

COMPLEXITY MANAGEMENT THEORY: MOTIVATION FOR IDEOLOGICAL RIGIDITY AND SOCIAL CONFLICT

Jordan B. Peterson and Joseph L. Flanders

Department of Psychology, University of Toronto

ABSTRACT

We are doomed to formulate conceptual structures that are much simpler than the complex phenomena they are attempting to account for. These simple conceptual structures shield us, pragmatically, from real-world complexity, but also fail, frequently, as some aspect of what we did not take into consideration makes itself manifest. The failure of our concepts dysregulates our emotions and generates anxiety, necessarily, as the unconstrained world is challenging and dangerous. Such dysregulation can turn us into rigid, totalitarian dogmatists, as we strive to maintain the structure of our no longer valid beliefs. Alternatively, we can face the underlying complexity of experience, voluntarily, gather new information, and recast and reconfigure the structures that underly our habitable worlds.

Key words: belief, religion, ideology, category, object, novelty, anxiety, frame problem, exploration, amygdala, prefrontal cortex

COMPLEXITY IS MORE UNAVOIDABLE THAN DEATH

There is a long line of speculation in the classic psychoanalytic and early cognitive literature regarding the essential role that belief systems play in emotional regulation (Becker, 1973; Freud, 1928/1991; Kelly, 1955). Freud posited that human beings constructed complex, irrational systems of fantasy, designed to rationalize and repress unacceptable and terrifying aspects of existence, such as death. For Freud, ideas of deity, immortality and even transcendent morality were childish conceits, products of a superstitious and unscientific past, designed to shield the individual from the truth: everyone is ultimately vulnerable. The well-defended individual does not face such truth, does not allow such facts into consciousness, because that realization would produce too much anxiety (Freud, 1928/1991). The neo-Freudian cultural anthropologist Ernest Becker (1973) extended Freud's beliefs about religion to ideology, integrating Freud's ideas, recast and arguably improved, with those of Otto Rank.

Becker believed that the emergence of self-consciousness rendered the individual's existential position in the world permanently intolerable. The individual aware of his or her mortal limitation has to hide from reality:

I believe that those who speculate that a full apprehension of man's condition would drive him insane are right, quite literally right.... Who wants to face up fully to the creatures that we are, clawing and gasping for breath in a

universe beyond our ken? ...Everything that man does in his symbolic world is an attempt to deny and overcome his grotesque fate. (p. 27).

Becker therefore presumed that human character was of necessity a “vital lie ... a necessary and basic dishonesty about oneself and one’s whole situation” (p. 55). Such self-deception is inevitable, because the world is a “hall of doom,” in Carlyle’s words (p. 55), a “nightmarish, demonic frenzy in which nature has unleashed billions of individual organismic appetites of all kinds – not to mention earthquakes, meteors and hurricanes, which seem to have their own hellish appetites” (pp. 53-54).

Becker realized that there was something pathological about such “necessary” and “inevitable” dishonesty; knew that the trivialization of reality came at the cost of dignity and self-respect. He believed that too much exposure to reality produced an intolerable chaos, that too little produced a narrow and unbearable restriction, and that the middle ground constituted a form of far-from-admirable but perhaps necessary “philistinism” (p. 81). Nonetheless, like Freud, Becker consistently reduced the highest human strivings to the need for yet another defence against the reality of finitude, and identified even the greatness of genius with the search for illusory immortality: “The genius repeats the narcissistic inflation of the child; he lives the fantasy of the control of life and death, of destiny, in the “body” of his work” (p. 109). The best that the creative mind can do, in consequence, is to “heroically ... create new illusions” (p. 188).

Becker was therefore finally sceptical of the benefits of psychotherapy, in general (“psychology as self-knowledge is self-deception, because it does not give what men want, which is immortality. Nothing could be plainer” (p. 271)), and, more broadly, of the value granted to insight itself:

Can any ideal of therapeutic revolution touch the vast masses of this globe, the modern mechanical men in Russia, the near-billion sheep-like followers in China, the brutalized and ignorant populations of almost every continent? ... Forget it. In this sense again it is Freud’s sombre pessimism ... that keeps him so contemporary. Men are doomed to live in an overwhelmingly tragic and demonic world. (p. 281)

This sophisticated neo-psychoanalyst saw the world, finally, as a place of existential catastrophe, from which human beings are protected by a shield of religious and ideological delusion – the delusion being first that life has some transcendent and ultimate value and second that human beings, qualitatively different from mere animals, somehow partake in that value. Similar notions have more recently been put forth by theoretical (Greenwald, 1980) and experimental social psychologists, most particularly in the form of terror management theory (McGregor et al., 1998; Pyszczynski et al., 1997).

Freud’s rationalist position, extended by Becker, must be given its due: religion and ideology can most definitely be used as a shield against weakness. Much of what passes for religious and ideological thought is mere defensive oversimplification and rationalization. However, it is still not clear that it is mortality itself that motivates necessary defensive repression, and poses the

central problem of life. The fundamental reality of human vulnerability and limitation appears, on further consideration, as something even broader and more fundamental than mere subjugation to death – which means, in a paradoxical sense, that the terror management theorists are too optimistic. Human beings are restricted spatially, for example, as much as temporally. We can only occupy one place at a time. Furthermore, the place that we occupy is shaped by the specific peculiarities, constraints and biases of our particular cultures. We are products of the social processes that shaped us, products of our time. Finally, we are characterized by profound intrinsic limitations on our perceptual and cognitive processing power. We can only make sense of a fraction of the information that constantly presents itself to us. The stability of the sense that we make is therefore fragile. Our models of experience are limited, incomplete, and chronically prone to failure. Our essential existential problem can thus be more accurately conceptualized as vulnerability to complexity, with subjugation to death appearing as a non-trivial but far from identical consequence of this more basic vulnerability.

George Kelly (1955) first hinted at the uncomfortable relationship we all hold with complexity at the beginning of the cognitive movement, insisting that human beings had an arbitrary, essential, unequivocal desire to be right – right once and for all, without question. Kelly regarded human beings as instinctive scientists, consistently and intuitively hypothesis-testing, but not as strict Popperians, consistently searching for falsifying information. The natural human scientist had instead a profound tendency to restrict or otherwise repress any “data” that threatened to invalidate his or her extant conceptual models. Why? Kelly states (1969, p. 283):

A major revision of one’s construct system can threaten with immediate change, or chaos, or anxiety. Thus it often seems better to extort confirmation of one’s opinion – and therefore of the system that produced them – rather than to risk the utter confusion of those moments of transition.

Kelly’s insightful theorizing highlights a fundamental behavioral tendency of human beings – a tendency that appears closely associated with animal territoriality (Peterson, 1999a, 1999b): we are perfectly willing to utilize aggression to defend those things we believe in or identify with. However, Kelly’s cognitively-oriented thought has been justly criticized for its lack of attention to motivation (Rychlak, 1982). He states, directly, that to be wrong is to encounter chaos, but he does not say why the encounter with chaos poses such a formidable existential problem. However, it is the formidably problematic “nature of chaos” that sits at the crux of the critical questions: Why are human beings motivated to protect their beliefs? Why are we motivated enough to use violence? And more – why will we use violence when its use is pragmatically counterproductive (Goldhagen, 1996)? Why will we use violence when its use clearly violates our own moral codes (Browning, 1993)?

We have enough information currently at hand to provide an answer to such questions – but not without reconsidering many of the presuppositions we currently accept as fact (and not without detailed reflection upon ontological,

religious and philosophical matters generally considered outside the purview of modern psychology). We need to determine exactly what it means to be exposed, without defence, to the underlying complexity of the world. We also need to reconceptualize “belief” – because it does not mean “description of the objective and independently extant world.” Finally, we need to invert our understanding of anxiety, and come to understand it as our default position in the world; come to understand it as something painstakingly brought under partial control, in consequence of effortful learning, and not something added through learning to a normative background of calm competence and security. We will start our journey towards such reconsideration in the realm of robotics, discussing the most intractable conundrum in artificial intelligence: the “frame problem.”

AI, the Frame Problem, and the Effortful Construction of Objects

Since the 1960's, AI researchers have consistently failed in their attempts to create a machine that could function – that is, model, evaluate, and act – in dynamic, real-world conditions. This failure is a consequence of the emergence of a “new, deep, epistemological problem” (Dennett, 1984, p. 129) – the unexpected difficulty of specifying what should be ignored, and what attended to, with regards to a particular action. This unexpected difficulty, the “frame problem,” was originally formulated, somewhat more narrowly, by McCarthy and Hayes (1969): how might logic-based event representations be specified, so that artificial information processors could make appropriate associations to and draw meaningful inferences from such events? A machine utilizing a properly framed representation could conceivably perform any action “A,” and infer all those and only those changes directly associated with “A” (Dietrich and Fields, 1996). A robot moving a toy block from one point to another in a given room, for example, might reasonably infer that its hand is occupied while the task is being undertaken. It should not spend too much time attending to the complex variations of illumination that characterize the toy block while it is moving, however. Neither should it check the walls of the room it inhabits, to see if they have changed color, or shape, or temperature (or density, or taste, or elasticity, or emotional state) during the procedure. These additional potential changes are not germane to the initial problem (and some of them are simply impossible). Unfortunately, determining what is and what is not germane turns out to be the “essence of intelligence” and the “hard part of the problems beings solved” as Brooks (1991b) points out. There are an infinite number of ways to perceive or construe a given situation, and an infinite number of potential consequences of a given action or event. The frame problem has therefore come increasingly to occupy center stage in discussions of intelligent machinery (Dennett, 1984).

Most simply, the frame problem is the problem of *relevance* (or, more accurately, *irrelevance*): there are a vast number of things that will not change in the course of any limited action, at least not in any important manner, but there are some that do. How can what does not change be determined, and ignored, while what does change, and what is relevant, be attended to and processed (Dietrich and Fields, 1996)? More complexly, however, the frame

problem is also the problem of the object: how can a given entity be segregated from the parts that compose it, the situation of which it is a part, and the other entities to which it is related? There is no simple answer to either of these problems – indeed, there appears to be no “answer” at all. Relevance cannot be determined, objectively. Furthermore, the defining parameters of a given entity are far from self-evident, far from simply given by the environment. The stimulus does not speak for itself. The staggering and as-of-yet unresolved difficulty of solving the frame problem taught AI researchers a very profound lesson: even apparently simple events are not bounded in any simple way (Dennett, 1984). Events are simple and distinct only insofar as their relevant features are framed, *a priori*, by the constraints of an operative context.

