Journal of Cultural and Evolutionary Psychology, 3(2005)1, 87–105

THE PARADOX OF NARRATIVE THINKING

FRANCIS F. STEEN
University of California Los Angeles

Abstract. Why do human beings show such a strong preference for thinking in narratives? From a computational perspective, this method of generating inferences appears to be exorbitantly wasteful. Using students' responses to the fairy tale of Little Red Riding Hood, I argue that narrative comprehension requires the construction of idiosyncratic imagery, but that the cognitive yield is structural and shared. This peculiar method of information processing, I suggest, is the outcome of evolutionary path-dependence. The narrative mode of construal is an expert system taking its input from the display of conscious experience, but producing results that are largely unconscious. Drawing on examples from rhesus play, I argue that the core features of narrative thinking have biological roots in strategy formation. Finally, I return to the fairy tale to illustrate the operation of a series of peculiar design features characteristic of human narrative thinking.

Keywords: narrative, evolutionary theory, consciousness, global workspace theory, simulation, pretense, play, unintelligent design theory

INTRODUCTION

We tell each other stories every day, effortlessly, without stopping to wonder what it is we are doing. Yet the ubiquitous practice of narrative is a remarkable human achievement. Through narrative, the young child reaches out to a new friend, soliciting contact through the confirmation of a shared imaginative reality; the lawyer pleads the innocence of his client by arousing the emotions of the jury; the storyteller entertains his audience by conjuring up absent or imaginary men and beasts. In the prototypical narrative, we establish relations between the actions of social agents, accounting for outcomes, linking causes to effects, and assigning credit and responsibility.

In cognitive terms, forming a narrative is an act of connecting a succession or mere co-occurrence of agents and objects into a causally ordered, intuitively graspable whole. The work of the narrator generates a structure that can be reused – as BATESON (1980) put it, "a pattern that connects", a prototype of understanding and

intelligibility. Listeners can selectively uncompress the meaning of this prototype by projecting it back onto their own real and imagined pasts and futures. In this manner, narratives can function as abstract models that structure, simplify, and lay out causal connections among otherwise indefinite and unintelligible events. By trying out a succession of stories on a particular situation and seeing which of them most credibly generate and fit with the observed facts, narratives function as testable hypotheses in heuristic thought experiments. In these and other ways, narrative helps us orient ourselves as agents in a complex natural and social world and make our experience meaningful (BRUNER 1987).

The presence of material that has a narrative format is commonplace in any compendium of anthropological source material, such as the Human Area Relations Files. From autobiographies and histories to folktales and myths, the *story* is used to remember, to persuade, to entertain. Proverbs, found in nearly all cultures, present moral lessons that must be unpacked into a narrative to be understood (HERNADI and STEEN 1999). The pervasive presence of narrative in human cultures throughout history indicates that the capacity to generate stories is tightly integrated into the everyday operation of the human mind. This integration expresses itself in part as a spontaneous preference for narratives over other forms of symbolic representation, such as mathematics or object mechanics, in part in the wonderful ease and skill with which people make sense of their world by narrating it. Through narrative, we know how to make impressively effective use of the information we glean from a situation.

This claim will strike many people as trivially self-evident. However, it remains a puzzle that human beings should be so designed. In what sense are narratives computationally effective? Stories are not a favored form of representation for our artificially constructed thinking machines, even though computers are built by people and might by default be expected to share the unconscious cognitive biases of their designers. Unlike people, computers do not think it terms of stories. BATESON (1980) proposed this capacity is diagnostic of human intelligence, imagining a future moment in which the computer scientist asks his computer, “Do you compute that you will ever think like a human being?” If, after a long pause, the machine were to respond, “That reminds me of a story...”, and proceed to illustrate, in a manner that combined the idiosyncratic stamp of individual experience with a generic and universally understood narrative structure, a specific case in which this question received – implicitly or explicitly – a meaningful answer, then and only then might we be justified in claiming that machines have achieved our level of intelligence. The fact is that today’s vastly powerful computers employ nothing like the narrative method for organizing, storing, and communicating information, or for generating inferences, highlighting the oddity of the human reliance on stories.

In the following, I suggest that the spectacular narrative performances we see in every human culture is made possible by a complex suite of well-established and

tested adaptations with a deep biological history. In a nutshell, I argue that narrative in its elementary form is an evolved mode of construal, a systematic method for making sense of specific aspects of existence, notably those that involve the task of predicting what agents will do. This mode of construal plays a key role in interpreting as well as in generating strategic action, in play and pretense as well as in functional interactions. Finally, we find it in the sophisticated mental simulations that are the hallmark of human cognition. Cultural uses of narrative are able to piggyback on and recruit a set of neurobiological circuits that were subject to natural selection over various periods, some relatively recent and others stretching all the way back to the early mammals. It is this continuity of function, I argue, that produces the paradox of narrative thinking: the simultaneous juxtaposition of the universal preference for narrative as an efficient and effortless method of organizing information, and the cognitive analysis suggesting it is in computational terms extravagantly expensive and for many purposes strikingly inefficient. Using my students' responses to the fairy tale of Little Red Riding Hood, I begin by giving an illustrative analysis of the paradox of narrative thinking, focusing on the fictive variety. I then develop a theoretical model that identifies the narrative mode of construal as an expert system taking its input from the display of conscious experience. Drawing on examples from rhesus play, I argue that the core structure of narrative fictions has deep biological roots. Finally, I return to the fairy tale to illustrate the operation of a series of impressive yet idiosyncratic design features of the human mind.

