

**Integrative Thinking:
The View from Cognitive and Social Psychology,
Circa 2005 AD.**

Mihnea Moldoveanu
Director, Marcel Desautels Centre for Integrative Thinking
Joseph L. Rotman School of Management
University of Toronto

Synopsis

We establish and formulate some disciplinary foundations for the study of integrative thinking, drawn from studies of individual perception, memory, cognition, intuition, meta-cognition and situational reasoning. We summarize a large-scale research program in cognitive and social psychology that has firmly established the occurrence, prevalence and resilience of 'counter-integrative' tendencies in thinking – such as sub-optimal convergence to early cognitive commitments to ideas, perceptions and models and tendencies to preserve such sub-optimal commitments in the face of strong refuting signals. We unveil the logical structure of this research program and highlight the often hidden assumptions on which its largely pessimistic conclusions are predicated in 'counterpoint' sections throughout the document. We articulate opportunities for research ('Research Questions: RQ') that should lead to the articulation of both selection criteria for individuals more likely than others to overcome the counter-integrative proclivities of their own minds through training and the design of interventions meant to produce integrative patterns of thinking and ways of being.

1. Sketch of Integrative Thinking.

Integrative thinking can be conceptualized as a faculty of 'mind' to interact with self, world and others successfully in predicaments that are unstructured, ill-structured, ambiguous or calling for the integration of disparate knowledge structures and behavioral patterns into a single cognitive schema and behavioral action plan. The integrative way of being is characterized by successful resolutions of the tensions between the need to learn and adapt and the need to act decisively and purposefully. Cognitive narrowings and bracketings of managerial predicaments along 'disciplinary' lines of thought – via the use of the models, myths and metaphors of 'normal science' – epitomize patently 'non-integrative' ways of being. A working model for an integrative way of being [Martin, 2003] is one predicated on the ability to consider 'more' salient variables in a decision process, the ability to internalize and conceptualize nonlinearly-structured temporal and structural relationships between the salient variables and the ability to harmonize and synthesize alternative standpoint rather than choose between them, all while retaining the ability to act decisively and punctually. Another possible working model is one that stresses meta-cognitive abilities such as the ability to entertain alternative beliefs while suspending the need to believe or disbelieve them, the ability to change fundamental epistemological, logical and ontological commitments quickly, playfully and non-committally, the ability to hold several mutually contradictory theories or models in mind at the same time and the ability to create, appropriate and

manipulate large-scale, complex knowledge structures and behavioral patterns, all while retaining the ability to function successfully in the pursuit of personal ends.

These working models bear some resemblance to models of 'open-mindedness' [Rokeach, 1968] articulated by psychologists as a scientific response to the doctrinaire ethos of the decades in the middle of the last century and to models of intellectual honesty [Lakatos, 1970] and epistemic rationality [Kuhn, 1983] in the philosophy of science, but differs from these models insofar as it includes an actionability, intuitiveness and ways of knowing that lie beyond the realm of language and representational cognition (which are the domain of cognitive psychology and the philosophy of (largely natural) science). It is in need of direct operationalization, an end that would be greatly helped by an overarching theory of action (including cognitive action) in the face of complexity, ambiguity and representational incommensurability that would serve to turn integrative thinking into an 'established' concept and a conceptual hub for real world interventions (including but not limited to pedagogical tools).

2. 'Counter-Integrative Thinking': The View from Social and Cognitive Psychology, Circa 2005 AD.

Cognitive and social psychology are not unitary or mono-paradigmatic fields. They feature many different conceptualizations of thinking, each leading to different research programs. There is, however, a stream of research that concentrates on uncovering the mechanisms by which the human mind indulges its proclivities to bracket, narrow, schematize, simplify and reduce its perceptual, cognitive and representational fields, and this literature provides a coherent and powerful picture of counter-integrative thinking: in its average everydayness, the mind as portrayed by this research program seems most prone to tendencies that run precisely counter to the cognitive and epistemic virtues sketched above. The adaptive, integrative way of being – of it is to emerge – must be brought out by interventions aimed at precisely annulling the counter-integrative tendencies that many psychological empiricists believe to be a constitutive part of the human condition.