The frame problem is not restricted in its importance to explorers of silicon and electricity. On the contrary, AI’s encounter with the frame problem exposed the fundamental inadequacy of the naïve realism underpinning cognitive science and, by logical extension, psychology (Brooks, 1991a, 1991b; Dennett, 1984). It has become painfully clear that whatever constitutes the definitive boundary of the objects that we manipulate both concretely and abstractly is not something simply intrinsic to those objects. Every “object” can be classified, even perceived, in an infinite number of manners (Medin and Aguilar, 1999). Every object must therefore be regarded as something infinitely complex – objectively, intrinsically. We live in a virtual sea of complexity and uncertainty. We frame an object by cutting away vast swathes of information, indisputably associated with that object, but irrelevant to our current purposes, however those might be subjectively construed. Lacking subjectively determined constraint, therefore, even the most powerful information processors cannot compute the explosive complexity of the object – cannot even reliably define the edges of a regular solid, say, under changing conditions of lighting or position.

This means, to put it bluntly, that the concept of “object” itself is somewhat illusory. The world does not present itself neatly packaged into pre-existent categories, available for direct human perception and manipulation. Nothing at all can be understood in the absence of a structured and subjective frame of reference. How then do we frame our conceptions and our actions, to make even perception possible? This question can best be answered from a specifically developmental perspective, as our frames are acquired over time. In this manner, we can profitably consider individual socialization and learning, as well as the evolutionary processes that make such learning possible – and we can consider both as exemplars of variation and selection, unfolding over markedly different temporal spans.

The developmental psychologist Piaget regarded individual adaptation to the “environment” as a consequence of two processes, assimilation and accommodation – and it should be noted that the Piagetian environment was an emergent property of exploratory behavior, rather than an objective given (Evans, 1973, p. 20). Assimilation, for Piaget, meant incorporation of novel or anomalous information within the structures already underlying representation, habit and skill. Accommodation, by contrast, means reconstruction of representation, habit, and skill, in consequence of assimilation. In the early 1960’s, the pioneering Russian neuropsychologist E.N. Sokolov worked out

several fundamental propositions that may be regarded as a commentary on the Piagetian perspective. These propositions are cybernetic in their basic structure – predicated on the view that the organism is both fundamentally goal-directed, and responsive to environmental feedback indicating success or failure – as Sokolov was influenced directly by Norbert Wiener, the father of cybernetics.

Sokolov (1969) believed that the nervous system was a “mechanism” that changed its internal structure, while modeling the external world. The models so generated, isomorphic in structure with that external world – although somehow simpler: “apparently affecting only those relationships of interest to the organism in adapting to its surroundings” (p. 673) – could in principle be altered by the modeller, to enhance prediction of external events, and to enable active behavioral adaptation. Sokolov based his belief in such internal models on the existence of the “orienting response.” He noted that creatures exposed to novel or anomalous stimuli responded with eye movement, or alterations in galvanic skin response, or “depression of brain-wave rhythms” (p. 673), and believed that these alterations were not due so much to “incoming excitation” as to signals of discrepancy which develop “when afferent signals are compared with the trace formed in the nervous system by an earlier signal” (p. 673). Sokolov noted that these orienting responses disappeared after multiple instances of the phenomena that originally produced them. He assumed that the internal nervous system model updated itself to account for the anomaly, and that perceived discrepancy therefore vanished. Sokolov believed that such an update might occur in two manners: by improving the quality of extrapolation (from current models, one might presume) by securing additional information, or by changing the “principles by which such information is handled, so that the process of regulation will prove more effective” (p. 683). The parallelism with Piaget’s thought is clear.

It is difficult to determine how an organism might manage anything as complex as an orienting response – which, as Sokolov described, might be elicited by “the slightest possible change” (p. 673) in a given stimulus – without first constructing an elaborated and detailed model of the world. However, that process of modeling is far more difficult than might be reasonably first considered, given the apparently self-evident manner in which objects manifest themselves to us. The standard naïve realist view of the world, predicated on this self-evidence, is that “objective reality” is composed of independently existing entities, directly apprehended by our sensory systems. Out of these direct perceptions a model like that proposed by Sokolov is constructed. We attribute value to particular configurations of these perceptions, think and plan by internally manipulating our object perceptions, and then implement our plans in the real world – successfully, if our models are accurate; unsuccessfully, if they are not.

Something along these lines appears obviously true – even self-evident. However, many obvious truths have revealed themselves as profoundly inaccurate in the past – and such inaccuracy reigns in the present case. A given entity or object is actually very difficult to “perceive” – not so much because things in themselves lack structure, as classical nominalists might have it, but because that structure is so rich and variegated that it may be endlessly and

variously construed (Brooks, 1991a, 1991b; Hacking, 1999; Medin and Aguilar, 1999). Although there appear to be certain “basic level” categories that “leap out at us, and cry out to be named,” in the developmental psycholinguist Roger Brown’s terminology (1986) – so we naturally apprehend the table, instead of each of its four legs and its single flat surface – we do not know precisely how our perceptual and cognitive systems manage this apprehension. The fact of our physical embodiment and its evolutionarily-determined structure appears to play some critical role, defining for us a reality that best meets our needs, in a truly biological sense (Brooks, 1991a, 1991b; Brown, 1986; Lakoff, 1987). We apprehend the world from a perspective shaped above all by evolution, and we are not and perhaps cannot in principle be primary modellers of an objective world.

We do not know what an object is, “in and of itself,” because it may be so many things. The incredible complexity of the “environment” means that even a problem as simple as classifying a “modest-sized” set of entities can be solved in a “limitless number of ways” (Medin and Aguilar, 1999). This is at least partly because two things differ and are the same in as many ways as there are potential things to which they might be compared: books in a library, for example, might be categorized by the total number of “e’s” they contain, or by their age, or thickness, or by the number of atoms of selenium on the first page of their preface, or by how closely they approximate the weight of the average book. It might well be objected: such classificatory strategies are *ridiculous* – but the problem with that objection is that the lack of utility of a given strategy is a *judgement of value*, not a necessity drawn by logic or by any conceivable objective standard (as any “objective” judgement requires the *a priori* establishment of value-based criteria to judge by). And, if it is judgement of value that determines the validity of classification (or even of perception), then it could easily be that functional utility determines the nature of the object.

A chair is not a chair because it shares a set or even a subset of identifiable objective properties. A chair is a chair because a person can sit on it. This makes a chair a tool, – a tool useful for specifically human purposes – and not a thing. Human beings are not only tool-*using* animals. We are as well tool-*perceiving* animals. We see the pebbles that make up gravel because we can throw pebbles. We see the pen, and not the four or five parts that typically make up the pen, because we can write with the pen. In the absence of a specific goal, or at least the possibility of a specific goal (which means in the absence of arbitrary and value-predicated constraint) the universe does not reveal itself as structured, or it reveals itself as too complexly structured, which is very much the same thing. “Objects” are therefore not the simple constituent elements of the objective world, directly and simply perceived, but tools apprehended with difficulty, in the service of specific goals (and, furthermore, tools that may be perceived at very different levels of resolution).

You are sitting at a computer, for example, typing – enmeshed in a world delimited by current motivation, consisting for the present of the words on your computer screen, and the keyboard upon which you are typing. Suddenly, the computer crashes. The screen flashes, and goes blank. What is going on? – which means, what should you see? Negative emotion emerges. Your world of

apprehension expands. The computer itself, rather than the content of your thought, becomes an object of particular attention, as you check the hypothesis of basic mechanical malfunction. You turn the computer's power switch on and off, and check the wiring. Is there a loose connection? Has one of the mysterious components of the computer burned out? Can you smell overheated circuitry? No....

You flip the switch for the ceiling light, so you can better see what you are doing. But no light appears. Aha! You burned out a fuse. The "object" expands, once again – this time, to include the mechanics of the household wiring. You walk to the fuse box, and investigate it thoroughly. But nothing seems wrong. The mystery grows. You decide to hike down to the corner store, to pick up a pack of cigarettes, so you can smoke, while you consider the situation. You step out the door. Traffic is snarled up, all the way down the street. The traffic lights don't seem to be working. The neon sign on the corner store is not lit up. Aha! The power itself is out. And then you read next day that a storm on the sun, ninety three million miles away, produced an immense solar wind, intense enough to knock out the entire province-wide electrical grid.¹ But even that does not make you truly understand that you could not perceive your computer as a discriminable object in the absence of the stability of the sun. And if you don't believe that the computer's status as a definable object depends on its position in a complex, but invisible, system, ask yourself: why do you have to purchase a new one, every three years, even if the old (and formerly functional) one still "works"? Where exactly are the boundaries of a computer?

From an evolutionary perspective, we appear primarily concerned with action, and its consequences, rather than with representation (a perspective echoed, interestingly enough, by existential philosophy). The fundamental requirement to survive and replicate provides us, at some indeterminate remove, with an *a priori* set of values or goals, and with a ceaseless although bounded challenge: how can we achieve our intrinsically determined ends in an endlessly complex environment? From the time of Augustine (at least according to Wittgenstein, from whom the following ideas are derived) we have tacitly assumed, in accordance with the naïve realist stance alluded to previously, that words or categories were labels for things (Wittgenstein, 1968). Wittgenstein posited, by contrast, that a word was a tool; proposed that a word played a role in a "game"; observed that a word was more like a knight or a bishop in a chess match (Wittgenstein, 1968). "The meaning of a piece is its role in the game..." (pp. 150) – a game with both "rules" and "a point" (p. 150). This appears to us to be a position that is radically Darwinian, and therefore appropriate from a broadly scientific perspective: we label and communicate to foster the attainment of necessary ends, rather than for descriptive purposes, as such.

Heidegger (1927/1975) argues, with regards to this problem, that the phenomenological or experiential world is subjectively constituted, from the onset of perception – argues that the "object," which appears simply given, already contains the subject. We bring our psychological constraints and motivations to the world from the very beginning of the perceptual process.

¹ Something very like this happened in Quebec, Canada, in 1989 (see Kappenman et al., 1997).

Heidegger expresses this insight by describing the human being as “the Dasein” (the “being there”) – by describing the individual as an organism that is always and already “being-in-the-world.” The Dasein, primarily motivated and goal-directed, brings *a priori* constraint to the world. Our specific goals provide us with a certain disposition or orientation. When we are hungry, therefore, we do not “see” the infinite complexity of being. Instead, objects in the world manifest themselves to us in categories associated with our desire to attain food. Our goals therefore determine not only what we do with the world, but how the world appears to us, from the beginning. In consequence, the world is “always and already” meaningfully, subjectively structured, for the conscious subject.

What we perceive, “naturally” – that is, the objects of our conceptual universe, those things that cry out to be named – are not therefore so much self-evident and therefore easily nameable things, given to us by the nature of reality, as tools for the attainment of biologically-relevant goals, painstakingly extracted from an infinitely complex and dynamic background. This process of extraction is aided in the first place by perceptual systems whose operations have been shaped, phylogenetically, under evolutionary pressure (Gibson, 1977), so that certain phenomena of invariant importance across diverse environments “present themselves” to us in the course of minimal learning, and is aided in the second place by the ontogenetic processes of exploration, which allow us to construct up from these relative invariants those useful things we casually and erroneously regard as objects.

