. Neurosci.* **6**, 297–311 (2005).
116. Riklan, M., Cullinan, T., Shulman, M. & Cooper, I. S. A psychometric study of chronic cerebellar stimulation in man. *Biol. Psychiatry* **11**, 543–574 (1976).
117. Koch, G. *et al.* Repetitive TMS of cerebellum interferes with millisecond time processing. *Exp. Brain Res.* **179**, 291–299 (2007).

Acknowledgements
I would like to thank the RIKEN Brain Science Institute for their continuous support of my research on the cerebellum.

DATABASES

Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
CADPS2 |

FURTHER INFORMATION

Masao Ito's homepage:
http://www.BRAIN.RIKEN.JP/en/m_ito.html

ALL LINKS ARE ACTIVE IN THE ONLINE PDF