LITTLE RED RIDING HOOD

Prompted by the wolf, Rotkäppchen "opened her eyes and saw the sunlight breaking through the trees and how the ground was covered with beautiful flowers" (GRIMM 1812). As she does, the listener conjures up in his mind's eye images of what it is she sees. These images must be constructed out of the reader's memories, and will vary from individual to individual. When I asked my undergraduate students in one class to report on their subjective phenomenologies in response to this sentence, the accounts varied dramatically. In one person's mind, the scene was reconstructed from above, looking down on the little blonde girl standing amid the green grass, who was gazing up on the beams of sunlight illuminating the flowers ahead of her, to the left of the path, which led off the frame at around 110 degrees. In another's, the scene was seen from a position ahead of the little girl on the path, observing her brown ringlets as she looks for flowers on either side. In yet another's, the perspective was that of the little girl herself, the light coming down in shafts through dense fir trees, sheltering clusters of deep blue wood anemones. The forest itself was typically the listener's prototypical forest, depending on their childhood experiences –

light birchwoods, or dense redwoods, in some cases a specific location the student remembered, and into which he or she placed the little girl. Some of them, having spent their entire lives in large cities, reported seeing a cartoon-like forest, of a kind they had seen in the movie theater. In brief, if we could plug a television into the visual cortex of each listener and display their mental movies in a bank of screens, each monitor would show a different movie.

In spite of the fact that not a single frame will be identical, however, the listeners to the story of Little Red Riding Hood will be very confident that they have all heard and understood the same story and participated in a common and shared experience. For the purpose of comprehending the story, it appears that the specifics of the mental imagery can be overlooked. This is the central puzzle of narrative: different movies, and yet with great conviction the same story.

On all the television screens, there will be a little girl with red headgear walking alone in a forest, picking flowers. It will not be the same girl or the same forest in any normal sense of the word "same" – that is to say, each screen will display an individual with a distinctive and different set of features, and not a single tree or flower will stay constant across two monitors. Yet these differences are not attended to; they are not subject to communication, and this startling fact is in turn not even noticed. Such differences in mental imagery are not a matter of lacking precision, a subtle misunderstanding, or missing information – rather, the details are in a radical sense not part of the story. In spite of the rich internal phenomenology the story generates, they carry little or no part of the significance. At a simple propositional level, they do not mean anything.

This is an extravagant state of affairs. When the brain is generating composite, moving, three-dimensional images on the fly, it is doing something that the most sophisticated computers still struggle to accomplish. Hollywood does produce 3D movies, but these are painstakingly put together frame by frame and then converted into moving images using renderfarms of hundreds of clustered computers. The software and computational power required to generate a coherent movie using several sources of stills and episodic images, modified instantly to fit the scene, and composited on the fly, are either not available technologies or beyond the budgets even of the big studios. Now, the brain is doing this at the drop of a hat – and for the purpose of a shared understanding, the whole show do not seem to matter one bit.

Narratives are individually instantiated in sophisticated and detail-rich inner worlds. We do not talk about them, and the attempt to do so may lead to discomfort, as if we were exposing to social comparison and competition an intrinsically private and emotionally treasured world. One might infer, of course, that two people in fact do not ever reach a shared understanding of a story, and instinctively paper over their differences to minimize unpleasant disagreement, or worse, cover up the terror of an unbridgeable chasm between solipsistic minds. Yet, we really have no reason

to think that countless generations of children and their grandparents have had any problems reaching a shared understanding of the story of Little Red Riding Hood. The reason they do not talk about their private imagery is simply that this does not contribute to the common understanding of the narrative; it would only introduce a distraction. Rather than assuming a solipsistic understanding of narrative, where each listener is locked into his or her impenetrable world of private imagery, we need to account for the mind's ability to understand a story in terms of an underlying structure, a structure that in itself is uninstantiated in consciousness, a set of abstract relations.

THE SEARCH FOR NARRATIVE STRUCTURE

A large literature from PROPP (1928) onwards examines the distinction between a set of underlying narrative structures and their surface manifestation in particular stories. In the search for the common structures present in the instantiation of a large number of stories, the more basic question is hardly ever asked: what is it that allows two or more people to derive a shared common structure out of their individually idiosyncratic instantiations of the same story? To pursue the computer analogy, what we have to explain is how the central elements of the narrative can remain invariant across the multiple television screens that we have plugged into the TV-out port of our listeners' brains. It would be extremely difficult for a computer program to pull out this structure, as not a single pixel remains invariant across two screens. It is not a matter of seeing the same object from different angles, at different distances or speeds, under different types of lighting. These are also demonstrably different objects, at best vaguely similar. To pull the story out of these displays, you would need highly sophisticated dedicated equipment, far exceeding anything currently in existence in visual analysis software.

Of course any child would do fine at this task of understanding the story. She would know, without realizing that she knew, that stories are about people facing some difficulty, and needing to come up with a way of overcoming this difficulty. Seeing a movie of Little Red Riding Hood, her mind would effortlessly abstract these elements and make a series of complex inferences regarding appropriate strategies of action in some class of similar circumstances. Listening to the story read aloud, she would generate her own imagery and utilize it in a similar manner to generate inferences – to construct a series of implicit morals, similar to the one that is explicitly provided in PERRAULT's 1697 version of the tale. Human minds are wired precisely in the requisite manner to solve this particular task – in fact they show a perverse preference for processing information in this roundabout fashion.

The point could be put even more strongly. The child appears to rely on a mental representation of the agent (Little Red Riding Hood), the setting (the forest), the goal

(survival), the obstacle (the wolf) and the little girl's resources (her imperfect understanding of the danger) in order to generate the appropriate inferences. Even though the details of the mental imagery do not matter, the imagery itself is mandatory. That is to say, the story must undergo a peculiar form of processing to be understood.