Barsalou (1983) has taken pains to describe how goals serve as the organizing principles for categories. He first discusses ad-hoc or one-off categories, such as “things to be taken from an apartment during a fire.” The disposition to preserve valued objects, in this particular case, constitutes the basic framework within which the world is simplified, prior to the execution of a given action sequence. This category scheme is clearly constructed on the fly, and is evidently not composed of “objects” that share much in the way of similarity of feature (one might take children and wedding pictures from a burning apartment, for example). Barsalou suggests further, however, that ad-hoc categories, common groupings or concepts (“household items”) and more firmly established perceptions are similar or even identical in structure, although they become represented at different “depths” of memory, or stages of automaticity – so that a functional category applied habitually becomes something more and more directly perceived, as an object, rather than something that has to be consciously grouped, as a concept. Something like this clearly happens when an individual learns to read, and first effortfully perceives features of letters, then individual letters; then effortfully perceives words, then perceives them automatically; then builds phrases, effortfully, and then even sees phrases, instantly, as “objects.”

From such a perspective, all categories emerge in relationship to goals, and begin in an ad-hoc manner. Some goals are transient. These produce transient categories, generated through conscious, effortful cognitive processing. Other goals emerge frequently, within individuals and across them (as a consequence of the operation of stable within- and between-person motivational and emotional states). The functional categories generated during pursuit of these

common goals become increasingly stable – increasingly automatized, within individuals, increasingly perceived instantly as objects – and also increasingly communicable, across individuals, depending on their degree of interpersonal commonality. With constant repetition, exploration, and interpersonal exchange of information, the process of extracting the useful and relevant features becomes automatic. At that point, the “object” is no longer a mass of complex and ultimately undefinable potentiality, but a particular identifiable object. So the developmental progression is complexity → goal → category → object. This process appears to be something akin to but more extensive than Piaget’s notion of the development of object permanence among children.

Categories functional enough to aid survival across many different contexts become increasingly easy to learn, even to perceive, as those fortunate enough to easily perceive them are more likely to survive, and reproduce. Categories that are sufficiently conventional, given the needs (or values) of particular organisms, come to be seen as stable, enduring aspects of the world – come to be seen as objects. This implies that categories organized around goals common across individuals become conventions, shared assumptions, and cultural presuppositions. Furthermore, this implies that cultural world-views (even ideologies) constitute accumulations of partial, inductively-derived solutions to the general problem of complexity, as much as or more than defences against existential anxiety – although they may also be that.

We appear to perceive things of maximum cross-situational utility, or maximal “relevance,” as we pursue our biologically-predicated goals. Formulation of the goal helps simplify the world massively, right from the onset: a given environment can in principle be parsed as a consequence of goal establishment into two very broad functional categories: those things relevant to goal attainment, and those things irrelevant. The latter category, which might be regarded most simply as ground, is by necessity the broader, as it contains the entire world, with the exception of the few phenomena apprehended as tools (or obstacles) specifically appropriate to the job at hand. Ground is what may be regarded as a constant, for the purposes of present operations. As long as something behaves predictably, it may be eliminated from attentive awareness. The former category – relevant things – must be carefully constructed, partially in tandem with goal-specification: it is unlikely that we can handle more than some arbitrary and small number of objects at any given moment. Miller (1956) estimated that number at seven, plus or minus two (see also Shiffrin and Nosofsky, 1994). Although this estimate is unlikely to be precisely accurate, it is clear that the number is small – perhaps even as small as four (Cowan, 2001) – and the arbitrary number seven will suffice for the purposes of the current argument.

So we appear necessarily determined at each moment to choose a goal that will allow the derivation of a conceptual “world” that consists of no more than seven objects (tools or obstacles). Otherwise, we posit a sub-goal, whose selection will allow for such derivation. The validity of these chunked object categories (Miller, 1956; Shiffrin and Nosofsky, 1994) – the validity of their inclusion and exclusion criteria – is subject to determination by assessment of their current functional utility (Simon, 1956). The world for a typist (assuming

computer) therefore is, for example, keyboard/keys/letters/monitor, and whatever specific verbal thoughts might constitute the subject matter for what is being typed. The great diversity that constitutes everything else is zeroed out or ignored (Lubow, 1989). Ignored, that is, unless trouble arises; ignored until the unexpected and undesired interruption of the ongoing sequence of goal-directed activity and the conceptual schema that is part and parcel of that sequence.

The Revenge of the Unjustly Ignored

How might such a goal-directed conceptual system work? Well, we know that the most fundamental aspects of motivation might be regarded as approach and avoidance – and that this is true far down the phylogenetic chain (Maier and Schneirla, 1935; Schneirla, 1959). This implies that behaving organisms, human beings included, are essentially linear creatures: we move forward and backwards, and our environment may be reasonably configured as a line. We strive to approach, and to consume: we move from comparatively undesirable point “a” to comparatively desirable point “b” (Adler, 1956; Peterson, 1999a). What we choose to value – that is, what we choose as the “content” or “locale” of point “b” – varies between individuals, within a broad but constrained domain (Rolls, 1999), as we must all eat and drink and breathe to live, as we must regulate our body temperatures, as we tend to value sexual behavior, social activity and dominance-hierarchy maneuvering or their abstracted or perhaps metaphorically or categorically identical equivalents. This means that we can understand each other without indefinite explanation, as we share a grammar of universal value (“I was angry with my brother” invokes “why were you angry?” in the course of conversation, not “what is anger?”), but that we may still differ very much, as there may be a practically infinite number of solutions to the ill-posed problem of attaining things of value.

To identify some end as valuable means essentially to grant it consummatory status, in the broad and narrow sense – broadly, as “end” implies consummation; narrowly, in that “consummatory reward” has attributes that are well understood and relevant to the current discussion (Rolls, 1999). Human capacity for abstraction means, however, that the merely hypothetical, arbitrary or symbolic may come to function as consummatory reward – to serve as goal; to indicate satiety, so that the acting organism can end its current sequence of behavioral operations; and to frame ongoing environmental events as “objects,” evaluated specifically as incentives, threats and punishments (Adler, 1956; Carver and Scheier, 1998; Gray, 1982; 1987; Gray and McNaughton, 1996; Oatley, 1992, 1999; Oatley and Johnson-Laird, 1987; Peterson, 1999a). These consequences of goal-setting appear universal, regardless of the “content” or specifics of the goal. This means, first, that the cortex may exercise modulatory control over the more evolutionary ancient motivational systems (Panksepp, 1999) by substituting abstractions, when possible, for more fundamental goals and, second, that goals might be considered *as a class*, rather than as specific exemplars. This latter point means that the diversity of potential goals that actually exists may be conveniently rendered irrelevant, and that the nature of “the goal” as such might serve as the abstracted object of discussion.

So we establish our point “b,” which is the ideal, motivated endpoint of our linear goal-directed activity, and we specify and evaluate both our point “a,” our starting point, and our means of movement, in reference to that currently operative ideal. And we strive to transform “a” into “b,” formulating and testing multiple possible partial solutions to the bounded frame problem presented by any motivational state, and we become anxiety-ridden or collapse in frustration or even die as a consequence of our failures, which are manifold and common. But now and then we embody a solution, as a consequence of favourable mutation, and reproduce, saving that mutation, or we stumble across a behavioral or conceptual solution, by overproducing potentially valuable ideas and percepts and then selecting those that are successful, in a quasi-Darwinian manner (Campbell, 1960), and then we communicate our successes. So our solutions to the frame problem are Darwinian, both phylogenetically (as we see things “properly” and allow our successful genes to accumulate over time) and ontogenetically (as we think many useless thoughts and try many useless approaches, but appreciate, conserve and communicate those that culminate in desired ends).

Point “b” is the desired future, the “goal;” point “a” is the unbearable present. The world is parsed up into seven plus or minus two functional categories, as the cognitive/perceptual part of the goal attainment process. Those categories serve as the objects of action (the tools and the end) for the plans designed to attain the goal. We apply procedures that we have painstakingly constructed, or that we have imitated or otherwise learned. And when we reach a goal, successfully, another emerges, from the menu comprised of our motivational or emotional states or their abstracted analogs (Peterson, 1999a). But this neat description assumes a perfect world, or perfect goal-oriented categorization of that world, which is essentially the same thing. Murphy’s law (a comical derivative of the second law of thermodynamics) unfortunately reigns supreme, however, in the real world: whatever can go wrong, will – and this is true even when we follow dutifully in the footsteps of tradition.

This all means (1) that we will make mistakes (because we do not understand everything, and because things we do “understand” change) and (2) that whenever we make a mistake, we encounter what we have not properly categorized and are presently ignoring (since, had we categorized it, and properly paid attention to it, we would not have made a mistake). And, because we ignore virtually everything when we are attempting to get something done, and because something important was ignored whenever we make a mistake, we must necessarily encounter the very large class of all currently ignored phenomena, whenever things go wrong. This “revenge of the unjustly ignored” immediately and thoroughly complicates our simple functional worlds.

The class of all currently ignored phenomena can be parsed, for the purposes of the current argument, into two major subclasses. The first subclass is everything ignored that resides “inside” our functional categories or presumptive objects. By definition, a category (and we would say, a presumptive object) contains things of a kind, considered in relationship to a goal. Things of a kind may be treated as if they were identical. The inner workings of an answering machine may all be treated as homogeneous and identical “parts,” for example –

as elemental atoms, metaphorically speaking – as long as the machine is performing as planned, expected or desired. This makes the answering machine something that may be treated as a unit, as a “single thing,” occupying limited cognitive, categorical, emotional and perceptual resources. This means that the object category “answering machine,” and object categories in general, might be regarded as functional, low-resolution images of the reality they are attempting to encapsulate. It is frequently the case, however, that one or more of our current categories or presumptive objects contains things that may not successfully be treated as a kind, for the purposes of our immediate goal-directed operations. This happens, for example, when a “thing” does not perform its implicit, desired, and predicted duty, because of its inherent and often invisible complexity. Such failure indicates the inadequacy of our current low-resolution take on the world.

The second subclass of currently ignored phenomena is the presumed-homogeneous set of “things” that reside “outside” the boundaries of our currently discriminated categories or presumptive objects, in the theoretically irrelevant “ground.” It may be that the category of “things that may be ignored” during current goal-oriented operations actually contains things that may not, in fact, be ignored – not if we wish to attain our ends. This is true virtually without fail, for example, when something very complex has been objectified, such as the value of a company’s stock, or the behavioral possibilities of a person. “Irrelevant background” is a particularly rich source of ignored information, containing as it does virtually everything that exists, outside the narrow domain of our current plan and the categories that we are utilizing to make that plan work.