2.1. The Phenomenology of Premature Cognitive Commitments: Seizing and Freezing. Kruglanski and Webster (1996) synthesize a large research program that studies the ways in which people form, validate, select, justify and retain beliefs by suggesting that "closing" of the mind is motivated and takes the form of seizing – usually sub-optimally – on beliefs, models and theories that come to mind early in the reasoning process, and by freezing – again, sub-optimally – commitment to such beliefs, thereby curtailing opportunities for learning and cognitive change. Both seizing and freezing are active processes, supported by significant cognitive effort: to shut out, distort, filter or discount new information, and to justify the current cognitive commitment. Both, seizing and freezing – related processes seem to be resistant to pedagogical interventions that aim to highlight the cost of suboptimal early cognitive commitments. Thus, while the closing of the mind may be motivated, this does not mean that it is under conscious control, which, in turn, means that 'opening' the closed mind is not something that can necessarily be consciously willed.

Counterpoint CP2.1.1. It is important to realize that cognitive psychologists make inference about human minds in general on the basis of noisy studies of individuals drawn from a restricted population (usually, college undergraduates). 'Noisy' means that at least some of the experimental subjects tested did not exhibit the effect

associated with cognitive narrowing. The restricted sample population makes it possible that the effects are artifacts of background, self-selection, institutional setting and developmental stage of the individual participants. The observations bring up:

RQ2.1.1.1 What explains individual differences in findings regarding motivated closure of the mind? Are there stable personality traits or developmental variables that predict immunity to closure-of-the-mind biases? If so, then these trait and state variables can be turned into valuable predictors of 'open-mindedness'.

RQ2.1.1.2 Are there situational differences that influence motivated closure of the mind? If so, what are they? Meditative practices, mindfulness interventions, epistemological training, practice in psychotherapeutic and psychoanalytic techniques are all possible mind-opening situational variations, but have yet to be turned into experimental interventions that undo the effect making up the 'motivated closing of the mind' literature.

Counterpoint CP.2.1.2 We know 'about' the open mind via our intuitions and observations of the doings of the close mind, but not directly; we speak, generally, of being more or less closed minded but have little insight into degrees of openness, much less into degrees of openness subject to the constraints of actionability. In short, a pragmatically informed theory of open-mindedness is missing, which brings up:

RQ2.1.2.1 What is an open mind, technically speaking? How do we know it by its signs?

RQ2.1.2.2 What can we say precisely about the trade-off between openness and actionability? What constraints does actionability impose on openness? At one end of the spectrum, one cannot act on account of having to maintain oneself even ready to change the beliefs on which purposive action must rest. At the other end of the spectrum, one cannot allow oneself to think and feel one's way around a situation on account of having to shut down inquiry and commit early to a particular set of beliefs. What are the tradeoffs between these extremes entailed by different approaches to inquiry?

2.2. Ego-centric Filtering processes: Self-Induced Bracketings of Cognitive and Perceptual Domains. Greenwald (1980) distills almost a century of psychological research on self-serving, self-protecting and self-enhancing biases by positing the self as a totalitarian information processing regime and an image that the regime props up. A totalitarian regime suppresses and distorts uncomfortable information and, just as importantly, edits and revises its own history – the narrative account of its past deeds, consistently with the observation – due to George Orwell – that he who controls the past controls the future. In similar fashion – argues Greenwald – empirical evidence on self-protecting and self-enhancing proclivities can be made send of by positing an agent – the self- that at once is an image and a process aimed at maintaining and justifying this image, by editing and re-interpreting memories and by filtering and distorting new information – particularly information that may contradict central tenets and dogmas of the self. Greenwald's practice of cognitive proclivities cautions that generally prevalent "seizing and freezing" tendencies may be amplified when the object of cognition is the self: Higher "affective temperature" begets higher levels of (motivated) cognitive activity.