But why would natural selection produce a computational device that uses moving 3D images to make inferences about something that has nothing to do with the specifics of the images produced? It would be one thing if you generated complex imagery that would visualize the information for you, but this is not at all what is happening in narratives – on the contrary, almost every feature of the visualized information has nothing to do with the target inferences. Only on a very abstract level can we say that the mental imagery instantiates the structural components in such a way as to make them intelligible. The general puzzle is, why do we want information presented in a narrative format? In what sense could this possibly be an efficient method of information processing? It is as if a computer, to make the right inference about 2 + 2, had to represent the numbers as dancing penguins trying to escape a polar bear. Inherent in its processing loop is the requirement that it generate moving three-dimensional dancing penguins, and when this display is sensed by some second component, it generates the correct inference: 4. While it is absurd from a computational perspective, showing dancing penguins may not be a bad procedure for teaching very young children how to add. This tells us something important about human cognition – but what exactly?

Similarly, the structural lesson that a child abstracts from a narrative may lend itself to a rather elementary representation. In the case of the fairly tale, for instance, let's say the target inference is "Do not stray from the path" and "Do not stop to talk to wolves". Or you might argue the inferences are really more complex; PERRAULT (1697), for instance, suggests that the "wolf" must be understood as a friendly but wicked man out to take advantage of young girls. Yet however complex these morals are, it is not hard to state them briefly. Why not skip the story and just provide the moral? Why this fantastically circuitous route?

THE ARCHITECTURE OF SENSORY CONSCIOUSNESS

The brain we have today is an accretive structure, a result of a series of successive innovations added onto a continuously operating plant. New capacity is typically built as extensions of existing operations, constraining the scope of available solutions. Due to path dependence in the evolutionary history of adaptations, computational solutions that in a global perspective would be the most efficient may well become unavailable, leaving far less optimal methods as the most efficient. When narrative, in spite of its extravagant resource consumption, remains a favored use of the human cognitive machinery, the causes must be sought in our biological history.

At a very basic level, the central role of the brain is to enable and dispose the organism to respond to its environment in a manner that promotes survival and reproduction. The simplest way to accomplish the task of connecting sensory information to appropriate action is a response system that is triggered by a particular range of stimulus values. Under adaptive pressure, genetic mutations may arise that build cognitive mechanisms to broaden the scope of the data acquired, improve its quality, and produce a better targeted response. In mammals, sensory data acquisition systems are complemented by perceptual processing systems that refine the incoming data stream, extracting meaningful patterns, mapping the data onto a spatial grid, adding color, and multiplexing sound, vision, smell, and touch, and outputting the result to what may loosely be spoken of as a locus of subjective phenomenology, a perceptually-based form of consciousness.

As EDELMAN (1992) explains it, primary consciousness is “a state of being aware of things in the world”, one in which we experience “a ‘picture’ or ‘mental image’ of ongoing categorized events” (p. 12). This elementary form of consciousness, present in mammals, in itself “lacks an explicit notion or a concept of a personal self, and it does not afford the ability to model the past or the future as part of a correlated scene” (p. 24), in contrast to the distinctive form of “higher-order consciousness” characteristic of humans.

EDELMAN’s account implies that what appears in sensory consciousness is already *categorized:* an initial level of processing has already taken place prior to any conscious experience. Our subjective phenomenology as we open our eyes and perceive the world is that it is not a confused jumble of sounds, lights and colors; it consists of an ordered world of objects in space and time. This initial construction of conscious experience, however, is not the highest level of analysis available to us. Thus, in the case of visual processing, MARR (1982) showed that conscious experience is characterized by what he called a “$2^1/_2$D sketch”. The world as we perceive it is a vast improvement over the two-dimensional image that falls on our retinas, but its full three-dimensionality is not represented in consciousness. We know, for instance, that people have back-sides as well as fronts, but our conscious experience presents only what is visible to us from our particular perspective. Sensory consciousness displays the intermediate results of visual analysis; the mind generates higher levels of analysis that do not get displayed.

The critical issue here is the relation between the intermediate and the higher levels of analysis. In MARR’s model, consciousness is not assigned a function; it is simply the receptacle for an intermediate level of visual analysis. At the time MARR published his landmark study, consciousness was not considered a legitimate subject in its own right, but this began to change with the publication of BAARS’ 1988 book, *A Cognitive Theory of Consciousness*. It presents a model that assigns consciousness the function of a “global workspace” (BAARS 2002). Global Workspace theory uti-

lizes a metaphor of the mind as a theatrical production: "Consciousness resembles a bright spot on the theater stage of Working Memory (WM), directed there by a spotlight of attention, under executive guidance (BADDELEY 1993). The rest of the theater is dark and unconscious. "Behind the scenes' are contextual systems, which shape conscious contents without ever becoming conscious ... Once a conscious sensory content is established, it is broadcast widely to a distributed 'audience' of expert networks sitting in the darkened theater..." (BAARS 2003).

In this view, consciousness functions within the mind as a type of display or broadcast. DENNETT and KINSBOURNE (1992) have argued against the notion that consciousness is what they term a "Cartesian theater", a model where the mind's performance of meaning is watched by a homuncular viewer. Such an arrangement, while easily grasped and intuitively attractive, merely displaces the problem of consciousness to the mind of the homunculus, and so on *ad infinitum*. In the Global Workspace model, however, the "audience" of the theater of consciousness is a large number of higher-level inference systems that themselves are unconscious. The "presentation" provides them with the richly and appropriately structured information they use as input and require to operate.

