Mental Magnitudes

C. R. Gallistel

Department of Psychology and

Rutgers Center for Cognitive Science

Rutgers University

New Brunswick, NJ 08854-8020

email: galliste@ruccs.rutgers.edu

Text: 2,742 words

Box 1: 404 words, 1 Figure

Box 2: 235 words, 1 Figure

Box 3: 196 words

As is evident from the other chapters in this volume, the experimental study of the mind's foundational abstractions has become an important part of cognitive science, particularly in human cognitive development and in animal cognition. Prominent among the abstractions that have been experimentally studied are space, time, number, rate and probability. They have all now been shown to play a fundamental role in the mentation of nonverbal animals and preverbal humans [1-5]. These results have moved cognitive science in a rationalist direction. In an empiricist theory of mind, these concepts are somehow induced from primitive sensory experience. Because language has often been thought to mediate the induction, these abstractions were often supposed to be absent in the mentation of nonverbal or preverbal beings. In a rationalist epistemology, by contrast, these abstractions are foundational. They make sensory experience possible. I suggest that the brain's ability to represent these foundational abstractions depends on a still more basic ability, the ability to store, retrieve and arithmetically manipulate signed magnitudes [6]. If this is true, then the discovery of the physical basis of this ability is a sine qua non for a well-founded cognitive neuroscience.

By magnitude I mean computable number, a magnitude that can be subjected to arithmetic manipulation in a physically realized system. I use 'magnitude' to avoid confusions that arise from 'number'. 'Number' may denote the numerosity of a set, or it may denote the symbols in a system of arithmetic, which may or may not refer to numerosities. The symbol '1' may denote the numerosity of a set, or the height in meters of a large dog, which isn't a numerosity, or the multiplicative identity element in the system of arithmetic. Insofar as magnitude is physically representable, the computable numbers represent it. The representation of the other foundational abstractions rests, I argue, on the brain's ability to represent computable number.

Computational Implications of Behavioral Results

The representation of space arises in its most basic form in the process of dead reckoning (aka path integration), which is the foundation of an animal's ability to find its way back whence it came and to construct a representation of the locations of landscapes and locations relative to a home base [2]. It requires summing successive displacements in an allocentric (other-centered) framework, a framework in which the coordinates of locations other than that of the animal do not change as the animal moves. By summing successive small displacements (small changes in its location), the animal maintains a representation of its location in the allocentric framework. This representation makes it possible to record locations of places and objects of interest as it encounters them, thereby constructing a cognitive map of its experienced environment. Computational considerations make it likely that this representation is Cartesian and allocentric. Polar and egocentric representations rapidly become inaccurate, because they integrate the step-by-step errors in the signals for the direction and distance of displacements (cite PLoS paper if it gets accepted). Representations of experienced locations are vectors, that is, ordered sets of magnitudes. A fundamental operation in navigation is computing courses to be run. A course to be run is the range and bearing of the destination location from the current location. Assuming that the vectors are Cartesian, the range and bearing are the modulus and angle of the difference between the destination vector and the current-location vector. This difference vector is the element-by-element differences between the two vectors. Thus, the representation of spatial location depends on the arithmetic processing of magnitudes.

The representation of time takes two forms: the representation of phase (location within a cycle) and the representation of temporal intervals (the temporal distance of one event from another). Nonverbal animals represent both [2]. Nonverbal animals compute signed temporal differences: how long it will be, and how long it has been. An elementary behavioral manifestation of the first (how long it will be) is the anticipation of time of feeding seen when animals are fed at regular times of day [see for review 2]. An elementary behavioral manifestation of the latter is the cache-revisiting behavior of scrub jays: Their choice of which caches to visit first depends strongly on their knowledge of how long it has been since they buried what where and their acquired knowledge of the (experimenter-determined) rotting-times for the different foods they have cached [7].

It has been widely supposed that the representation of temporal intervals is generated by an interval-timing mechanism [8]. There is, however, a conceptual problem with this supposition: The ability to record the first occurrence of an interesting temporal interval would seem to require the starting of an infinite number of timers for each of the very large number of experienced events that might turn out to be “the start of something interesting”--or not. Because, there is no knowing what may follow an event that might or might not mark the onset of an interesting interval, one would need to start a timer in anticipation of any eventuality, but the eventualities are infinite. If one cannot record the first occurrence of an interval, then every recurrence is effectively the first. If you do not remember that food followed the bell after 30 seconds the last time you heard the bell, then you have no idea how long it will be the second time you hear it, nor any way of realizing that the second interval was the same as the first. Resolving this paradox seems to require the assumption that temporal intervals are derived from the representation of temporal locations, just as displacements (directed spatial intervals) are derived from differences in spatial locations. This, in turn leads to arithmetic operations on temporal vectors (see Gallistel, 1990, for details).

