What conditions should a theory of consciousness meet?

Bernard J. Baars
SUNY Stony Brook

This paper is written in the conviction that theories of consciousness today are in the same situation that semantic theories were in a decade ago. At that time there was a widespread belief that semantics was an essentially insoluble puzzle, that much more data would have to be collected to constrain adequate theory. In the event, it turned out that the real obstacles were conceptual rather than empirical. Once the actual conditions to be met by an adequate semantics are specified, the theoretical options are vastly reduced. Further, conditions on semantic theory were not difficult to find: all such theories must be able to handle discourse reference, paraphrase generation, question answering, the detection of anomaly and contradiction, and the ability to resolve ambiguities at all levels of analysis. Until people looked at these actual constraints, no progress could be made, and the problem was treated as insoluble. Similarly, until we actually look at widely-accepted constraints on theory of consciousness, the topic will be treated as fuzzy, mystical, and insoluble.

What are the conditions for an adequate theory of consciousness? First, we can specify some pre-theoretical criteria:

1. An adaptive construct. Consciousness should be treated as a cognitive construct much like any other, with an adaptive information-processing function.

2. Relationship to other constructs. The proposed construct should be distinct from others, but explicitly related to perception, memory, intentionality, executive functions, automaticity, availability, the "internal monologue", the subjective observer, and especially attention.

3. What is UNconscious? A theory of consciousness should give a (principled) explanation of the difference between conscious and unconscious processes.

4. Respect for common-sense psychology. There is a world of difference between bootstrapping one's way beyond common sense and blindly ignoring it. Common sense is our starting point, and we cannot even ask about the nature of consciousness without it.

5. Empirical reference. Finally, a clear, empirically-based domain should be specified. Tables 1 and 2 show a number of contrasting pairs of claims about similar conscious and unconscious processes. These claims command a wide consensus among psychologists. The job of theory is then to fit some explanatory model to the constraints in the simplest possible way.

Table 1: Capability Constraints on a theory of conscious contents.

Conscious Processes	vs. Unconscious processors
1. Computationally inefficient.	Highly efficient in specialized tasks.
2. Great range, & relational capacity.	Limited domains & relative autonomy.
3. Apparent unity, seriality, & limited capacity.	Very diverse, parallel, and together have great capacity.

Table 1 shows some Capability Constraints, which indicate the capacities and limitations of conscious vs. unconscious phenomena. For example, if we use the term "computational efficiency" to mean the ability to work out some algorithm quickly and without error, it is clear that wholly conscious processes are not computationally efficient. Even simple addition or subtraction is performed slowly and with a good chance of error. It appears that the great bulk of fast and efficient processing is done by a large set of specialized processors. There is much neurophysiological evidence to this effect as well (Geschwind, 1979).

The great range and diversity of conscious contents seems to compensate for these efficiency limits. People can be conscious of virtually any energy pattern impinging upon any sensory system, down to single protons hitting single visual receptors. By means of conscious biofeedback, one can gain voluntary control over the actions of two-neuron spinal motor units. Etc. Or one can be conscious of events that require enormous cooperative activity between many millions of neurons.

Further, conscious contents always appear internally consistent at any one time, even if this consistency is spurious. This is in agreement with the fact that conscious capacity is limited and that the contents of consciousness appear serially. These facts belong together: If there must be internal consistency at any one time, then there is a clear capacity limit for incorporating mutually inconsistent contents, and such mutually exclusive contents must also appear serially.

In contrast to conscious processes, there is reason to think that unconscious processors can operate autonomously in their specialized domains without difficulty, because they are isolated from each other. They seem to be highly diverse, operating fast, efficiently, and in parallel. Taken as a whole, the set of all unconscious special-purpose processors has a very great capacity indeed.

These Capability Constraints have led me to associate consciousness with a well-known information-processing configuration: A global data-base, operating in a very large, distributed system. The global data base is essentially a central information exchange which permits otherwise autonomous specialized processors to interact with each other. Representations in the global data base are globally distributed, so that any one of a myriad specialists can respond to the global information, and some set of specialists can cooperate in return to create another global representation.

In this system, global processes are inefficient, slow, and error-prone because they require cooperation between different sets of specialists. Yet global information will have great range, diversity and context-sensitivity precisely because it involves interaction between many specialized processors in the system. Global information will show apparent unity, because inconsistent global representations will lead to competition between mutually exclusive specialists, which will cause the global representation to become rapidly unstable. There will thus be a narrowly limited capacity to display mutually competitive contents at any one time, and these will have to be displayed serially. In this way, all the Capability Constraints of Table 1 can be shown to apply to global representations in a distributed system. But this is not the whole story.