Now, why would these high-level inference systems need a display-type solution? The force of the metaphor of a display is that the coupling is loose: there is a many-to-many relation between the information that is shown on the display and the behavior generated in response. A loose coupling may be favored in situations where there are multiple inference systems working in parallel – if you do not know beforehand which one will give you the response you need, you need to respond to several types of information at the same time, information gathering resources are finite and under pressure, and you must constantly reassign them new relative priorities.

More generally, the need for a broadcast-type functionality is due to the modular architecture of the mind (HIRSCHFELD and GELMAN 1994). A key argument in evolutionary psychology is that natural selection will tend to produce highly specialized cognitive subsystems, each of which is optimized for solving recurring problems within a narrow domain (TOOBY and COSMIDES 1992). A corollary of this argument is that the resulting architecture will over time generate a persistent and pervasive communication and coordination problem within the mind, as different specialized subsystems operate on locally optimized information formats. Global Workspace theory presents a simple solution to this emerging design problem. It implies that consciousness was selected for as a result of the adaptive problem of increasing modularity, which creates obstacles for the effective integration and sharing of information among the different functions within the brain. By adapting to the *lingua franca* of the representational format of consciousness, highly specialized expert systems are given the means to interoperate effectively.

We have no grounds for assigning this architecture to a hominid innovation. Behavior indicative of sensory consciousness – an active search for multiple dimensions of information, accompanied by a state of elevated suspension as multiple interpretations and options are being weighed, and leading to flexible behavioral responses – is common among mammals. It is constructed moment to moment from sense data by "backstage" processes that are fast, informationally encapsulated, and mandatory, as described in FODOR's *The Modularity of Mind* (1983). Rather than "one great blooming, buzzing confusion", as William James famously characterized the world of the infant, our conscious perceptual experience is the fine-tuned product of hundreds of millions of years of mammalian evolution, presenting an orderly world of objects, agents, and events. Recent work in developmental psychology indicates that this is already true for very young infants (BAILLARGEON 1987; GOPNIK, MELTZOFF and KUHL 2000; SPELKE and HERMER 1996).

It is worth distinguishing between the kind of modularity advocated by FODOR (1983), which is restricted to the *pre-conscious* processing of sensory data, and the "massive modularity" (SPERBER 1996) of TOOBY and COSMIDES (1992), HIRSCHFELD and GELMAN (1994), and others, which is focused on *post-conscious* inference systems. Fodorian type modules are not only cognitively impenetrable – that is to say, their processes, though not their products, are unavailable to consciousness – but also informationally encapsulated, or incapable of accepting input from higher-level mental processes (FODOR 1983). While higher-level expert systems, or inference engines, are typically also cognitively impenetrable, they are not informationally encapsulated, in that they accept inputs from consciousness. Indeed, according to BAARS (1988, 2003), and contrasting with MARR's model, consciousness is the *only* source of input to these systems.

For the purpose of clarity in the following argument, I will adopt this strong version of Global Workspace theory and assume that higher-level inference engines, such as the ability to generate a narrative interpretation of events, rely for their inputs exclusively on that global broadcast of processed information we call conscious experience. This simplified account provides us with a powerful tool to understand the paradoxical nature of narrative thinking.

THE NARRATIVE MODE OF CONSTRUAL

Recall our patient subjects, listening to the story of Little Red Riding Hood while wired up to a bank of high-resolution wall-mounted plasma displays that broadcast the content of their visual consciousness to us experimenters. The challenge we began to formulate for ourselves was to design the software, and thus to model the cognitive processes, capable of taking as input the content of one or more of these idio-

syncratically different displays and generating a high-level common understanding of the story. We can now see that the project of working directly on the individual pixels, trying to discover correlations across screens, when the subjects' imaginative construction of the story differ in every detail, is not only a dead end; it is a misformulation of the problem. The preceding discussion of a sensory consciousness allows us to recast the task at hand, making it considerably more tractable. First of all, our subjects must be hooked up to a three-dimensional holodeck rather than a bank of flat screens. The world we experience in the privacy of our minds is not a two-dimensional, unlabeled matrix of sensations. Rather, it is a $2^1/_2$D sketch of a world with three dimensional properties, a world already parsed into Kantian categories of objects and agents, located in space and time. Within this architecture, the narrative mode of construal operates as a high-level expert system that takes its input from this pre-processed and orderly presentation, one in which a girl is walking through a forest.

To peal away a level of fictive remove and mimic this model of sensory consciousness, let us imagine that the forest is real, and that our subjects are perched in the trees. Privileged observers of the fairy-tale drama, they broadcast their conscious experiences via wireless broadband to our observational holodeck at central command: Operation Little Red Riding Hood. The holographic projections now show what any person would recognize as the *same* scene, only viewed from multiple different angles. What is experienced in consciousness is the little girl, Little Red Riding Hood, a path on which she walks, a meadow full of flowers, a forest of trees. The point is that this is the first we see: there are no *priors* to this experience, no conscious display of the complex Fodorian processes whereby the gestalt of the individual trees and flowers, the path and the approaching wolf, are abstracted from the matrix of sensory perception and presented in consciousness. What we see on the holodeck is a coherent world ready to be interpreted. The narrative mode of construal is an expert system that takes this world as its input.