Counterpoint CP. 2.2.1. What of consciousness of oneself? Does it – will it – make a difference to the natural proclivities of the self? Once again, given that motivated behavior is not necessarily purposeful and consciously willed, it is not clear that conscious awareness – at a theoretical level – of the self’s totalitarian tendencies will make a difference to these tendencies. It is however, plausible that if the self (as an image) becomes expansive and complex enough to incorporate awareness of its own tendencies into its own body (so to speak), than there will be an impact on these very same tendencies. More generally, it is to be expected that the nature of the self (as image) will interact with its activities: fragmented, diverse, dissociated selves will not likely pursue the same totalitarian tendencies as those pursued by unitary, coherent selves, which suggests:

RQ 2.2.1.1 What are the effects of individual differences in self-image on the incidence of totalitarian proclivities? Will Buddhists, psychotherapists, schizophrenics, entrepreneurs exhibit the totalitarian cognitive proclivities observed by psychologists in their sample populations? What are the personality-level trait correlates of totalitarians and counter-totalitarians?

RQ 2.2.1.2 What are the optimal trade-offs between actionability and decisiveness (which totalitarian tendencies often favor and facilitate) and openness and responsiveness to the new (adaptivity), which affectively “weaker” selves may have easier access to?

2.3. Ego-centric Processing of Allocentric Information: The Self as a Problem. When interacting with others, it is at once indispensable to form beliefs and conjectures about what they think and very difficult (and unlikely) to form them correctly. Nickerson (1999) posits that the experimental evidence on the ways we think about what others think suggests that we start from the simple – and often misleading – ‘default’ knowledge of what we think; we assume, thus,, that others’ minds are like ours – both in fundamental ways (in the rules of grammar and inference they obey) and in specific and particular ways (in knowledge of facts, models, theories, beliefs, etc.). We take an egocentric approach to the processing of allocentric information – of information about the other – thereby making false inferences and incurring the often high costs of misunderstanding begetting alienation, hostility and resentment.

Counterpoint 2.3.1. Is egocentric processing of allocentric information an artifact of cultural and developmental background? It is conceivable that egocentric processing of the ‘other’s mind’ is adaptive in a culturally uniform environment (such as middle 20th century North America) but will not be observed in other – more diverse socio-cultural environments. It is also possible that the experience of cultural transplantation – of forced change of socio-cultural environment – undoes the egocentric bias in the processing of allocentric information. This suggests:

RQ2.3.1.1. What are the cultural and developmental determinants of individual-level differences in egocentricity in the creation of an image of the other’s mind?

Counterpoint 2.3.2. When Jean-Paul Sartre affirmed, “L’enfer, c’est les autres,” he may have been commenting on the personal difficulty of relating to another person authentically and meaningfully. This difficulty may be of a cognitive or affective nature and may stem from the absence of basic skills or personal technologies for intervening

successfully upon the mind of another. On the cognitive side, the kind of interactive reasoning required in game-theoretic contexts, for instance, may 'transfer' to generalized interpersonal situations. On the effective side, a personal commitment to an interaction principle such as 'don't guess, ask!' [Martin, 2004] when in doubt about what someone else thinks may be useful in overcoming the emotional barrier to inquisitive behavior that mires interpersonal interactions into tacit mutual outguessing and false imputation.

RQ 2.3.2.1. Can cognitive skills be transferred from interactive game-theoretic reasoning to interpersonal situations? If so, then under what conditions?

RQ 2.3.2.2. Can training in the fundamental fallibility of knowledge in general and interpersonal knowledge in particular induces or facilitate the initiation of interpersonal inquiry as a basic mode of communication?