Table 2: Boundary constraints on the contents of consciousness.

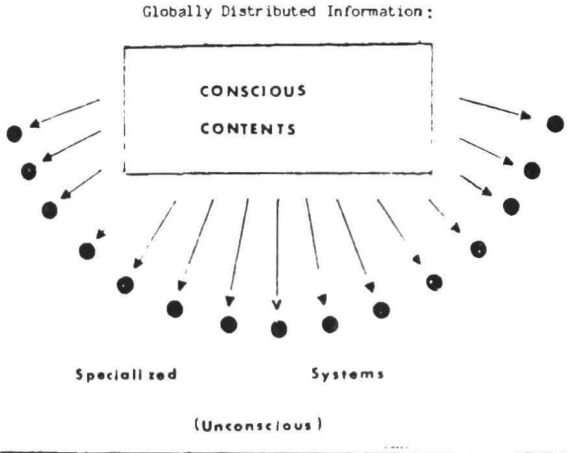
Conscious phenomena	Unconscious phenomena
Simultaneous cases:	
1. Percepts.	Context needed to organize percepts.
2. Input consistent with context.	Input inconsistent with context.
Diachronic cases:	
3. Percepts.	Pre-perceptual processes.
4. Any change in a habituated stimulus.	Habituated percept.

Consider the conscious-unconscious contrasts in Table 2, entitled Boundary Constraints --- a well-known set of facts about perception. Percepts are conscious, but the rapid pre-perceptual hypothesis-testing that is needed to establish the conscious percept is unconscious. Furthermore, the physical stimuli that lead to the conscious percepts are only conscious if they are defined within a stable set of contextual constraints. These contextual constraints are needed to provide the presuppositional background for the percepts, and they are not conscious. The physical stimuli themselves are not conscious either if the appropriate contextual background is missing. Further, percepts that are thoroughly analyzed drop out of consciousness (habituation and automaticity).

These facts imply that conscious contents involve more than just global information. They need to be stable and coherent , to accomodate the fact that rapid pre-perceptual processing is not conscious. Further, for a global representation to be conscious, it must be able to trigger widespread adaption in the nervous system --- once this adaptation has occurred, the conscious percept becomes unconscious, presumably because it has now become a part of the stable, presuppositional background. To put it all together, then: conscious contents must be global, coherent, and informative .

A first approximation to a system that fits these constraints is shown in Figure 1.

Figure 1

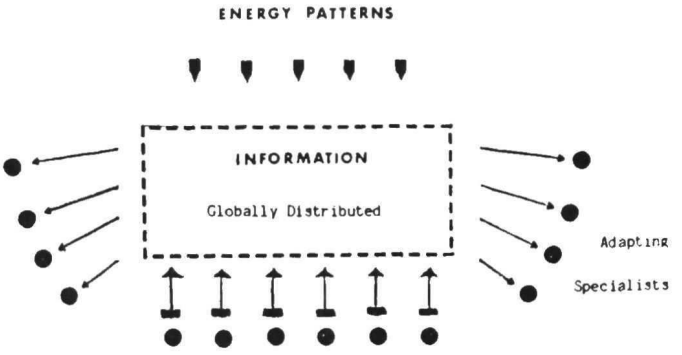


Note that conscious contents are globally available, but most detailed information processing is performed locally by a large set of specialized, distributed processors. The specialized processors maintain the processing initiative.

Conscious contents are not only global, but als coherent, because internally inconsistent contents would imply competition between different sets of specialists. Any active competition would rapidly remove the incoherent representation from the global data base.

Figure 2 presents a more refined system to fit the constraints set out in Tables 1 and 2.

Figure 2.



The global representation is now shown to be bounded by a set of stable contextual specialists which provide the presuppositions within which the global representation is defined. However the contextual specialists are not themselves conscious. Other specialists are in the process of adapting to the global representation. Once adaptation is complete, the global representation fades from consciousness; it becomes part of the contextual background, though it continues to constrain other conscious contents.

This final model can handle all the major empirical constraints on theory listed above: Conscious contents must be global, internally coherent, and informative (i.e., able to trigger widespread adaptation).

This theoretical approach is treated in much greater detail in Baars (in press) and Baars and Mattson (in press).

Baars, B.J. Conscious contents provide the nervous system with coherent, global information. In R.J. Davidson, G. Schwartz, and D. Shapiro (eds.) Consciousness and self-regulation Vol. III , NY: Plenum, in press.

Baars, B.J. and Mattson, M.E. Consciousness and intention: A framework and some evidence. Cognition & Brain Theory , in press.

Geschwind, N. Specializations of the human brain. Scientific American, 241 (3) .1979,180-201.