So we simplify the world, to operate in it, by presuming functional homogeneity of relevant and irrelevant objects – but the presumption of homogeneity is subject in both cases to error. Error – that is, the failure of a goal-directed sequence of action and its accompanying schema to transform the world as desired – therefore either means (1) “the current functional categories utilized to simplify the world into multiple low-resolution objects and uniform but irrelevant ground are incorrect,” in which case they must be unpacked into their constituent elements at some presently unspecifiable level of resolution and reconstructed, or (2), analogously, “the motor procedures currently applied to transform the world are inappropriate,” in which case new procedures must be originated, constructed and put into place (see Carver and Scheier, 1998, for an elaborated and detailed description of the manner in which these processes might be related). Either way, a truly multi-stage process, fraught with potentially serious complication, must be initiated and undertaken. Things that were once regarded as understood may no longer be so regarded. Things that were ignored have now in some mysterious manner become relevant. Actions that were once habitual can no longer be unthinkingly applied. This is the terrible consequence of conceptual insufficiency, manifested as error (subjectively defined: I made a mistake), or of anomaly or complexity (objectively defined: the world appears as other than what I presumed).

Accept for a moment that state “b” may be regarded as the goal of action proceeding from current state “a.” A given behavior, presumed to alter the

“environment” in some desired manner, may therefore appropriately be judged with regards to its suitability according to its consequences. This is the intrinsic morality of goal-directed conceptualization. If the current behavior produces the results that are “expected” (“desired,” more accurately) then it is regarded as “correct,” that is, situationally appropriate (Simon, 1956). If something untoward occurs, however, as goal-directed behaviors manifest themselves, the brain circuitry underlying response to anomaly or uncertainty kicks into action (Davis and Whalen, 2001; Gray, 1982; Gray and McNaughton, 1996; Peterson, 1999a; Sokolov, 1969). “Untoward,” in this context, merely means “unexpected or anomalous.”

This means: I am performing an action, designed to obtain a specified end, in a specified place, at a specified time. The action does not produce the result intended. Instead, something else happens. The nature of that “something else” constitutes a mystery, and not a differentiated object or tool; constitutes something that is, in the most initial stages, uncategorized (except in the most general possible sense, as anomaly). At the very least, that something comprises a novel occurrence, in the context defined by my starting position, my goal and my behaviors. But a novel occurrence is in reality something exceedingly complex, as it is the world, which had been ignored while I was acting – excepting the seven plus or minus two tools I had carved out for my goal-directed purposes. The “re-emergence of the ignored world” is therefore an occurrence rife with potential meaning, ranging in its valence from extremely positive through irrelevant to terribly negative, ranging in its magnitude from trivial to catastrophic. This is the re-emergence of hidden complexity, the re-emergence of the frame problem – something that manifests itself first not as category and certainly not as object but as complexly dysregulated affect (with a decidedly negative, protective, defensive twist). The fact that things are not unfolding according to plan may mean virtually nothing – may signify only a trivial error in presupposition, easily repairable, and worthy of nothing but the investment of a few second’s corrective thought. Something entirely new and beneficial may even have manifested itself, consequent to my insufficient conceptualization. I may learn some new and strikingly useful categorization system, or some more productive and efficient habit, because of my mistake. Conversely, in the most unpleasant circumstance, I may discover some fatal error in my calculations, some utter failure – may come to realize that my goals are unattainable, my plans irreparably flawed, my self-conception totally inadequate, my fundamental preconceptions in serious and immediate need of reconstruction. Initially, however, none of these possibilities may be discriminated from one another. That initial state of indiscriminability or undifferentiation makes the unexpected or anomalous phenomenon a very complex and unsettling thing (see Neumann, 1955, for an elaborated and metaphorical discussion of this thing).

Adaptation as the Cautious Mining of Complexity

What is the most broadly functional response to a category that contains phenomena with such a broad range of potential import – with the full range of

potential import, in fact? Such a question seems impossible to answer, in principle, as the full range of potential import spans the spectrum of meaning, or implication for action. But it is a paradoxical fact that the very unpredictability of the anomalous occurrence may be regarded as a sort of constant – a constant predicated on the fact of its ineradicable and situation-independent motivational significance. The fact of this constancy means that appropriate conceptual schemes and habits specialized for dealing with unpredictability may be constructed and utilized, or even selected for, at least in principle, as a consequence of evolutionary pressure. The typical and functionally appropriate default response to unexpected plan-and-goal violation appears to be behavioral inhibition, and the accompanying emotion of anxiety. It appears generally best, for the purposes of continued short-term healthy survival, to cease carrying out a flawed sequence of activity, and to respond to new and unspecified situations with caution. Then it appears best to explore, cautiously, since long-term survival is dependent on adaptation to novel environmental contingencies, so that new threats are minimized, and new sources of information and resources properly exploited (Dollard and Miller, 1950; Gray, 1982; 1987; Gray and McNaughton, 1996; Peterson, 1999a).

Why behavioral inhibition, first? Why caution and anxiety? Well, if you don't hit the target, when you are aiming for it, it is not there – or, there is something wrong with your bow; or, more seriously, there is something wrong with you. In any case, you are in trouble, and you cannot tell initially how much trouble. So you stop what you are doing, become cautious. But what of the long term? How can you deal with the fact that an important or even vital goal-directed sequence of activity has not been successfully undertaken? If you are hungry, and frightened into immobility, you are still hungry: all the caution in the world will not feed you. So anxiety protects you, but it does not solve your problems. However, novelty is not only threatening. An undesired occurrence is proximally frightening, but distally rewarding (distal considered spatially, or temporally) (Dollard and Miller, 1950) – and that reward is, technically, incentive (Gray, 1982; 1987). Incentive reward – a cue that something good may be obtainable – motivates exploratory behavior. So that means that an undesired occurrence first motivates withdrawal, anxiety, caution, by default – and then exploration, all other things being equal, assuming that nothing additionally terrible or undesired occurs. And motivated exploration may extract from the previously unrevealed domain of anomaly or complexity precisely that useful and delimited information necessary to re-establish integrity of category and functionality of habit.

This does not mean that the psychological significance of exploration is even yet precisely clear. We presume, for the purposes of a given operation, that a given group of diverse elements may be treated as if it were one thing. “One thing” means “as if it will perform a single duty, under specified conditions, during a specified time frame, and in a particular locale.” When things do not go according to plan, the functional unity of a given category must immediately be called into question: a car that will not start is no longer a car (all delusional protestations to the contrary). It is instead “a group of problematically-yoked-together subelements, whose failure to function constitutes a mystery of

unspecified potential seriousness.” The emergence of the negatively-valenced unexpected, in a given situation, produces a state of behavioral inhibition, accompanied by anxiety. Mismatch between desire and revealed actuality means “stop doing what you are doing, because it is not producing the results intended.”

Mismatch means, by definition, that something is wrong – not so much that a model of the objective world has been falsified, but that a means is no longer useful, or that an end, whose attainability is a predicate of any means, is no longer attainable (Peterson, 1999a). The mere emergence of the anomalous error in behavior or presumption, however, does not provide information regarding the locale or nature of that error. Instead, error, novelty or anomaly generates anxiety, which can be regarded as a non-specific message of caution (caution: you’re not where you think you are – or, worse, you’re not who you think you are). This emergence of anxiety may or may not be followed by the desire to explore, which is more latent response to the second formal property of anomaly, or complexity: its incentive-reward status, as previously described. It is the process of incentive rewarding error-or-anomaly-motivated exploration that generates new and detailed information regarding the precise reason for the error or anomaly. This is all to say: functional information – which is the only kind that really counts – is not just there for the taking. It has to be extracted from the environment, as a consequence of careful, cautious, thoughtful, effortful, metabolically-demanding processing (Friberg, 1991; Ohman, 1979; 1987; Roland et al., 1987).

It is the unpacking and repacking of the implicit subelements of our categories that constitutes much of such difficult exploratory behavior (the intransigent car described earlier, for example, could easily now be something best considered in *ad-hoc* manner (Barsalou, 1983) as “a thing worthy of cursing and kicking,” “an unpredictable piece of junk,” “an uncontrollably expensive nightmare,” or, more specifically, as “something fit only to be towed to a junkyard” – and the initial affective response to its failure is response to all these indiscriminated possibilities). “Explore” therefore means “gather information, as a consequence of active interaction with the elements of the experiential world; unpack and re-structure categories, so that they are once again functional; and, finally, modify actions so that desire once again finds consummation.” We generally presume that we act as scientists, while we are exploring (following Kelly, 1955). We gather more information about the objective nature of things. We formulate new hypotheses, and test them. Our current model of anomaly-driven explanation appears to strongly support such a supposition. But there is a more accurate, pragmatic alternative: we are engineers, more than scientists. When we explore, we try to find out what operations work, more than what things are. In fact, we can not find out what things “are,” because they are too complex. We constantly strive, instead, to determine how the difficult and finally incomprehensible circumstances currently obtaining might be bent more effectively towards fulfillment of our biologically-grounded ends.

This means that we gather more information about the properties of things and situations through direct, hands-on manipulation of the world, as well as through active decomposition and reconstitution of the categories that make up

our objects of apprehension, and the habits that make up our potentials for action. We take things apart, in new ways, and put them together, in new ways, and therefore reveal properties of function that had been hitherto hidden from us, as a consequence of the simplified functional nature of our object representations. We do the same thing with our more abstract world-specifying concepts: the seven-plus-or-minus-two things that make up our fields of cognitive apprehension constitute categories with contents. These contents are the currently implicit constituent elements of the category (as a “car” has constituent elements: motor, transmission, body; as the motor, transmission and body are pistons and valves, gears and shafts, windows and doors) all packed up into a “unity” whose structure as a unity is violated whenever something that is not desired occurs.

Functional information is extracted, in the course of this careful, demanding processing, by directed attention to and exploration of the domain of potential or latent things, “inside” and “outside” of current categorical judgement and object apprehension. Directed attention and exploration therefore also necessarily means functional or explicit specification of the presuppositions guiding goal-directed behavioral maneuvering in the now error-ridden context, and their tentative, experimental restructuring. These presuppositions constitute implicitly or invisibly chunked object categories, or implicit-when-functioning-properly subroutines of goal-directed behaviors. Such invisibly chunked categories are, to say it again, groups of phenomena deemed equivalent because of their “similarity,” which must be for the sake of practicality and simplicity equivalent currently-goal-directed relevance or significance. Exploration thus means reconstruction of previous category or behavioral habit such that the probability of similar error in equivalent contexts is reduced or eliminated, at least in principle, in the future.