2.4. Credulity and the Failure of Discernment. The 'average mind' in its everyday behavior seems to be sub-optimally credulous and recalcitrantly viscous; it does not seem to disentangle understanding a fact from accepting the fact as true and often proceeds as if the fact is true even after being given reason to think it false (Gilbert 1991). This 'empirical regularity' is fundamentally problematic for those who would like to portray the mind as a marvel of adaptability, because grasping the new should precede (rather than coincide with) evaluating it and responding to it purposefully. At the 'other end', forgetting false alarms quickly (facts that seemed true but turned out to be false) is also important to the optimally adaptive mind, for the purpose of maintaining sharpness of focus and efficiency of storage in working memory.

Counterpoint CP.2.4.1. 'Facts' are not simple objects. They are claims to attention and awareness (a scarce resource) and claims to validity on 4 different levels: ontological, epistemological, substantive and pragmatic. They establish a set of objects that exist (ontological), presuppose ways of validation or refutation (epistemological) state "what is the case" (substantive) and presuppose or prescribe a certain interpersonal context or course of action (pragmatic). Given this fact about the inherent complexity of facts, it is important to understand the underpinnings of credulity and viscous discounting of errors more carefully. For some fact, understanding may require a predisposition to act in accordance with them, which would mean that it is not possible to understand without experiencing the emotional state of acceptance. For others, detached understanding without either acceptance or rejection may be neither required nor desirable. Similarly, the irreversible nature of erasing erroneous facts from memory may make it rational for deletion decisions to be tentative (at least for a period of time) and to have a finite, but non-zero, extinction period. Understanding the dynamics of credence will require the development of a model of facts that incorporate their four-dimensional structure and of a model of the processes of understanding and credence that is responsive to this complex view of facts. Once this model is developed, interventions that discourage or inhibit sub-optimal attachment to erroneous or nonsensical facts can be articulated, tested and refined, and the following questions can be tackled:

R.2.4.1.1. Can sub-optimal conflation of understanding and credence be diminished by training in the basic operations of deconstruction that psycholinguistics and analytical philosophy has developed? (Implicitly, are physicists and analytic

philosophers less prone to conflate understanding and credence than are people without similar training?

R.2.4.1.2. *Will interventions aimed at inducing a playful mindset (along the ontological, epistemological, substantive and pragmatic dimensions) – mindfulness interventions, for instance – diminish the proclivity to associate understanding with credence? Will training in ‘mindful’ ways of being change dispositions to converge prematurely on sticky cognitive and perceptual structures?*

3. The Meta-Psychology of Counter-Integrative Thinking: Automaticity and the Difficulties of Thought Control and Self-Regulation.

We introduce a somewhat artificial distinction between the research discussed in the previous section and that discussed in the current one, namely, that between studies concerning themselves with cognitive processes, mechanisms and tendencies and studies that concern themselves with meta-cognitive processes, mechanisms and tendencies. The mind, uniquely, can reflect on itself, perform work on itself, do or undo itself because it possesses the capacity to monitor and act upon its own output. The research we are about to attempt to understand asks questions about just this capacity: does it exist or is it itself only a creation of the mind? If so, by what mechanism and to what end has it been created? Why, when and how do self-monitoring and self-control processes work in the realm of thinking? Is the mind truly an uncontrollable organ, as Tom Schelling has suggested? If not, what are the limits of its controllability? Is control of the mind something we can hope to teach? Develop? Select for? If so, how? Can we reasonably contemplate a science of the economics of mind, a science of successful self-control – of ‘egonomics’ (Tom Schelling’s phrase again) – and artificial science, no doubt – but one that allows us to allocate scarce resources (attention, calculation, perception, reflection) more successfully, to ameliorate the rate at which we fall into traps of thinking, of seeing, of doing and not-doing?