What are the core elements of such a narrative mode of construal? In an evolutionary perspective, the operative question is that of function. According to MILLIKAN (1984), the biological function of a cognitive device is the work it is designed to accomplish by virtue of its past successes. What is the adaptive problem that the narrative mode of construal solved? In a broad sense, it is a given that the problem is one of extracting certain kinds of inferences from a series of events. Specifically, I suggest that the core challenge addressed by a narrative mode of construal is that of interpreting the past and the present in order to generate predictions about the likely future behavior of other agents. The problem of prediction is one of life and death: it is on this narrative level of accurately predicting and anticipating behavior that Little Red Riding Hood is outfoxed by the Wolf, with such lethal consequences. In a world of interacting agents, then, elementary forms of narrative serve the function of gen-

erating *predictions* about what types of *strategies* agents will pursue, based on inferences about their *goals*, by tracking the *obstacles* to achieving these goals and the *resources* available for overcoming them. These elements, I suggest, constitute the primitives, the basic categories, of an evolved narrative expert system. The narrative mode of construal takes its input from the moving $2^1/_2$D broadcast of consciousness, already parsed into objects and agents, and applies the filter of strategies and goals, resources and obstacles. This is not a sophisticated theory of mind module (BARON-COHEN 1995; LESLIE 1987); it involves only a basic construct of a goal, a rudimentary version of intentionality oriented towards the generation of immediate and short-term behavioral predictions.

As an example of how I envisage a narrative expert system operating to generate predictions, consider Don SYMONS' (1977) filmed analysis of rhesus macaque play-fighting. Symons begins by establishing the common narrative structure to such fights:

> Aggressive play may appear to be unordered or haphazard, but it is not. During playfights, each monkey attempts simultaneously to bite its partner and to avoid being bitten. A monkey achieves its goal to the extent that its partner fails. It is the players' working at cross-purposes to each other that makes playfighting so fast-paced, so complex, and so variable. (02:30 min)

By attributing goals to the agents – in this case a set of mutually exclusive goals, defining an agonistic relationship – Symons is able to make sense of a vast range of otherwise unpredictable behavior. What generates the complexity of the behavior, however, is not agonism itself but the size of each monkeys' repertoire of moves. The combinatorial space of all available moves is stupendously large, and the challenge of the playfight from the participants' perspective is to generate and carry out contextually appropriate sequences of moves that bring you closer to your goal. The critical skill here is the *length* and *complexity* of the behavioral sequences you can carry off – at the lower end, we call them tactics, and the higher end, strategies.

In Symons' blow-by-blow analysis of a playfight between A, a three-year-old male, and B, a two-year-old male, we see the younger monkey pursuing a rapid succession of tactical sequences with a time horizon of split seconds:

> B initiates the playfight by leaning forward and biting A on the chest. A's mouth is open, in preparation for biting. With his right foot, B pushes A's face away, and prevents A from biting.

The older monkey, in contrast, is working on developing strategies that involve complex, contingent sequences that have a time horizon of several seconds:

> A attempts simultaneously to roll B to his left and step to B's right, and thus attain a position behind B. Although B attempts to twist to face A, A uses his hands and left foot to roll B onto his side.

Strategic action is guided by the monkeys' *planning* – that is to say, their "attempt to achieve positions favorable for biting, and to avoid positions that render them susceptible to being bitten" (2:40 min). Achieving the behind position is a key strategic goal, as the behind money can bite at will and the monkey in front cannot bite at all. The successful completion of such a strategy in turn can be predicted based on the relative *resources* available to each agent, resources that in an agonistic relationship function reciprocally as the other's obstacles, the simplest measure of which is size:

> When the behind position is actually achieved in a playfight between two males, it is almost always the larger monkey who achieves it. (03:30 min)

There is no doubt that Symons possesses sophisticated expert systems enabling him to parse rhesus playfighting in a manner that provides him with a powerful generator of behavioral predictions. Yet, the reason his analysis needs to be so sophisticated is that the rhesus macaques themselves act strategically to reach goals. They generate complex behaviors that consist in moves drawn from a large repertoire, assembled into orderly and contextually contingent sequences designed to reach intermediate goals. The narrative mode of construal utilized to *predict* behavior, based on modeling the agents' goals and resources, has its counterpart in planning, a high-level narrative structuring of behavior. What we see in rhesus play is the development of proto-narratives in the form of multiple superficially varied concrete instantiations of strategies with an underlying structural design. Admittedly, from a human perspective this design remains simple: a sequence of moves aimed to achieve a behind position, with a time horizon of now more than a few seconds. Nevertheless, this analysis suggests that rhesus playfighting (in which aggression expresses itself in a deliberate sequence, but remains unconsummated) has the *structure* of fictive narratives.

In playfights, rhesus macaques occupy a first-person role in an exciting and aboriginal drama. By fighting with a larger and more experienced individual, younger monkeys are challenged to anticipate their opponent's moves. To master this task, they must construe these moves in narrative terms and grasp the underlying plot. In the safe environment of pretense, the players are given a low-cost opportunity to mine the possibility space of moves by understanding the narrative structures being developed by their opponents and by increasing the sophistication of their own.

The dual, complementary uses of the narrative mode of construal – to anticipate the behavior of others, and to achieve complex goals oneself by means of well-designed strategies – can thus be seen in a rudimentary form in non-human mam-

malian play. Key pieces of the human cognitive architecture appear to be in place: the monkeys are capable of creating demarcated pretend spaces where the skills required for high-stakes agonistic encounters can be practiced safely. From their behavior, we can infer that they parse the world into agents and objects, and there is little reason to deny them a sensory consciousness very similar to ours. In the present model *(Figure 1),* the production of conscious experience involves sophisticated, fast, mandatory, and informationally encapsulated processes located *below* the threshold of consciousness itself. Strategy development takes conscious experience as its input and parses it according to conceptual primitives that include a goal, obstacles to achieving this goal, and the development of strategic sequences of moves for marshaling available resources to maximize one's chances of overcoming these obstacles. We observe such strategy development in mammalian play, and this model formalizes our narrative intuitions that allow us to understand and anticipate their behavior. While the narrative mode of construal itself, according to this model, is one of a series of higher-level expert systems whose operations are *above* the