This exploration-guided category or habit construction or reconstruction is development of “personality,” so to speak, in the literal sense – that is, expansion or improvement of the current repertoire of functional categories and skills as a consequence of the voluntary incorporation of information left previously latent in the “world.” No such expansions or improvements just “happen” as a consequence of exposure to anomaly, except in the case of very simple or elementary errors (and even then the simplicity is only apparent: the “answer” is only “at hand” because of previous personal exploration, because the requisite knowledge was garnered and then socially transmitted by someone at some point in time for whom the problem was not simple, or because the problem has been solved for us by processes unfolding over evolutionary time scales) (Peterson, 1999a). “Habituation” is far from automatic.

The Neuropsychology of Uncertainty and Exploration

Anomaly – complexity – appears first in the form of anxiety, not in the form of object. The neuropsychological underpinnings of such appearance have become increasingly well understood, as a consequence of recent animal and brain-imaging work. It appears probable, for example, that it is the “limbic” substructure known as the amygdala that is primarily responsible for producing

the affective marker for the unknown (Damasio, 1994; Davis and Whalen, 2001; Ledoux, 1996). Gray, who originally and influentially posited that the septal-hippocampal system was responsible for anxiety (1982; 1987; Gray and McNaughton, 1996), appears somewhat in error regarding the precise neuroanatomical locale of the emotion-production mechanism – although it remains clear that the hippocampus is in fact involved in novelty detection and processing (Grunwald et al., 1998; Knight and Nakada, 1998; Strange et al., 1999). Gray’s broader theory regarding the generation of anomaly-anxiety, however, remains exceedingly informative (a theory can be incomplete or even wrong at one level of resolution, and right at another or many others).

Gray believes that the septal-hippocampal system is characterized by reaction to specified threats, as well as to the absence of expected rewards and to the presence of unexpected obstacles. The septal-hippocampal system is integrally involved (1) in analyzing spatial location and its abstracted equivalents – which means context – and (2) in the movement of events captured in short-term attention to long term memory (Eichenbaum, 1999; O’Keefe and Nadel, 1978). It is therefore in a prime position to identify what environmental events constitute deviations from desire, as what is expected or desired has to be (1) context-specific and (2) constructed as a potential object from memory. Gray presumes that the septal-hippocampal system tracks the relationship between expectancy or desire and the current status of the world (and this would be the world simplified by goal-positing) and then responds with behavioral inhibition and production of anxiety to mismatch (following Sokolov, 1968; Vinogradova, 1961).

Perhaps what the septal-hippocampal system does, instead, is specifically or peripherally disinhibit the function of the integrated amygdala/right-hemisphere systems responsible for anxiety (Davis and Whalen, 2001; LeDoux, 1996; Peterson, 1999a; Tucker and Frederick, 1989) when the current goal-directed “map of the environment” (O’Keefe and Nadel, 1978) fails – and if this neuropsychological localization/conceptual representation proves to be somewhat simplistic, the essential point still remains: anxiety may well be the *default* response to the unknown, inhibited by learning. This implies that it is security that is learned (Peterson, 1999a), and that such security may be “unlearned,” in a specific or more generalized manner, under the pressure caused by the emergence of complexity or anomaly. Freezing is a typical response, after all, to sudden placement in a novel environment (Gray, 1982; 1987). It is only after animals so placed have explored and “habituated” (a process that likely occurs only as a consequence of exploratory behavior and the information-gathering and model-updating that occurs in its wake) that they become “normally” calm. We confuse the post-exploration-adapted and therefore fearless animal with our theoretically stable, normal, emotionally-regulated selves, forgetting that our general complacency is a function of successful exploration conducted by ourselves, or by others, in the past.

Consider Hebb and Thompson’s words on the subject (1985, p. 766):

“One usually thinks of education, in the broad sense, as producing a resourceful, emotionally stable adult, without respect to the environment in

which these traits are to appear. To some extent this may be true. But education can be seen as being also the means of establishing a protective social environment in which emotional stability is possible”

Hebb and Thompson note, furthermore, that education changes the psychological structure of the individual, making him or her more “stable,” but also ensures that appearance and behavior in the social context will be more uniform. It is this inculcation of uniformity – socially-negotiated mutual agreement not to act or even think in a manner that would violate fundamental social-cognitive categories – that removes the impetus for dangerous, unpleasant and unpredictable emotional outbursts (p. 766):

“On this view, the susceptibility to emotional disturbance may not be decreased. It may in fact be increased. The protective cocoon of uniformity, in personal appearance, manners, and social activity generally, will make small deviations from custom appear increasingly strange and thus (if the general thesis is sound) increasingly intolerable. The inevitable small deviations from custom will bulk increasingly large, and the members of the society, finding themselves tolerating trivial deviations well, will continue to think of themselves as socially adaptable.”

In support of such notions, we know that decorticate animals, stripped of their capacity for inhibition, manifest highly emotional reactions to the slightest provocation (reviewed in LeDoux, 1996); know that rats exposed unexpectedly to a predator under naturalistic conditions cannot relax until they have re-explored the territory where the predator had appeared (Blanchard and Blanchard, 1989); know that the right hemisphere (particularly the right prefrontal cortex) appears integrally involved in the initial stages of novelty analysis, prior to partially linguistically mediated left-hemisphere routinization (Goldberg et al., 1994); know that individuals who have sustained right-hemisphere damage can no longer use even dramatically anomalous information to update their fundamental conceptual systems (Damasio, 1994; Ramachandran, 1996); and know that the lateral bed nucleus of the stria terminalis (an extension of the central nucleus of the amygdala) is associated not with phobia-like specific fear, but with uncertainty-anxiety (Davis and Whalen, 2001). It is Whalen’s work (Whalen, 1998; Davis and Whalen, 2001), in particular, that appears to lend most direct support to the notion that the amygdala responds to emergent complexity. The amygdala is most active during the early phases of conditioning or when stimulus contingencies change, but habituates rapidly with repeated presentations of the conditioned stimulus, even when such presentation continues to elicit fear behaviours (Buchel et al., 1998; LaBar et al., 1998). Consistent, predictable stimulus contingencies, by contrast, do not appear to maintain amygdala conditioned responses (Kapp et al., 1990).

Whalen (1998) therefore argues, additionally, that the amygdala is involved in modulating vigilance, in the presence of ambiguous stimuli. An ambiguous stimulus is one that has more than one possible meaning, more than one possible implication for action. To resolve this ambiguity, more information about the

nature of the stimulus must be gathered. The central nucleus of the amygdala appears to modulate precisely this kind of gathering. If the predictive value of a given stimulus is not understood, the amygdala potentiates vigilance in the perceptual systems. This may occur through the activation of cholinergic neurons in the basal forebrain that lower response thresholds of neurons in diverse sensory cortical areas as a consequence of acetylcholine release (Davis and Whalen, 2001).

It is of great interest to consider LeDoux's work on the amygdala more specifically, from a perspective modified by knowledge of Davis's work. LeDoux (1996) points out that the amygdala receives inputs from "a wide range of levels of cognitive processing" (1996, pp. 170). Inputs from the sensory areas of the thalamus can, for example, produce amygdalic response to "low level" stimulus features. These "low-level" features appear to potentially include those to which fear can be easily "conditioned": staring eyes, bared teeth; movements, shapes or other features characteristic of snakes, or eels, or spiders; blood, dismembered or immobile bodies; fire and, perhaps, dark or enclosed places (see Peterson 1999a for an extended discussion). Higher processing areas, by contrast, allow more complexly-constructed and difficult-to-recognize objects and events (LeDoux, 1996) to disinhibit anxiety. The sensory cortex may help with complex object recognition. Hippocampal inputs might allow both for the influence of contextual information (it is possible that contexts or situations, which "cannot be named," according to Wittgenstein, might be regarded as very transient objects, which can only be understood at very high levels of integrated processing) and for the interaction of memory and fear (in combination with the rhinal or transition cortex). The medial prefrontal cortex, higher yet up the processing hierarchy, has been implicated in "extinction" (LeDoux, 1996). At such higher and therefore more open and flexible levels (Panksepp, 1999), it makes increasing sense to consider such extinction and "habituation" as a consequence of active exploration, and the behavioral and conceptual generation and reorganization that emerges as a consequence (Peterson, 1999a), rather than as some simple automatic process of "failing to respond to."

The fact of this multiple-level input provides some anatomical foundation for our speculations regarding the nature of anomaly, or emergent complexity. Any given phenomenon is first encountered in a very primitive low-resolution manner, and reacted to as an exemplar of that primitive conceptual "category" (see also van der Kolk's and Fisler's (1995) discussion of highly emotion-laden and fragmentary memories in post-traumatic stress disorder). LeDoux (1996) uses the following illustrative story: a hiker is walking through the woods. He abruptly encounters a snake, coiled up behind a nearby log (p. 166):

"The visual stimulus is first processed in the brain by the thalamus. Part of the thalamus passes crude, almost archetypal, information directly to the amygdala. This quick and dirty transmission allows the brain to start to respond to the possible danger signified by a thin, curved object, which could be a snake, or could be a stick or some other benign object" (pp. 166).

The thalamus also passes visual information to the visual cortex, which

creates a more detailed representation of the stimulus. Why not use this more detailed information? Simply put: it takes longer to generate. Because snakes are fast, it is better to jump and be wrong (“oh, it’s only a stick!”) than to wait around a few hundred milliseconds and be dead. So it is clearly the case that one can know that something is up (“unexpected/ undesired (complex) thing” → “dangerous thing” → “dangerous animal” → “maybe snake”) before one knows what it is precisely that is up.

And it should be pointed out, as well: even the category “danger” or “potential snake” or whatever it is that the thalamus has conceptualized is something perhaps more well-developed, more specific, more processed and less primitive than an error message merely indicating the failure of a plan (something that appears processed, potentially, by the bed nucleus, referred to previously). But even that “more primitive” and unrevealed world of error is still something that must be responded to. It is of great interest to note, in this line, that recent research directly indicates that the amygdala can respond, via a subcortical midbrain-thalamus pathway, to visually presented but masked and literally “unseen” emotional stimuli (Morris et al. 1998; 1999); interesting as well that Bechara et al. (1999) have demonstrated separability of amygdala-generated emotion and ventromedial prefrontal action-oriented (decision-making) responses to that emotion.

I am moving from point “a” to point “b,” both specified by me, according to plan. But while I am acting, something I do not expect occurs; something that I have not encapsulated in my currently operative categorical system. I do not know what the undesired thing signifies, except for the inescapable but complex significance of the fact of its occurrence: *my operative plan is wrong*. Where my plan is wrong, I do not know; why it is wrong, I do not know; how it might be rectified, I do not know; and what may happen in consequence, I do not know. My mistake could be something of virtually any significance, however, as I have essentially excluded the world while immersed in my current goal-directed operation. It is certainly possible, therefore, that my mistake indicates the possibility that I am in great danger. Emotion emerges as a default response.