Rats represent rates (numbers of events divided by the durations of the intervals over which they have been experienced) and combine them multiplicatively with reward magnitudes [9]. Both mice and adult human subjects represent the uncertainty in their estimates of elapsing durations (a probability distribution defined over a continuous variable) and discrete probability (the proportion between the number of trials of one kind and the number of trials of a different kind) can combine these two representations multiplicatively to estimate an optimal target time [1]. Human adult subjects generalize from the proportion between two durations to the proportion between two integers without being instructed to do so [10]. These are but a few of the many experimental results that imply that the analog magnitude system represents both discrete and continuous quantity and brings to bear on the symbols that refer to these abstract aspects of our experience the representational power of the arithmetic field (the system of arithmetic that is closed under addition, multiplication and their inverses).

Constraints on Mental Magnitudes

Closure. Closure is an important constraint on the mechanism that implement arithmetic processing in the brain. Closure means that there are no inputs that crash the machine. Closure under subtraction requires that magnitudes have sign (direction), because otherwise entering a subtrahend greater than the minuend would crash the machine; it would not be able to produce a valid output. Rats learn directed (signed) temporal differences; they distinguish between whether the reward comes before or after the signal and they can integrate one directed difference with another [11]. When humans are asked to tap out the estimated sums or differences between two sequences of rapid arhythmic flashes, one seen on the right side of the screen and one on the left, the mean numbers that they make are proportional to the true sums and differences and the variability increased in proportion to the magnitude of the operands [12]. The subjects indicated the sign of a difference by the side on which they tapped. Both the mean taps and the variability were smooth as the difference approached and passed through 0. Thus, the mechanism that subtracts analog magnitudes does not appear to have to engage in any special processing as the differences pass through 0 and reverse their sign. Closure under division requires that there be magnitudes that represent non-integer proportions, including proportions less than one.

Large dynamic range. Weber's law was experimentally established at the dawn of empirical psychology. It has traditionally been seen as a fact about sensory discrimination: the discriminability of two sensed magnitudes (e.g., two weights) is a function of their ratio. However, it turns out to apply just as much to the abstract magnitudes of number and duration, whose apprehension is not rooted in sensation [8, 13]. This suggests that it is a fundamental aspect of the brain's machinery for representing magnitude. What it implies about that machinery is, however, still not understood. Is it a feature or a bug? In thinking about this question it is important to bear in mind that one constraint on the encoding of magnitude is that it function over a very large range: 10 orders of magnitude or more, because the distances and durations that must be encoded range from fractions of a centimeter and fractions of a second to thousands of kilometers and years.

One ancient but still widely advocated explanation is that the mapping from an objective magnitude to its subjective counterpart is logarithmic [14, 15]. This could be seen as making Weber's law a feature rather than a bug, because a logarithmic mapping would enable a brain mechanism with a limited dynamic range to represent a much larger objective range. However, logarithmic mapping makes computation problematic: Unless recourse is had to look-up tables, there is no way to implement valid arithmetic addition and subtraction, because the addition and subtraction of logarithmic magnitudes corresponds to the multiplication and division of the quantities they represent [see 16 for an experimental exploitation of this]. Special processing at 0 will obviously be required because the logarithm goes to infinity as the magnitude goes to 0, and dealing with sign (direction) will be a problem, because there logarithm of a negative magnitude is not defined. Logatihms do, of course, have sign, but the negative logarithms represent the proportions between 0 and 1, so they cannot be used to represent negative magnitudes.

An alternative explanation is scalar variability: the noise or variability in the representation of magnitude is proportionate to the magnitude [17]. This would make Weber's law a bug, not a feature: mental variability (noise) is simply proportional to magnitude, as it often is for physical quantities. The problem with this suggestion is that it would require a dynamic range in the physical realization of the encoding for which it is hard to imagine a plausible mechanism.