3.1. Automaticity of Higher Cognitive Processes: The Ubiquity of Mental Habits. Bargh and Ferguson (2000; see, also, Bargh and Chartrand, 1999) argue from experimental evidence produced by themselves and others that we get along ‘just fine’ (i.e. are behaviorally competent) in a large number of personal and interpersonal contexts without the need for conscious, purposeful processing of external stimuli. Functioning ‘on autopilot’ is prevalent as well as potentially functional. They marshal an array of experimental results that purport to establish the ubiquity of cognitive automaticity: automatism is the rule rather than the exception; early deterministic approaches to the study of behavior (‘behaviorism’ and the behaviorist tradition) had gotten it right – it seems – with respect to the deterministic nature of the right explanation of behavior, but wrong with respect to the locus of the cause, which is not the external stimulus but the internal, unconscious process that causes the response to the stimulus. Thus, even as behaviorism fades away into the darkness of the annals of science as an outmoded mechanistic model of human ways-of-being, determinism shows up again as a powerful explanatory tactic, re-embodied in models that stress the automaticity of higher cognitive processes. Wegner and Wheatley (1999) go deeper down the same explanatory path by positing (Figure 1 of their paper) a model whereby an unconscious mental process leading to (an unconsciously caused) action is the ‘real’ causal mechanism underlying behavior and the experience of the ‘will’ simply has to do with a judgment regarding the plausibility of a particular perceptual or cognitive state (‘thoughts’, ‘images’) given the actions one has just observed oneself undertaking – a sharper re-

statement of earlier, Bemian self-perception theories in psychology, whereby we infer our preferences from our choices and behavior, rather than vice-versa.

The implication of these findings are powerful: counter-integrative tendencies are 'sealed' by higher-level cognitive routines that look, feel and operate like habits of the mind, or mental tics. Closure of the mind of the type contemplated by Kruglanski and Webster (1996), sub-optimal perceptions of understanding and approval of the type described by Gilbert and self-defeatingly self-centered biases in the processing of ego-centric and allocentric information (Nickerson, 1999) are protected from outside intervention by (1) the automaticity of precisely the higher cognitive processes that could potentially be used to 'unfreeze' counter-adaptive cognitive and perceptual tendencies, (2) our unconsciousness about these automatisms and (3) the ability of the mind to manufacture illusions of conscious will and purposeful direction that mask the underlying automaticity of the higher mental processes. One can only be pessimistic, on statistical grounds, about the possibility of training-based interventions that undo the proclivities that we associate with counter-integrative thinking. The self-defeating tendencies that one attempts to redress or correct, behave – in the laboratory – 'as if' they are hard-wired, a feat accomplished by the concatenation of habituation and self-deception that the psychological empirical record attests to.

Counterpoint 3.1.1. Individual and Developmental Differences. Any set of experiments produces noise, and the noise produced by psychological experiments is often far more informative than the noise produced by experiments in the natural sciences. It is often a curious kind of noise, wherein all unexplained data points are aggregated together. In spite of the illuminating results produced by the analysis of individual differences in judgment formation [Stanovich and West, 2000], which show what *kind* of people (personality-wise; intelligence-wise) are likely to produce the 'incorrect' response patterns on standard tests of epistemic rationality, no similar research program in individual differences in automaticity of higher and lower cognitive processes has been initiated (in part due to the fact that it is only recently that research on automaticity has congealed into a research program or a paradigm). It is possible, however, that stronger ego-defensive mechanisms and stronger egocentric biases may be more closely correlated with automatic processing of stimuli than weaker or many-faceted self-schemata and similar differences may be observed between exploitation and exploration-oriented individuals (a personality trait). Moreover, cognitive development and developmental influences should also be expected to matter to the occurrence and prevalence of automatisms. Much of 'higher education' unfortunately corresponds to the development of cognitive habits (and relatively few cognitive counter-habits) that are optimally (but, statically) adapted to the routinized processing of information presented in standard formats and resting on standardized sets of assumptions. These cognitive habits are drilled into sub-liminal consciousness by sharp incentive structures that stress speed and cognitive efficiency over adaptiveness. It is not unreasonable for rational agents to develop cognitive 'production functions' that resemble mass-manufacturing systems rather than nimble customized design systems. Thus:

RQ3.1.1.1. What difference do individual personality and general intelligence make to the occurrence, prevalence and persistence of automatisms? What difference does the presence of lower-level automatisms make to the occurrence of higher level cognitive automatisms? What difference does the presence of perceptual automatisms make to the occurrence and prevalence of higher level automatisms?