The desire/world mismatch, detected by the hippocampus, disinhibits the amygdala, activating circuitry in my right hemisphere (Peterson, 1999a; Tucker and Frederick, 1989), inhibiting positive-emotion and approach behavior governed by the left-hemisphere (Davidson, 1992). My current goal-directed actions cease (Gray, 1982), my autonomic nervous system is activated, my heart-rate rises (Fowles, 1980), cortisol floods my bloodstream (Gray, 1987). I feel anxious; I do not know who I am, where I am, or what is going on. *This is the signal of the emergence of complexity*. There is nothing within that complexity that must immediately reveal itself as object, concept or idea. There is only what was once but is no longer known (that is, my evidently-flawed previous goal-specific plan), what was unknown but has now been revealed (that is, whatever caused my error), “consciousness” of error (manifested in emotion), and a complex and information-laden territory, comprising the unknown occurrence, that might be explored and forced to reveal its secrets (its implications for the modification of action and representation).

This terrible, complex and information-laden “territory” is essentially

unbounded in its potential implications and, therefore, in its potential for generating negative affect. Information generated or released as a consequence of error-motivated exploration may cause cascades of concept failure, down the presupposition hierarchy (Peterson, 1999a). A plan rendered no-longer-operative may well comprise a key foundation block for many other equal, lesser or greater plans: as the proverb has it – for want of a nail the shoe was lost, for want of a shoe the horse was lost, for want of a horse the battle was lost, for want of the battle, the Kingdom was lost (see Carver and Scheier (1998) for a usefully extended hierarchical model of belief, in this vein). So the tendency to remain ideologically committed to a given position (associated with failure to explore and update in the face of anomaly) is also motivated by the desire to maintain the current superstructure of belief and tradition, in the face of evidence that a currently-unspeccifiably-large portion of it has been rendered dangerously and troublesomely invalid (Peterson, 1999a) (and, what is worse – dangerously and troublesomely invalid, by its own criteria).

Courageous, Creative Exploration as Process, not State, of Adaptation

Of course, every error does not produce infinite anxiety. The magnitude of the initial affective response is, all things considered, something proportional to the “size” of the goal-directed plan and conceptual system currently implemented. It is far more devastating to fail an important examination or to miss a long-sought-after promotion than to stumble into a chair that someone moved in your office. This is because larger-scale goal-directed schemas and their associated procedures or habits stabilize larger areas of “territory,” conceptualized both as space and time. Larger territories simply take more effort to explore and to map, so the consequences of large scale plan failure are more devastating. The problem is amplified when the larger territory so disrupted is shared, and therefore mapped, socially – when agreement with the principles of use and occupation have previously been carefully, dangerously and painstakingly negotiated with others, personally and historically. This point is particularly important for understanding ideological rigidity, and ideologically rigid response to alternative viewpoints. Two competing ideological systems may easily be predicated upon fundamentally opposed axioms – in which case to give credence to the opponent’s viewpoint is simultaneously to disrupt the basic substructure of current belief, and its attendant restriction of complexity.

It is one thing to make an error while acting out a given, bounded plan. It is another to encounter an alternate culture, or belief system, whose structure rests on axiomatic presuppositions at fundamental odds with yours. Mircea Eliade, the great religious historian (1965, 1985), has pointed out that explosive periods of creativity often occur in the aftermath of contact between previously isolated cultures. He is also careful to mention, however – like Tolstoy (1887/1983) – that the encounter between competing beliefs produces a psychological state akin to abstract death in the psyches of individuals unfortunate enough to constitute their spiritual battle-ground. This point has also been thoroughly developed by Carl Jung (1967, 1968) and Nietzsche (1968, pp. 301):

“In an age of disintegration that mixes races indiscriminately, human beings have in their bodies the heritage of multiple origins, that is, opposite, and often not merely opposite, drives and value standards that fight each other and rarely permit each other any rest. Such human beings of late cultures and refracted lights will on the average be weaker human beings: their most profound desire is that the war they *are* should come to an end.”

Belief does not shield death-anxiety, primarily, although it may serve that function. More properly, belief regulates and constrains complexity. This constraint is generally adaptive, pragmatically speaking, although it is sometimes merely defensive. Individuals are therefore motivated to maintain the structure of their belief systems, because those belief systems are painfully constructed abstracted patterns of action, designed to meet desired motivational ends, in a world complex and anxiety-provoking beyond understanding. The Freudian psychoanalysts presume that all cultural constructions are necessarily illusory, because the existential position of man is in the final analysis unbearable. Becker (1973) believes, for example, that identification with culture shields the individual from fear of death, defends him or her from overwhelming anxiety, but provides only a vague causal mechanism: the provision of a culturally-acceptable forum for symbolic immortality, either in the form of contribution to a transpersonal historical edifice, or in the form of life after death. The truth – typically – is far more complicated and interesting.

Cultural identity provides a mode of adaptation to the vicissitudes of life that is far from illusory. It does so by providing traditional categories of conceptualization and patterns of habit that directly serve their stated (and unstated) functional purposes (Peterson, 1999a). These purposes include (1) the provision of a stable and universally accepted mode of interpretation and habit, so that social interactions are rendered predictable and mutually beneficial; and, simultaneously, (2) the provision of diverse socially-acceptable means of personal attainment. In this dual manner, individual security may be obtained, and individual desire fulfilled, within a context that in the ideal remains both flexible enough to allow for update, and stable enough to allow for predictability. The fact that all culturally-determined categories and patterns could be other than they are, in some ways, and still function, does not demonstrate that they are illusory: it is possible to attain considerable real security and success as a physician or as a lawyer, for example, or as a Christian or a Jew, despite the differences in approach, value and belief that characterize these different modes of being. This flexibility of modes of adapted being is merely one more illustration that the world is sufficiently complex to bear being parsed up, functionally, in many different manners.

Furthermore, the “symbolic immortality” offered by such cultural-religious systems is far from merely defensive, and has not been properly understood, or even attended to, by academic psychologists or cognitive scientists. Becker (1973) attempted to provide “closure of psychoanalysis on religion” (pp. xiv). He essentially ignored Jung’s contribution to this topic, however, because the meaning of Jung’s work on alchemy (Jung, 1963, 1967, 1968), which occupied the latter half of Jung’s life, remained opaque to him: “I can’t see that all

(Jung's) tomes on alchemy add one bit to the weight of his psychoanalytic insight" (Becker, 1973, pp. xiv). There is no doubt that Jung's alchemical writings are difficult, but this is in part because they are revolutionary, at least from the perspective of modern psychology. Jung split with Freud on the topic of religion (see Ellenberger, 1970). Freud believed that religious thinking was defensive, in the same way that a neurosis was defensive – believed that religious thinking was self-deceptive, and necessarily and usefully supplanted by a sceptical rationality. Jung believed, by contrast, that religious thinking constituted mankind's essential but metaphorically-predicated adaptation to the totality of existential or phenomenological reality (although such thinking could be petrified, so to speak, into dogma and used in a purely defensive manner). His publication in 1911/1912 of the original German version of *Symbols of Transformation* (Jung, 1952), the first of his mature, alchemy-related works, was precisely the act that made his viewpoint qualitatively different from Freud's, and ensured his break in personal relations with Freud.

Jung's perspective on alchemy is extraordinarily difficult to summarize (see Peterson, 1999a, for a differentiated analysis), but its essential features can perhaps be laid out comprehensibly. He predicated his argument on the idea that cognitive categories necessarily transform over time, and demonstrated that the pre-empirical idea of "matter" therefore bore little resemblance to its modern counterpart. Matter for the pre-experimentalist was something more like chaos, psychologically speaking (something more like the unknown, or the undesired, or the emotion-inspiring, or, more particularly, something like complexity or anomaly): something more like what we mean when we say "it matters" or "that is a weighty matter" or "what does it matter?" or when we note that the "object" is precisely something that surprisingly "objects" to the realization of our desires. The anomalous matter of the object, from such a perspective, is *import*, before it is entirely manifested or, more fundamentally, *world*, before it is revealed. This is a conception with ancient roots. Reinhold Niebuhr (1964, pp. 6-7) describes Aristotelian concepts, for example:

"...since Parmenides Greek philosophy had assumed an identity between being and reason on the one hand and on the other presupposed that reason works upon some formless or unformed stuff which is never completely tractable. In the thought of Aristotle matter is 'a remnant, the non-existent in itself unknowable and alien to reason, that remains after the process of clarifying the thing into form and conception. This non-existent neither is nor is not; it is "not yet," that is to say it attains reality only insofar as it becomes the vehicle of some conceptual determination' (Jaeger, 1968, pp. 35)."

This perspective on matter is derived from a much more archaic and diversely-derived religious tradition (Eliade, 1978), predicated on the idea that the cosmos was derived from the interaction between the "Logos" or "Word" or seminal action of a creator-God, and the more basic, virtual, unformed "matter" of chaos. From the Jungian perspective (more accurately, from a traditional but mostly implicit religious perspective) the individual serves as the embodiment of that dynamic Word or seminal process, when he or she is fashioning the

structure of culture – when he or she is creating the comprehensible, secure and productive “world,” from the disorder or chaos of complexity or anomaly. This act of creation occurs when the latent “material” of nature is explored, and transformed into the functional categories and patterned behaviors that comprise familiar and secure territory – or, alternatively, when previously functional but now counterproductive concepts and actions are destroyed and recast (Peterson, 1999a). This makes the creative individual something akin to deity, in the sense implied in Genesis: man is made in the image of the process that extracts the world from its chaotic, undetermined, “material,” substrate. This idea echoes through the heroic/cosmogonic myths of the world, and is particularly evident in the creation stories of the ancient Middle East (Peterson, 1999a), which have played a determining role in shaping the structure and processes of modern consciousness and individuality. It takes no great leap of imagination, after all, to posit that the extant world described by such stories is the phenomenological world of experience, rather than the objective world of science, not least since conceptions such as the objective did not even formally exist when these stories and traditions were founded (Peterson, 1999a).

This means that our ancestors understood metaphorically at least five thousand years ago that the process of creative courageous encounter with the unknown comprised the central process underlying successful human adaptation, and that this process stood as the veritable precondition for the existence and maintenance of all good things. Such understanding, however, was implicit and low-resolution – at best, procedural, embodied, encoded in ritual and drama – and not something elaborated to the point we would consider explicit or semantic understanding today. We are constantly tempted to regard such understanding as superstitious, because of its continuing lack of explicitness, and to presume that our current modes of apprehension have rendered traditional beliefs superfluous. This attitude is predicated (1) on failure to recognize that empirical enquiry cannot provide a complete world description, because of the intractable problems of action, value and consciousness and (2) on an ignorance with regard to the content and meaning of pre-empirical or pre-experimental belief that is so complete, profound and unfathomable that its scope can barely be communicated.