There is a third explanation: autoscaling. Measuring instruments generally have limited dynamic range in their output (a limit on the display screen, a limit on the dial excursion, a limit on the number of digits in the read-out, etc). It seems likely that the brain's mechanism for representing magnitude is similarly limited. Measuring instruments convey information about quantities over a much larger dynamic range by adjusting their sensitivity to the magnitude of the input signal: the stronger the signal, the lower the sensitivity. Autoscaling keeps the representation of the input within the dynamic range of the instrument. The scale adjustment makes the minimally representable difference proportional to signal strength, but autoscaling does not implement a logarithmic mapping from input to the output. At any scale factor, the mapping from input to output is scalar, which means that the full range of arithmetic operations may be validly performed on this representation of the input. The autoscaling explanation puts Weber's law in the feature category. The system is so designed that the amount of information the representation of the signal conveys about it is independent of the magnitude of the input--a highly desirable design feature.

The autoscaling of sensitivity to stimulus strength has been shown to operate at the single neuron level with an efficiency close to that imposed by physical considerations [18]. The spike train in a single axon conveys a linearly decodable encoding of the yaw waveform (the back and forth swings of the visual scene), while at the same time signaling the scale factor. The scale factor changes quickly as the fly passes from turbulent conditions, which produce large yaws, to calm conditions, which produce only small yaws.

The autoscaling explanation suggests that the properties of mental magnitudes have been shaped over evolutionary time by the role that they play in conveying information about quantities in the world. This role imposes a two-fold constraint that enables us to say with unusual precision and explicitness how to recognize when one has found it in the machinery of the brain. The physical realization of mental magnitudes in the brain must allow them to enter efficiently into the basic arithmetic operations [19]. And, the mapping from quantities in the world to the mental magnitudes that represent those quantities must be such as to make the results of arithmetic processing validly applicable to the represented world, because the accurate representation of the behaviorally relevant aspects of that world is the function that has shaped the evolution of the arithmetic processing mechanism. This double constraint is particularly salient when one considers the unique mental magnitude that must serve both as the multiplicative identity element and as the representative of sets of numerosity one.

Combinatorial signal processing. Attempts to identify the neural mechanisms mediating the brain's representation of the experienced world must address both aspects of an effective representation: (1) reference, that is, the mapping from an objective aspect of the experienced world (e.g., the numerosities of sets) to the neural signal represents it; (2) compositionality, the combinatorial processing of the symbols (e.g., the mechanism that computes the difference in two numerosities or the range and bearing of one location from another location). Most neurobiological work focuses only on reference, but there is no point in establishing reference if the referring symbols are not be processed so as to make valid inferences about that to which they refer.

The arithmetically necessary signal-combining properties of the multiplicative identity element (the neural symbol for 1) should make it easy to recognize (Box 2). Moreover, it must also be the mental representative of numerosity one, a foundational element in the system of discrete quantity [20]. The scale factors relating mental magnitudes to the quantities they represent cannot generally be specified a priori. However, the mapping from numerosity to mental magnitudes is constrained by analytic considerations: numerosity one must map to the multiplicative identity element or the mental books won't balance. A constraint on any useful symbol-processing system is consistency. If numerosity one maps to any magnitude other than the multiplicative identity, arithmetic reasoning will be inconsistent. Suppose, for example, that numerosity one mapped to a magnitude that was not the multiplicative identity, say, for example, a magnitude twice the multiplicative identity. This would establish the scale factor for the mapping from numerosity to mental magnitudes: numerosity two would map to a magnitude 4 times the multiplicative identity, numerosity three to magnitude 6 times as great, and so on. Consider the effect of two operations on the mental magnitude for numerosity one that ought to have equivalent consequences: adding one or doubling it. Adding one (2+2) will give the magnitude that represents numerosity two, but doubling it—multiplying the magnitude that represents numerosity one by the magnitude that represents numerosity two (2 x 4)—gives the magnitude that represents numerosity four. Therefore, for analytic reasons, the scale factor relating numerosity to mental magnitude must itself be one (the multiplicative identity element).