RO3.1.1.2. What difference, if any, do developmental experiences such as education and training programs and experience make to the occurrence and prevalence of automatisms in higher-level cognitive processing? Are there differences in the degree to which various kinds of educational programs induce automatisms in higher level cognitive functions?

Counterpoint 3.1.2. Cognitive Interventions Not Explained by the Paradigm. Langer (1989) has developed, in the Western tradition, a line of thought introduced in the Eastern philosophical and religious traditions which focuses on states of consciousness ('mindful' states) that are at once not effortful to produce *and* seemingly not susceptible to cognitive and behavioral automatisms. Mindfulness interventions developed by Langer and her co-workers have been used to inhibit the formation of premature cognitive commitments, to heighten the subjectively experienced and objectively measured levels of personal control over complex predicaments, and to heighten the sense of 'presence' of the subject experienced by an observer. These interventions are simple to replicate and implement in the laboratory setting. They revolve around setting up a cognitive 'play situation', wherein the meaning of words (their extension and reference conditions), the nature and foundations of knowledge, perception and belief, the nature and function of rules, principles and laws are individually or conjointly 'played with' [Moldoveanu and Langer, 1999]. These interventions, however, have yet to be applied in a systematic fashion in order to test their ability to undo the automaticity biases documented by the tenacious research stream of Bargh, Wegner and their co-workers.

Similarly, the study of cognitive skill transfer [Barnett and Ceci, 2002] has generated a significant research stream documenting the selective ability of researchers and educators to produce the transfer of cognitive skill across different perceptual, pragmatic, social and contextual domains. The basic question of cognitive skill transfer research is simple: will learning to perform a particular mental task enhance performance on a different, only partially related task? Will learning statistical mechanics, for instance, lead to skill in developing new variants of the Black-Scholes option pricing formula, whose underlying mathematics is based on the analytical apparatus of statistical mechanics)? Will learning analytical geometry lead one to be more successful (than one was previous to learning it) in designing room layouts? Will learning a new Romance language lead to better performance in the learning of a new Romance language? Singley and Anderson [1989] report that cognitive transfer indeed *works* if the cognitive 'production' of the two tasks between which transfer is expected are similar. Cognitive productions are algorithmic representations of the mental tasks that underlie the two tasks. The cognitive production of figuring out the competitive equilibrium in a game, for instance, is the same as that of figuring out if a graph (network) has a clique of a particular size (a set of nodes that are all interconnected). The *fact* that cognitive skill transfer can be achieved in some settings can be interpreted as the starting point of a more over-arching criticism of a point of view of human thought as inevitably tied to unconscious automatisms: the transfer of cognitive skill is precisely about the modification of higher level cognitive processes that one would normally associate with a particular stimulus. In a successful case of cognitive skill transfer, for instance, one would have to change the way one would cognitively process a stimulus, which, if automatisms were unavoidable, would function as an inescapable trigger for automatisms that would inhibit the exercise of

newly acquired cognitive skills. Moreover, if adaptive algorithms representative of integrative thinking represented kinds of cognitive skills that can be transferred, one has reason to ask: would the successful transfer of such skills undo cognitive automatisms? Thus:

RQ3.1.2.1. Will play-oriented interventions disrupt higher-order cognitive automatisms? What mediates the effects of such interventions and why?

RQ3.1.2.2. Can adaptive algorithms be transferred in the form of cognitive skills? If so, under what conditions? What are the effects of the successful transfer of such skills on cognitive automatisms?