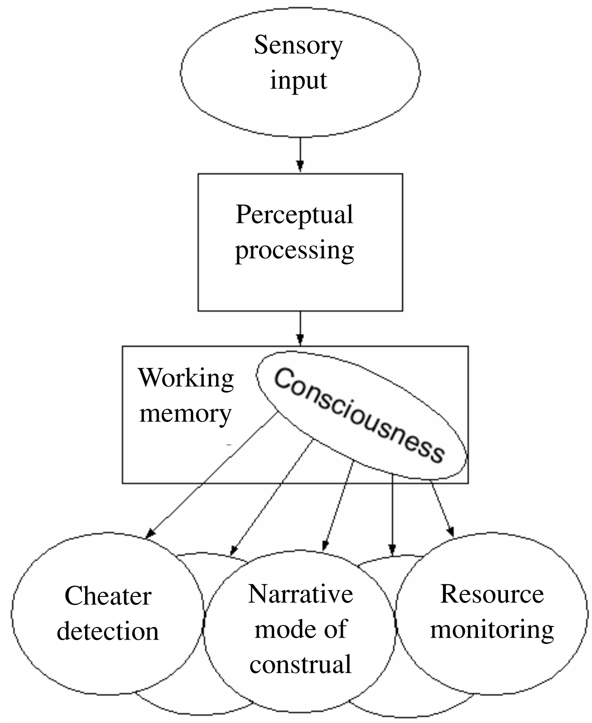

Figure 1. A Global Workspace model of sensory consciousness. Input from the senses is processed behind the scenes before it reaches the theatrical stage of working memory, on which is cast the spotlight of consciousness. A large audience of expert systems get their inputs from that portion of working memory that at any given time is lit up by consciousness

threshold of conscious awareness, its various products manifest themselves in consciousness as an anticipation of a move, as a feeling of an opportunity and an intention to act, as a mental image of a goal. As we share much of their evolutionary history, we can expect to share key features of rhesus minds.

However, impressive these achievements, major pieces of the human activity are missing from the drama of rhesus play. Although it is a play, it is not a performance. Although it is fictive, it does not involve the imaginative projection of oneself onto another agent. Although it has strategy development, the stories have time horizons of seconds rather than lifetimes. What is left for us to account for is how human narratives at once preserve and build on a set of preexisting adaptations, a sophisticated mammalian and primate cognitive architecture, and at the same time introduce dramatic innovations.

LITTLE RED RIDING HOOD REVISITED

In the story of Little Red Riding Hood, an aboriginal mammalian drama is given a hominid twist. The narrative retains the elementary structure of a chase: a vulnerable victim is spotted, pursued, and – depending on the version of the story you read – is either caught and eaten or gets away.

Yet, the story is not a chase. If a play chase is a simulation in action of a real chase, the story of a chase substitutes an imagined predator for the pretended predator. The mental image of the wolf is different for every child, yet every image emerges out of and makes explicit and comprehensible the concept of a wolf. Rather than relying on information from the senses, the maturing human child develops the ability to recall memories into consciousness, and to assemble these memories into episodes. By imaginatively constructing a pretend world from memory, cued by the words in the story, the child provides his or her higher-level inference systems with the input they need. The imagined world is a simulation of a sensed world, and conforms to the same format, the *lingua franca* of consciousness itself. This act of communication within the mind is required because our higher-level inference systems evolved to take processed perceptual information, presented in consciousness, as their input. The thrill of the chase is thus conveyed over to the physically passive act of listening to a story.

In the mind's eye, however, the listener must actively construct an inner geography in which the meaning and significance of the story can unfold. The story of Little Red Riding Hood becomes an event taking place in time and space: she walks from home through the forest towards the village where her grandmother lives, to bring her some food. She takes her time to enjoy the forest, "gathering nuts, running after butterflies, and making nosegays of such little flowers as she met with" (Lang

1889). The geographical level of the story must be generated before the more complex modeling can begin: it is required to feed higher-level inference engines the type of data they can handle.

The predation theme ubiquitous in mammalian play is put to novel and specifically hominid uses. The evocation of this ancient narrative serves first of all the dramatic purpose of activating the excitement, fear, and thrill of predation play, thus ensuring the child's rapt attention. This narrative usage is very different from the original biological function of predation play, which I argue elsewhere is that of providing an opportunity for practicing predator-evasion skills (STEEN and OWENS 2001). In the story, the predator – the wolf – has become a metaphor for a deceitful and ill-intentioned man. His is a blend of a wolf and a human being, selectively drawing features from each (cf. FAUCONNIER and TURNER 1998). While the predator features activate a primordial set of cognitive and physiological responses, the human features serve to explore aspects of the uniquely complex hominid possibility space. This possibility space is made possible by our expanded capacities for running recursive simulations for modeling other minds.

The wolf in the story, unlike real wolves, is a skillful mindreader, and he very subtly uses his skills to achieve his goals. When he first encounters Little Red Riding Hood in the forest, he does not attack her "because of some faggot-makers hard by in the forest" (LANG 1889). Now, why would the sound of woodmen discourage you? To make sense of the wolf's behavior, we must model the wolf's mind. As he encounters the girl, he quickly generates a conscious simulation of a possible future in which he at once attacks the girl. As he attacks, she screams in terror; he cannot stop her. Her screams are heard by the faggot-makers – the wolf now adopts the perspective of the out-of-sight woodmen, and infers that since he can hear them, they will hear the child's desperate cries. What will the woodmen do when they hear the cries? The wolf knows that in this particular species of primate, the default behavior of adults is to come to the aid of children threatened by predators, even when they are not the adults' own offspring; he knows that such a defense is going to be well coordinated, that these are strong and large males, and that they are armed with deadly axes. The wolf's higher-level inference systems compute quite accurately that this scenario is altogether unappealing; his emotions notify his whole body that such an attack is extremely risky and should be avoided; and he instantly abandons this initially promising course of action.