Conclusions

It seems likely that magnitudes (computable numbers) are used to represent the foundational abstractions of space, time, number, rate, and probability. The growing evidence for the arithmetic processing of the magnitudes in these different domains, together with the "unreasonable" efficacy of representations founded on arithmetic, suggests that there must be neural mechanisms that implement the arithmetic operations. Because the magnitudes in the different domains are interrelated--in for example, the representation of rate (numerosity divided by duration) or spatial density (numerosity divided by area)--it seems plausible to assume that the same mechanism is used to process the magnitudes underlying the representation of space, time and number. It should be possible to identify these neural mechanisms by their distinctive combinatorial signal processing in combination with the analytic constraint that numerosity 1 be represented by the multiplicative identity symbol is the system of symbols for representing magnitude.

Box 1. Why arithmetic is special

Eugene Wigner [21] called attention to "The unreasonable efficacy of mathematics in the natural sciences." Mathematics rests on arithmetic. As Wigner emphasized, representations constructed on this simple foundation are surprisingly successful in representing the natural world. Based on ethnographic and psychological evidence, Fiske has argued that humans, at least, also use mental magnitudes to represent social rank [22]. Figure 1 reminds the reader of the ways in which magnitudes may be used to represent space, time, number, probability, and rate. It uses the lengths of lines in place of number symbols, because length instantiates magnitude.

The representations shown are conventional and elementary. The brain's representations are likely more sophisticated, hence less transparent. See, for example, the Burgess chapter. Nonetheless, they, like those here portrayed, must enable arithmetic processing to be brought to bear in behaviorally useful ways.

In particular, the magnitudes that represent discrete numerosities must be computationally compatible with those that represent continuous variables, because symbols for continuous variables like proportion or probability must be generable from discrete magnitudes by division. Humans spontaneously transfer from a proportion between numerosities to the same proportion between durations **[10]**. Rats multiply reward rates by reward magnitudes to determine average stay durations in a matching task **[9]**. Mice and humans combine trial probabilities and a representation of their uncertainty about elapsing durations to determine optimal target intervals for switching from one option to another **[1]**.