The “kinship of the creative hero with deity” constitutes a phenomenon of tremendous import, as of yet radically uncomprehended: consciousness plays a world-constructing role, in a manner that is neither epiphenomenal nor trivial. It is for this fundamentally non-metaphysical reason that the individual cannot be sacrificed to the exigencies of social and political convenience, as those who live in western democracies have painfully come to realize: the “world-constructing capacity” of the individual must be respected and honored as something sovereign, lest the forces of chaos or complexity re-attain the upper hand, or the state rigidify and doom itself. The truly healthy individual comes to identify, over time, with the adaptive social structure generated by past heroes, by incorporating the hierarchical organization of that social structure into the self – but does not sacrifice his or her capacity for individual creativity, which is an “eternal and immortal” extra-social force, while so doing. This means not so much that the individual is protected against death-anxiety by the fact of culture

as that the individual is provided with a dual means of coping with vulnerable mortality in a meaningful and functional manner – first, as a consequence of his identity with social order and, second, as a consequence of his ability to voluntarily face chaos, complexity and anomaly, recast the protective strictures of tradition, and prevail.

This dual manner of coping is, to say it again, real, rather than illusory. The protection of culture is granted as a consequence of the provision of historically elaborated concepts and plans whose incarnation in behavior produces results that are necessary, intended and desired. This real protection is limited, however: the past is static by its very nature (is a “state”), and can therefore never provide complete information about the present or future. This means that the embodiment of past wisdom in present behavior will inevitably result in error, in anomaly, in “unrevealed world”, in chaos. In consequence, the healthy individual – however socially-adapted, must also play the hero, whose embodiment also provides real protection from the unknown, and who is therefore represented in traditional accounts as a divine psychological (“spiritual”) process (Peterson, 1999a). The individual must be willing to voluntarily face the consequences of the errors of the past, to mine the information embedded in the territory whose existence is revealed by those errors, and to reconstruct society and self, in consequence.

This all implies that those most likely to use identification with the current culture as a terror-management strategy (and to denigrate, punish or destroy those who threaten that protective culture) are precisely those who refuse to face the consequences of their own errors, and who directly and literally weaken the functional integrity of their personalities and the states they inhabit by doing so. The inevitable consequence of such weakening is increased existential anxiety, hopelessness, frustration, depression and anger, as poorly-laid plans produce results that are neither intended nor desired. Such weakening also engender ever-more intense desire to remain safely ensconced within the confines of the cultural world-model, and increases the probability that the capacity to deal with anomaly will become ever-more rejected and unlikely.

Ideological rigidity is therefore the tendency to avoid emotionally and cognitively-demanding exploration and information-gathering, subsequent to the receipt of an error message, in the interests of maintaining short-term emotional security. Events that indicate error in the pursuit of goals are negatively valenced, but informative. Ideologically rigid individuals sacrifice new and potentially useful information – and, therefore, personality and habitable world – to avoid short-term negative emotion. This makes totalitarianism of belief something that may be indulged in by default, so to speak – a sin of omission – and something that is potently reinforced, negatively, in the short term. This combination of ease and emotional relief might help explain the widespread prevalence of rigid, maladaptive belief. Dogmatic certainty is a condition that may be thoughtlessly and carelessly indulged in – a condition that lurks constantly as a temptation, as a second-rate alternative to the travail of authentic adaptation. This is the inauthenticity of the existential phenomenologists (Boss, 1963; Binswanger, 1968), the deadly spiral of the “adversarial personality” into chaos, and a process that inevitably breeds hatred for vulnerable existence

(Peterson, 1999a; 1999b). It is impossible to understand anything about the nature of the now-defunct Soviet Union, for example, without developing some appreciation for the integral causal interplay between individual capacity for rigidity and self-deception and genocidal totalitarian “illusion” (Solzhenitsyn, 1975).

Anomalies are unsettling because they represent everything that lies outside the domain of the understood world. Complexity lacks the simplifying and constraining boundaries defining the objects that characterize known territory. In consequence, we have profound, *a priori* motivation to avoid anomaly, to ignore complexity, and to maintain the structural integrity of our belief systems. Anything unexpected (new phenomena, new ideas, new people) re-introduces the overwhelming complexity that our beliefs simplify. This introduced complexity, in turn, threatens the stability and security that our beliefs tentatively confer on existence. Freud described religious beliefs as illusions, motivated by wish-fulfillment. Such beliefs can be more accurately understood as culturally-shared and accepted strategies for pragmatically managing complexity. It is necessary for us to generate simplified, functional models, in order to function in situations constantly beyond our understanding. However, this process of simplified functional modeling can be pathologized by individuals who are unwilling to allow any unconstrained complexity whatsoever to exist – pathologized, that is, by the existential cowards who make ideological purity the hallmark of existence. What Becker and the neo-Freudians describe as death terror can be more accurately conceptualized as *a priori* fear of unconstrained complexity. Mere belief cannot keep such complexity permanently at bay, because all models are simplified, functional representations of an exceedingly intricate underlying reality. This means that adaptation is by necessity a process (that is, the process of confronting complexity, voluntarily) as well as a state – despite the claims of committed ideologues to absolute comprehensiveness of current belief.

The adapted human being is protected by a functional shell of belief – but he or she is also able to repair that shell, or even to reconstruct it in its entirety, when it is challenged by a shift in circumstance or environment. Ideologically committed individuals, by contrast, make personal omniscience, in the guise of obeisance to a transpersonal belief system, an axiom of individual faith. In consequence, they undermine, then destroy, their personal capacity to explore and update, since everything that needs to be known has in their view been already discovered. Furthermore, they make their own suffering – a consequence of the insufficiency of their beliefs – heretical, and attempt to eradicate by whatever means possible anything that exposes their true pathology and weakness. Religious stories, occupying the necessarily metaphorical space at the base of our cultures of belief, provide the foundation for the dogmatic concepts and action patterns that structure our social interactions, and stabilize the territories that we all share. But this is not all protection against death anxiety. Much of it establishes the necessary “arbitrary” groundwork for shared social being, in the absence of real certainty, and instead of constant interpersonal conflict. More importantly, however, functional religious systems ensure that our shared beliefs are predicated on a concept of the individual that makes respect for the capacity of courageous, creative individual action in the face of

complexity the most fundamental and ineradicable of values (Peterson, 1999a). This predication makes faith in belief, as a state something necessarily subordinate to faith in courageous action as a process. This faith in process is a position that the totalitarian, who desires above all to be finally right, cannot tolerate, and attempts constantly to undermine and destroy. This faith in process is also a final answer, not to the problem of death, which cannot be solved, but to the more fundamental problem of complexity, and attendant vulnerability. Creative exploratory action in the face of anomaly and chaos generates, sustains and renews the world.

REFERENCES

- ANSBACHER HL and ANSBACHER RR. *The individual psychology of Alfred Adler: selections from his writings*. New York: Harper Torchbooks, 1956.
- ARNOLD MB. *Emotion and personality*. New York: Columbia University Press, 1960.
- BARSALOU LW. Ad hoc categories. *Memory and Cognition*, 11: 211-227, 1983.
- BECHARA A, DAMASIO H, DAMASIO AR and LEE GP. Different contributions of the human amygdala and ventromedial prefrontal cortex to decision making. *Journal of Neuroscience*, 19: 5473-5481, 1999.
- BECKER E. *The Denial of Death*. New York: Free Press, 1973.
- BINSWANGER L. *Being in the world*. New York: Basic Books, 1963.
- BLANCHARD DJ and BLANCHARD DC. Antipredator defensive behaviors in a visible burrow system. *Journal of Comparative Psychology*, 103: 70-82, 1989.
- BOSS M. *Psychoanalysis and daseinsanalysis*. New York: Basic Books, 1963.
- BROOKS A. Intelligence without reason. *MIT Artificial Intelligence Laboratory: Artificial Intelligence Memo 1293*, 1991a.
- BROOKS A. Intelligence without representation. *Artificial Intelligence*, 47: 139-159, 1991b
- BROWN R. *Social psychology* (2nd Ed.). New York: Free Press, 1986.
- BROWNING CR. *Ordinary men: Reserve police battalion 101 and the final solution in Poland*. New York: Harper Perennial, 1993.
- BUCHEL C, MORRIS J, DOLAN RJ and FRISTON KJ. Brain systems mediating aversive conditioning: An event-related fMRI study. *Neuron*, 20: 947-957, 1998.
- CAMPBELL DT. Blind variation and selective retentions in creative thought as in other knowledge processes. *Psychological Review*, 67: 380-400, 1960.
- CARVER CS and SCHEIER MF. *On the Self-Regulation of Behavior*. New York, NY: Cambridge University Press, 1998.
- COWAN N. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Brain and Behavioral Sciences*, 24: 87-114, 2001.
- DAMASIO AR. *Descartes' error: Emotion, reason and the human brain*. New York: Avon Books, 1994.
- DAVIDSON RJ. Anterior cerebral asymmetry and the nature of emotion. *Brain and Cognition*, 20: 125-151, 1992.
- DAVIS M and WHALEN PJ. The amygdala: vigilance and emotion. *Molecular Psychiatry*, 6: 13-34, 2001.
- DIETRICH E and FIELDS C. The role of the frame problem in Fodor's modularity thesis: a case study of rationalist cognitive science. In Ford KM and Pylyshyn ZW (Eds) *The Robot's Dilemma Revisited: The frame problem in Cognitive science*. New Jersey: Ablex Publishing Corporation, 1996.
- DOLLARD J and MILLER N. *Personality and psychotherapy: An analysis in terms of learning, thinking, and culture*. New York: McGraw-Hill, 1950.
- EICHENBAUM H. The hippocampus and mechanisms of declarative memory. *Behavioral and Brain Research*, 103: 123-133, 1999.
- ELIADE M. *Rites and symbols of initiation: The mysteries of birth and rebirth* (Trask WR, Trans.). New York: Harper and Row, 1965.
- ELIADE M. *The forge and the crucible* (Corrin S, Trans.) (2nd ed.). Chicago: University of Chicago Press, 1978.
- ELIADE M. *A history of religious ideas. Vol. 3. From Muhammad to the age of reforms*. Chicago: Chicago University Press, 1985.
- ELLENBERGER H. *The discovery of the unconscious*. New York: Basic Books, 1970.
- EVANS PI. *Jean Piaget: The man and his ideas*. New York: E.P. Dutton and Company, 1973.
- FOWLES DC. The three arousal model: Implications of Gray's two factor learning theory for heart-rate, electrodermal activity, and psychopathy. *Psychophysiology*, 17: 87-104, 1980.
- FREUD S. The future of an illusion. In *Civilization, society and religion*. New York: Penguin, 1928/1991, pp. 181-241.