Because the wolf, being a human blend and having the mental capacities of a human being, has already covered the overhead costs of building a powerful simulation machine, the marginal costs of running a particular simulation is tiny. It is cheap for him to run through this complex, counterfactual scenario; in his mind, he can explore possibility spaces that in real life would have been fatally expensive. The simulation of the woodmen that come to the child's aid is an embedded tale within

the tale itself. Within the world of the tale, it never actually happens; in fact, the story is not even explicitly told. Yet for the child to understand the wolf's behavior, she must model the wolf's mind modeling the woodmen's mind, and reach the same conclusion.

The wolf runs through this possibility and rejects it fast, while they are still approaching each other. Before they begin to speak, he initiates a second counterfactual scenario. In this alternative simulation, he projects a future further ahead, a situation where he will be able to attack her and eat her under conditions of his own choosing, in a location where the woodmen will not hear her. This complex series of moves produces far more desirable emotions in the wolf, as success seems far more probable. But how can the wolf, who intends to eat the little girl, reliably obtain information from her about where she is headed? He quickly realizes that if he communicates his intentions to her, or even allow them to shine through, she will become suspicious and afraid, and withhold this information from him, thwarting his new scenario. The wolf must not only delay his gratification to be able to carry out his newly formed strategy; he must as best he can conceal his intentions from her, and make her think he has a different set of intentions than he actually does. How can he accomplish this? He must adopt the role of a trustworthy adult, someone the child can implicitly rely on. By manipulating her mind in this subtle manner, he increases his chances of successfully killing her eventually.

Innocent little Rotkäppchen is no match for this superpredator. His complex and rapid simulations are invisible and inaccessible to her; all she sees is the outward behavior of a friendly man. She takes this appearance at face value and blindly provides him with the information he requests, thereby not only endangering her own life, but placing her grandmother in imminent and mortal danger. Ignorant of her fatal mistake, she fritters away her time gathering beautiful flowers for a woman whom she has already condemned to an imminent death.

The listening child, however, is shielded from the price that Little Redcap has to pay. He is given the tools to understand the wolf's intentions, as the story provides him with information about the wolf's mind that is unavailable to her. A key and distinctive function of hominid stories is to reveal the hidden connections between thought and action, so that the child can improve his skills at mental modeling. Narrative, for this reason, presents its readers with transparent minds (COHN 1978), rendering the sequence of events hyperintelligible. The listening child is led to infer that communication is not always a good thing. He needs to understand in a visceral manner that the wonderful gift of communication is also a peril. If you freely provide information to those who want to hurt you, you help them to destroy you and those you love. Little Redcap should have communicated to the woodmen that she needed protection, or she should have refused to speak to the wolf. If she really had her wits about her, she could have outwitted him by pretending in turn to believe and

accept his pose of friendship and then provided him with information that was incorrect, saving herself and her grandmother and sending him off to some other location, perhaps a place where he would encounter a stronger adversary and be killed instead. These are alternative stories open to the listening child, once he has understood the challenge posed by the tale.

The story itself is an enactment of pretense. In chase play, the chaser pretends to be a monster by emitting cues such as stalking, grasping, and growling. In the final scene of the story of Little Red Riding Hood, the roles are reversed: it is the wolf, the predator, that pretends to be the safe and beloved grandmother. Yet in the telling of the story, the storyteller – who may herself be the listener's grandmother – must pretend to be and enact the wolf pretending to be her. The situation is of course entirely absurd: no child would ever make the gross categorical mistake of confusing a wolf for her grandmother. What makes this absurdity tolerable, indeed perfectly natural and thrilling, is that the wolf *is* in fact the grandmother, or some other safe and loved adult, pretending to be the wolf. As the storyteller simultaneously recounts and enacts the story, the wolf is conjured into the present by invoking his salient features as cues. His big eyes, his big ears, and – climactically – his big teeth bring him alive and allows the child the experience the safe and terrifying thrill of being eaten while your grandma hugs you.

In listening to another, we construct a partly conscious simulation out of the raw material of our personal memories. On the one hand, this construction provides our higher-level inference systems with the material they need to respond to the story in some way as if it actually happened. On the other hand, the conceptual grasp of the story that allows us to affirm a shared understanding is prior to and not dependent on the details of the simulation. Concepts have an interesting relation to consciousness: they must necessarily be instantiated in a particular form, drawing on personal memories, in order to be present in consciousness. Yet, this instantiation is not in itself the concept. The image you utilize to represent the concept in consciousness does not exhaust the concept, which can be instantiated in an infinite number of ways. Most interestingly, human beings have what appears to be a very robust if entirely implicit understanding of the distinction between a concept and its simulated instantiation in consciousness. The ability to distinguish between a concept and its particular instantiation would appear to be a requirement for symbolic communication beyond some elementary level of complexity, since the instantiation cannot be communicated. This line of reasoning produces the somewhat surprising conclusion that we cannot be conscious of a narrative as such, if what we mean by this term is the shared understanding a group of people have of a story. What we are conscious of is only the individual instantiation of a narrative, an instantiation that in itself is uncommunicable.