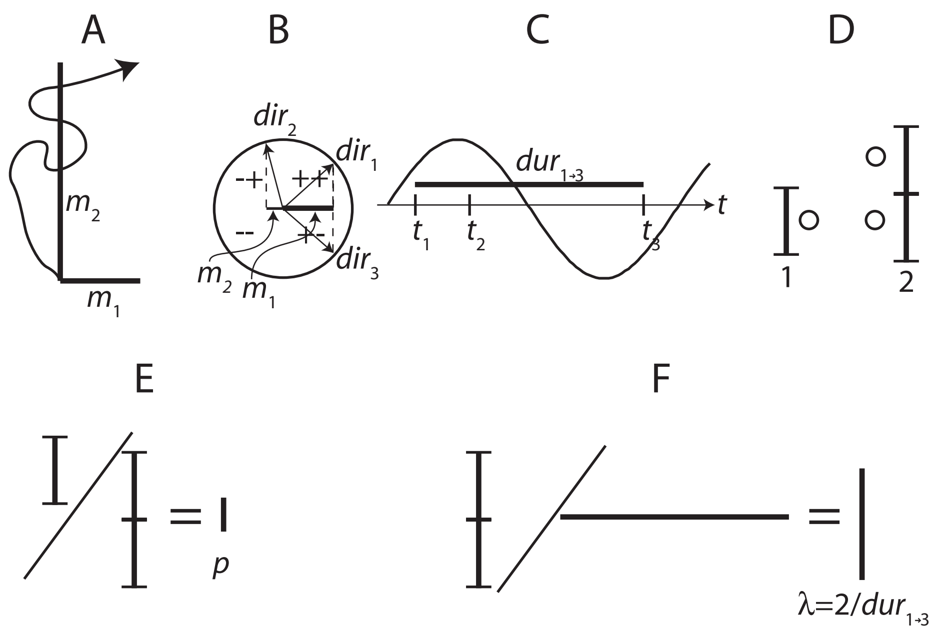

Figure 1. A. the representation or location in two or three dimensions, as in dead reckoning while foraging (trace ending in location arrow), may be mediated by a vector composed of two or three magnitudes (m_1 & m_2). **B.** The representation of direction, which is critical in dead-reckoning, may be reduced to a magnitude proportional to the cosine of the direction angle and two signs. The signs code the quadrant. The magnitude codes direction within it: dir_1 is encoded by $\langle m_1,+,+\rangle$, dir_2 by $\langle m_2,-,+\rangle$ and dir_3 by $\langle m_1,+,-\rangle$. **C.** Durations are represented by single magnitudes ($dur_{1\to3}$), which may be computed from differences in temporal locations (t_1, t_2, t_3): $dur_{1\to3} = t_3 - t_1$. Temporal locations may be represented by the phases of endogenous clocks, like the circadian clock **[23]**, and phase may be represented in the same way as directions: $\langle m_1,+,+\rangle$, $\langle m_1,+,+\rangle$ and $\langle m_1,+,-\rangle$ could as readily represent t_1, t_2, and t_3 as dir_1, dir_2 and dir_3. **D**. Numerosity is also represented by analog-like magnitudes **[see for review 24]**. **E.** Dividing magnitudes representing numerosity (discrete quantity) generates magnitudes representing probability and proportion (continuous quantities). **F.** Dividing agnitudes representing numerosity by magnitudes representing duration yields magnitudes representing rates.

Box 2. The Signal Processing Signature of the Multiplicative Identity Element

Identifying the physical mechanism of the mental magnitudes will go hand in hand with identifying the neurobiological machinery for the arithmetic processing of those magnitudes. That machinery must implement the basic arithmetic operations: addition, subtraction, multiplication, division, ordination and negation. In all but the last of these operations, the machinery must combine two inputs (two mental magnitudes) to produce an output that is itself capable of being processed in the same way by that same machinery, an output that is itself a mental magnitude. One mental magnitude, the multiplicative identity element, must have distinctive properties that make it easy to recognize, both by its behavior when input to the arithmetic processing machinery and by mapping from numerosities to it. Its behavior is illustrated in Figure 1. The reasons why it must be the representative of numerosity one are explained in the text.

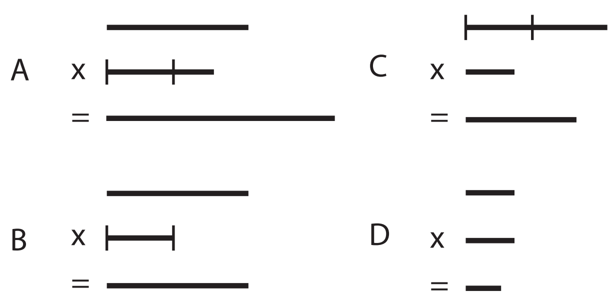

Figure 1. The behavior of the multiplicative identify element in the multiplication operation, the element whose magnitude is delimited by the two short vertical strokes. **A**. Any two magnitudes greater than this produce a magnitude greater than both of them. **B**. The product of the identity element and any other magnitude is the other magnitude. **C.** The product of an element less than the identity element and an element greater than the identity element is intermediate between the two. **D**. The product of two elements less than the identity is smaller than both.

Box 3: Unanswered Questions

A pressing question, the answer to which is likely to have profound consequences for neuroscience, is the coding question: What is the enduring physical change that encodes experienced magnitudes (experienced numerosities, experienced durations, experienced distances, experienced relative frequencies, etc), carrying this information forward in time in a computationally accessible form. It is very widely assumed that information is encoded in the nervous system in the form of altered synaptic conductances. However, there is almost no discussion of how altered synaptic conductances could code some particular piece of information, and there are reasons to be skeptical that “associative” changes in synaptic conductances are suitable for this purpose [25]. Focusing on how magnitudes may be enduring encoded is likely to be fruitful for two reasons: First, it poses the question very sharply, to wit, “What is the physical change by which individual neurons or perhaps neuronal circuits encode a computable number?” Second, a mechanism that is capable of encoding a number is capable of encoding any kind of information whatsoever. Thus, answering this question may answer the question how acquired information is stored in the nervous system, which is arguably the single most important question in behavioral neuroscience.