- FRIBERG L. Auditory and language processing. *Alfred Benzon Symposium*, 31: 44, 1991.
- GAZZANIGA MS and LEDOUX JE. *The integrated mind*. New York: Plenum Press, 1978.
- GIBSON JJ. The theory of affordances. In Shaw R and Bransford J (Eds) *Perceiving, Acting and Knowing*. New York: Wiley, 1977, pp. 67-82.
- GOLDBERG E, PODELL K and LOVELL M. Lateralization of frontal lobe functions and cognitive novelty. *Journal of Neuropsychiatry and Clinical Neuroscience*, 6: 371-378, 1994.
- GOLDHAGEN DJ. *Hitler's willing executioners: ordinary Germans and the Holocaust*. New York: Alfred Knopf, 1996.
- GOLDMAN-RAKIC PS. Architecture of the prefrontal cortex and the central executive. *Annals of the New York Academy of Sciences*, 769: 71-83, 1995.
- GRAY JA. *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. Oxford: Oxford University Press, 1982.
- GRAY JA. *The psychology of fear and stress* (2nd ed.). Cambridge: Cambridge University Press, 1987.
- GRAY JA and MCNAUGHTON N. The neuropsychology of anxiety: Reprise. *Nebraska Symposium on Motivation*, 43: 61-134, 1996.
- GREENWALD AG. The totalitarian ego: Fabrication and revision of personal history. *American Psychologist*, 7: 603-618, 1980.
- GRUNWALD T, LEHNERTZ K, HEINZE HJ, HELMSTAEDTER C and ELGIN CE. Verbal novelty detection within the human hippocampus proper. *Proceedings of the National Academy of Science USA*, 95: 3193-3197, 1998.
- HACKING I. *The social construction of what?* Cambridge, MA: Harvard University Press, 1999.
- HEBB DO and THOMPSON WR. The social significance of animal studies. In Lindzey G and Aronson E (Eds) *The handbook of social psychology*. New York: Random House, 1985, pp. 729-774.
- HEIDEGGER M. *The Basic Problems of Phenomenology* (Hofstadter A, Tans). Indianapolis: Indiana University Press, 1927/1975.
- JAEGER W. *The theology of the early Greek philosophers: The Gifford lectures 1936*. London: Oxford University Press, 1968.
- JUNG CG. Symbols of transformation: an analysis of the prelude to a case of schizophrenia. In Hull RFC (Trans) *The collected works of C.G. Jung*, Bollingen Series XX. Princeton: Princeton University Press, Vol 5, 1952.
- JUNG CG. Mysterium Coniunctionis. In Hull RFC (Trans) *The collected works of C.G. Jung*, Bollingen Series XX. Princeton: Princeton University Press, Vol 14, 1963.
- JUNG CG. Alchemical Studies. In Hull RFC (Trans) *The collected works of C.G. Jung*, Bollingen Series XX. Princeton: Princeton University Press, Vol 13, 1967.
- JUNG CG. Psychology and alchemy. In Hull RFC (Trans) *The collected works of C.G. Jung*, Bollingen Series XX. Princeton: Princeton University Press, Vol 12, 1968.
- KAPP BS, WILSON A, PASCOE JP, SUPPLE W and WHALEN PJ. A neuroanatomical systems analysis of conditioned bradycardia in the rabbit. In Gabriel M and Moore J (Eds) *Learning and computational neuroscience: Foundations of adaptive networks*. Cambridge, MA: MIT Press, 1999, pp. 53-90.
- KAPPENMAN JB, ZANETTI LJ and RADASKY WA. Geomagnetic storms can threaten electric power grid. *Earth in Space*, 9: 9-11, 1997.
- KELLY G. *The Psychology of Personal Constructs*. New York: Norton, 1955.
- KELLY G. The threat of aggression. In Maher B (Ed) *Clinical Psychology and Personality: The Selected Papers of George Kelly*. New York: Wiley, 1969, pp. 281-288.
- KNIGHT RT and NAKADA T. Cortico-limbic circuits and novelty: a review of EEG and blood flow data. *Review of Neuroscience*, 9: 57-70, 1998.
- LABAR KS, GATENBY JC, LEDOUX JE and PHELPS EA. Human amygdala activation during conditioned fear acquisition and extinction: A mixed-trial fMRI study. *Neuron*, 20: 937-945, 1998.
- LAKOFF G. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press, 1987.
- LEDOUX JE. *The Emotional Brain*. New York: Simon and Schuster, 1996, pp. 42-72, 138-178.
- LUBOW RE. *Latent inhibition and conditioned attention theory*. Cambridge: Cambridge University Press, 1989.
- MAIER NRF and SCHNEIRLA TC. *Principles of animal psychology*. New York: McGraw-Hill, 1935.
- MCCARTHY J and HAYES PJ. Some philosophical problems considered from the standpoint of artificial intelligence. In Meltzer B and Michie D (Eds) *Machine intelligence*. New York: Elsevier, Vol 4, 1969, pp. 463-502.
- MCGREGOR HA, LIEBERMAN JD, GREENBERG J, SOLOMON S, ARNDT J, SIMON L and PYSZCZYNSKI T. Terror management and aggression: Evidence that mortality salience motivates aggression against worldview-threatening others. *Journal of Personality and Social Psychology*, 74: 590-605, 1998.
- MEDIN DL and AGUILAR CM. Categorization. In Wilson RA and Keil F (Eds) *MIT Encyclopedia of Cognitive sciences*. Cambridge, MA: MIT Press, 1999, pp. 104-105.
- MILLER GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63: 81-97, 1956.

- MORRIS JS, OHMAN A and DOLAN RJ. Conscious and unconscious emotional learning in the human amygdala. *Nature*, 393: 467-470, 1998.
- MORRIS JS, OHMAN A and DOLAN RJ. A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Science USA*, 96: 1680-1685, 1999.
- NEUMANN E. *The great mother: An analysis of the archetype* (Manheim R, Trans). New York: Pantheon Books, 1955.
- NIEBUHR R. *The nature and destiny of man: A Christian interpretation*. New York: Charles Scribner's Sons, Vol 1, 1964.
- NIETZSCHE F. Beyond good and evil. In Kaufmann W (Ed and Trans) *The basic writings of Nietzsche*. New York: Random House, 1968, pp. 181-438.
- O'KEEFE J and NADEL L. *The hippocampus as a cognitive map*. Oxford: Clarendon Press, 1978.
- OATLEY K. *Best laid schemes: The psychology of emotions*. New York: Cambridge University Press, 1992.
- OATLEY K. Why fiction may be twice as true as fact: fiction as cognitive and emotional simulation. *Review of General Psychology*, 3: 101-117, 1999.
- OATLEY K and JOHNSON-LAIRD PN. Towards a cognitive theory of emotions. *Cognition and Emotion*, 1: 20-50, 1987.
- OHMAN A. The orienting response, attention and learning: An information-processing perspective. In Kimmel HD, Van Olst EH and Orlebeke JF (Eds) *The orienting reflex in humans*. Hillsdale, NJ: Erlbaum, 1979, pp. 443-467.
- OHMAN A. The psychophysiology of emotion: An evolutionary-cognitive perspective. In Ackles PK, Jennings JR and Coles MGH (Eds) *Advances in psychophysiology: A research annual*. Vol 2 Greenwich, CT: JAI Press, 1987, pp. 79-127.
- PANKSEPP J. *Affective neuroscience*. Oxford University Press, 1999.
- PETERSON JB. *Maps of meaning: The architecture of belief*. New York: Routledge, 1999a.
- PETERSON JB. Individual motivation for group aggression: Psychological, mythological and neuropsychological perspectives. In Kurtz L (Ed) *Encyclopedia of Violence, Peace and Conflict*. San Diego: Academic Press, 1999b, pp. 529-545.
- PYSZCZYNSKI T, GREENBERG J and SOLOMON S. Why do we need what we need? A terror management perspective on the roots of human social motivation. *Psychological Inquiry*, 8: 1-20, 1997.
- PYSZCZYNSKI T, GREENBERG J and SOLOMON S. A dual process model of defence against conscious and unconscious death-related thoughts. *Psychological Review*, 106: 835-845, 1999.
- RAMACHANDRAN VS. *The evolutionary biology of self-deception, laughter, dreaming and depression*. *Medical Hypotheses*, 47: 347-362, 1996.
- ROLAND PE, ERIKSON L, STONE-ELANDER S and WIDEN L. Does mental activity change the oxidative metabolism of the brain? *Journal of Neuroscience*, 7: 2372-2389, 1987.
- ROLLS E. *The brain and emotion*. New York: Oxford University Press, 1999.
- SCHNEIRLA TC. An evolutionary and developmental theory of biphasic processes underlying approach and withdrawal. *Nebraska Symposium on Motivation*, 5: 1-42, 1959.
- SHIFFRIN RM and NOSOFKY RM. Seven plus or minus two: A commentary on capacity limitations. *Psychological Review*, 101: 357-361, 1994.
- SIMON HA. Rational choice and the structure of the environment. *Psychological Review*, 63: 129-138, 1956.
- SOKOLOV EN. The modeling properties of the nervous system. In Maltzman I and Coles K (Eds) *Handbook of contemporary Soviet psychology*. New York: Basic Books, 1969, pp. 670-704.
- SOLZHENITSYN AI. *The gulag archipelago, 1918-1956: An experiment in literary investigation* (Whitney TP, Trans), Vol 2. New York: Harper and Row, 1975.
- STRANGE BA, FLETCHER PC, HENSON RN, FRISTON KJ and DOLAN RJ. Segregating the functions of human hippocampus. *Proceedings of the National Academy of Science USA*, 96: 4034-4039, 1999.
- TOLSTOY L. *Confessions* (Patterson D, Trans). New York: W.W. Norton and Company, 1887/1983.
- TUCKER DM and FREDERICK SL. Emotion and brain lateralization. In Wagner H et al. (Eds) *Handbook of social psychophysiology*. Chichester, UK: John Wiley and Sons, 1989, pp. 27-70.
- VAN DER KOLK BA and FISLER R. Dissociation and the fragmentary nature of traumatic memories: Overview and exploratory study. *Journal of Traumatic Stress*, 8: 505-525, 1995.
- VINOGRADOVA O. *The orientation reaction and its neuropsychological mechanisms*. Moscow: Academic Pedagogical Sciences, 1961.
- VITZ PC. The use of stories in moral development: New psychological reasons for an old education method. *American Psychologist*, 45: 709-720, 1990.
- WHALEN PJ. Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. *Current Directions in Psychological Science*, 7: 177-188, 1998.
- WITTGENSTEIN L. *Philosophical investigations* (3rd ed.) (Anscombe GEM, Trans). New York: Macmillan, 1968.