The proposal of this paper is that the reason conscious simulations play such a prominent role in our subjective experience of narrative, even though they play close to no role in our shared understanding, relate to the evolutionary origins of the narrative mode of construal as an expert system relying on consciousness for its inputs. Through millions of years of evolution, our ancestors' brains evolved a complex network of inference systems responding to highly processed information presented in sensory consciousness. By recalling memories into conscious awareness, mental simulations are able to tap directly into this machinery, activating the full range of cognitive responses as if (with appropriate caveats) the imagined event had been experienced and perceived in reality. From a pure information-processing perspective, this solution is hideously wasteful. It would be a much better engineering solution for the mind to operate directly on the conceptual structure of narrative and derive the appropriate inferences. The integrated architecture of the mind, in which our emotions, bodies, and thoughts are intimately tied to conscious sensations, appears to make this impossible. We might call this perspective "unintelligent design theory": natural selection, constrained by prior choices, may be driven towards very local optima.

This is no cause for grief. Nature abounds in poor design, and the everyday delights of narrative easily make up for the purely theoretical efficiency costs, however exorbitant they may be. In fact, the extravagantly idiosyncratic design of the human mind provides us with a significant competitive edge, since it renders our mode of thinking far less attractive as a paradigm for the artificial intellects of computers. We are so obsolete that we have become irreplaceable to each other.

REFERENCES

BAARS, B. J. (1988): *A Cognitive Theory of Consciousness.* New York: Cambridge University Press.

BAARS, B. J. (2002): The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Science,* 6, 1, 47–52.

BAARS, B. J. (2003): The global brainweb: An update on global workspace theory. Guest editorial. *Science and Consciousness Review,* October.

BADDELEY, A. D. (1993): Working memory and conscious awareness. In A. F. Collins et al. (eds): *Theories of Memory*. Hillsdale, NJ: Lawrence Erlbaum, 11–28.

BAILLARGEON, R. (1987): Young infants' reasoning about the physical and spatial characteristics of a hidden object. *Cognitive Development,* 3, 179–200.

BARON-COHEN, S. (1995): *Mindblindness: An Essay on Autism and Theory of Mind.* Cambridge, MA: MIT Press.

BATESON, G. (1980): *Mind and Nature*. London: Fontana.

BRUNER, J. (1987): *Actual Minds, Possible Worlds*. Cambridge, MA: Harvard University Press.

COHN, D. (1978): *Transparent Minds: Narrative Modes for Presenting Consciousness in fiction*. Princeton, NJ: Princeton University Press.

DENNETT, D. and KINSBOURNE, M. (1992): Time and the observer: The where and when of consciousness in the brain. *Behavioral and Brain Sciences,* 15, 183–247.

EDELMAN, G. M. (1992): Bright Air, Brilliant Fire: On the Matter of the Mind. New York, NY: Basic Books.

FAUCONNIER, G. and TURNER, M. (1998): Conceptual integration networks. *Cognitive Science,* 22, 2: 133–187.

FODOR, J. (1983): *The Modularity of Mind: An Essay On Faculty Psychology*. Cambridge, MA: MIT Press.

GOPNIK, A., MELTZOFF, A. N. and KUHL, P. K. (2000): *The Scientist in the Crib*. New York, NY: Perennial.

GRIMM, J. and W. (1812): *Kinder- und Hausmärchen*. v. 1. no. 26. Berlin. Trans. D. L. Ashliman. Available <http://www.pitt.edu/~dash/type0333.html>.

HERNADI, P. and STEEN, F. (1999): The tropical landscapes of Proverbia: A crossdisciplinary travelogue. *Style,* 33, 1, 1–20.

HIRSCHFELD, L. A. and GELMAN, S. A. (eds) (1994): *Mapping the Mind: Domain Specificity in Cognition and Culture*. New York, NY: Cambridge University Press.

LANG, A. (1889): *The Blue Fairy Book*. New York, NY: Longmans, Green and Co.

LESLIE, A. M. (1987): Pretense and representation: The origins of "Theory of Mind". *Psychological Review,* 94, 4, 412–26.

MARR, D. (1982): *Vision*. San Francisco, CA: W. W. Freeman.

MILLIKAN, R. G. (1984): *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, MA: MIT Press.

PERRAULT, C. (1697): *Histoires ou contes du temps passé, avec des moralités: Contes de ma mère l'Oye*. Paris: Barbin.

PROPP, V. L. (1928): *Morfologiia skazki*. [English tr. Morphology of the folktale.] *International Journal of American linguistics,* 24 (1958), no. 4, pt. 3.

SAUSSURE, F. de (1913): *Cours de Linguistique Generale*. Paris: Payot. English edition *Course in General Linguistics*. Translated by Wade Baskin. London: Peter Owen, 1974.

SPELKE, E. S. and HERMER, L. (1996): Early cognitive development: Objects and space. In R. Gelman, T. Kit-Fong et al. (eds): *Perceptual and Cognitive Development*. San Diego, CA: Academic Press, 71–114.

SPERBER, D. (1996): *Explaining Culture: A Naturalistic Approach*. Oxford: Blackwells.

STEEN, F. F. and OWENS, S. A. (2001): Evolution's pedagogy: An adaptationist model of pretense and entertainment. *Journal of Cognition and Culture,* 1, 4: 289–321.

SYMONS, D. (1977): *Rhesus Play*. Filmed and edited by John Melville Bishop. Written and directed by Donald Symons. Produced by Film Study Center, Harvard University.

TOOBY, J. and COSMIDES, L. (1992): The psychological foundations of culture. In J. H. Barkow, L. Cosmides and J. Tooby (eds): *The Adapted Mind,* 19–136.