Bibliography

1. Balci, F., D. Freestone, and C.R. Gallistel, *Risk assessment in man and mouse.* Proceedings of the National Academy of Science U S A, 2009. **106**(7): p. 2459-2463.
2. Gallistel, C.R., *The organization of learning*. 1990, Cambridge, MA: Bradford Books/MIT Press. 648.
3. Lourenco, S.F. and M.R. Longo, *General magnitude representation in human infants.* Psychological Science, 2010. **21**(6): p. 871-881.
4. Xu, F. and V. Garcia, *Intuitive statistics by 8-month-old infants.* Proceedings of the National Academy of Sciences (USA), 2008. **105**(13): p. 5012-5015.
5. Vallortigaria, G., et al., *Rudiments of mind: Insights through the chick model on number and space cognition in animals.* Comparative Cognition and Behavior Reviews, 2010(5): p. x-xx.
6. Walsh, V., *A theory of magnitude: Common cortical metrics of time, space and quantity.* Trends in Cognitive Sciences, 2003. **7**(11): p. 483-488.
7. Clayton, N., N. Emery, and A. Dickinson, *The rationality of animal memory: Complex caching strategies of western scrub jays*, in *Rational Animals?*, M. Nuuds and S. Hurley, Editors. 2006, Oxford University Press: Oxford. p. 197-216.
8. Gibbon, J., *Scalar expectancy theory and Weber's Law in animal timing.* Psychological Review, 1977. **84**: p. 279-335.
9. Leon, M.I. and Gallistel, *Self-Stimulating Rats Combine Subjective Reward Magnitude and Subjective Reward Rate Multiplicatively.* Journal of Experimental Psychology: Animal Behavior Processes, 1998. **24**(3): p. 265-277.
10. Balci, F. and C.R. Gallistel, *Cross-domain transfer of quantitative discriminations: Is it all a matter of proportion?* Psychonomic Bulletin and Review, 2006. **13**: p. 636-642.
11. Barnet, R.C., R.P. Cole, and R.R. Miller, *Temporal integration in second-order conditioning and sensory preconditioning.* Animal Learning and Behavior, 1997. **25**(2): p. 221-233.
12. Cordes, S., et al., *Nonverbal arithmetic in humans: Light from noise.* Perception & Psychophysics, 2007. **69**(7): p. 1185-1203.
13. Moyer, R.S. and T.K. Landauer, *Time required for judgments of numerical inequality.* Nature, 1967. **215**: p. 1519-1520.
14. Dehaene, S., *Verbal and nonverbal representations of numbers in the human brain*, in *Galaburda, Albert M (Ed); Kosslyn, Stephen M (Ed); et al (2002) The languages of the brain*. 2002, Harvard University Press: Cambridge, MA. p. 179-190.
15. Fechner, G., *Elemente der Psychophysik*. 1860, Leipzig: Breitkipf & Härtel.
16. Gibbon, J. and R.M. Church, *Time left: linear versus logarithmic subjective time.* Journal of Experimental Psychology: Animal Behavior Processes, 1981. **7**(2): p. 87-107.
17. Gibbon, J.G., *Scalar expectancy theory and Weber's law in animal timing.* Psychological Review, 1977. **84**: p. 279-325.
18. Fairhall, A.L., et al., *Efficiency and ambiguity in an adaptive neural code.* Nature, 2001. **412**(6849): p. 787-92.

19. Dehaene, S., et al., *Arithmetic and the brain.* Current Opinion in Neurobiology, 2004. **14**(2): p. 218-224.
20. Leslie, A.M., R. Gelman, and C.R. Gallistel, *The generative basis of natural number concepts.* Trends in Cognitive Sciences, 2008. **12**(6): p. 213-218.
21. Wigner, E., *The unreasonable effectiveness of mathematics in the natural sciences.* Communications in Pure and Applied Mathematics, 1960. **13**(1).
22. Fiske, A.P., *Four Modes of Constituting Relationships: Consubstantial Assimilation; Space, Magnitude, Time and Force; Concrete Procedures; Abstract Symbolism* in *Relational Models Theory: A Contemporary Overview*, N. Haslam, Editor. 2004, Earlbaum: Mahway, NJ. p. 61–146.
23. Crystal, J.D., *Nonlinearities in sensitivity to time: Implications for oscillator-based representations of interval circadian clocks*, in *Functional and neural mechanisms of interval timing*, W.H. Meck, Editor. 2003, CRC Press: New York. p. 61-76.
24. Gallistel, C.R. and R. Gelman, *Mathematical cognition*, in *Cambridge handbook of thinking and reasoning*, K. Holyoak and R. Morrison, Editors. 2005, Cambridge University Press: New York. p. 559-588.
25. Gallistel, C.R. and A.P. King, *Memory and the computational brain: Why cognitive science will transform neuroscience*. 2009, New York: Wiley/Blackwell.