Counterpoint 3.1.3. Non-Cognitive Interventions Not Addressed by the Paradigm. Lieberman [2002] argues for intuition as a non-conscious, non-explicit faculty, and for implicit learning as one of the functions of intuition that can be inhibited by explicit learning of the rules by which an implicitly learned behavior is produced. Rizzolatti and his co-workers [Gallese, Keysers and Rizzolatti, 2004] argue that a mirror-neuron system and a visceromotor system that are physiologically independent of the pre-frontal cortex (whose activities biologically ground higher cognitive processes) provide neurophysiologically plausible mechanisms that explain how we intuitively understand each others' actions and emotions, respectively. No attempt has been made thus far, however, to identify possible interactions between non-cognitive interventions (at the level of intuitive understanding) and the higher-level cognitive automatisms described in cognitive psychology. For example, it may be imagined that repeated exposure to a highly adaptive, cognitively and emotionally nimble creature can stimulate emulatory patterns of behavior (and thought) which, though not named or represented or codified in a formal system, nevertheless become embodied in a new set of patterns of thinking that escape the clutches of previous automatisms. It may also be conjectured that *living* a particular way of life, structured and constrained by interactions with people jointly following mental maps and models of interactions different from one's own can eliminate cognitive automatisms underlying the egocentric processing of information about other –people (using oneself as a model for the content of the mind of another), to which end a taxonomy of ways of being-towards-others and of logic of interaction [Fiske, 1992] can help with the design of interventions and programs of training meant to loosen egocentric automatisms that color our interactions with the world. Thus:

RQ3.1.3.1. Can higher level automatisms be undone by direct exposure to and interaction with behaviorally and cognitively non-routinized individuals? Can automatisms be implicitly undone, even if they cannot be explicitly accessed? If so, then, under what conditions?

RQ3.1.3.2. Will changes in the social environment and the need to interact with group of people that follow fundamentally different mental models of interaction from one's own lead to loosening of egocentric cognitive biases and the automatisms that insulate them from change? For example, Fiske [1992] describes four fundamental ways of interaction in social groups based on an extensive anthropological and ethnomethodological set of studies and review of the literature based, respectively, on the logics of equality matching (interpersonal comparisons), authority ranking (power), market pricing (efficiency and value) and communal sharing (common

ownership and welfare), which can help with the formulation of experimental paradigms for understanding the individual's mental map of the 'other' and the 'others' and with structuring canonical interaction patterns for exposing individuals exhibiting a certain logic of interaction to groups exhibiting fundamentally different logics of interaction.

3.2. Costliness of Self-Regulation and Self-Control and Failures of Thought Control. One possible explanation for the stubbornness of automatisms (both behavioral and cognitive) is the fact that self-regulation and self-control is a costly process [Baumeister and Heatherton, 1996; Muraven and Baumeister, 2000]. The problem of self-control and self-regulation appears whenever we try to get ourselves to do something we do not naturally and willingly incline to do *or* when we try to get ourselves to *not* do something that we naturally and willingly incline to do (i.e. that we would habitually do). The authors cite a large body of evidence in psychological research that suggests that the capacity for self-control and self-regulation functions as a limited resource. For example, individuals asked to abstain from eating cookies that they crave perform more poorly on controlled processing tasks such as naming the color of the ink in which a color-naming-written work that names a different color than that of the ink than individuals who do not have to manage their appetites in the same way. The authors argue that self-control acts like a muscle in two ways: it is depleted by strenuous use, and it is built up by exercise: a two-week training period in which self-control is repeated and strenuously exercised (by self-enforced adherence to postural rules and eating regimes, for instance) leads to better performance on controlled processing tasks. Ainslie [2001] takes a slightly different approach to the description and operationalization of self-control, based on his studies of addictive psychiatric patients, as well as 'normal' subjects. He argues that 'sane' preferences for long-run goods (good health, the end of addiction) are foregone for short-run temptations (like eating unhealthily, smoking or drinking or taking psychoactive drugs) because we discount rewards hyperbolically rather than exponentially and the short run reward always looms larger than the long run reward when it gets temporally close to us. Our preferences thus flip at the last moment the moment it comes time to actually abstain from the forbidden or undesired behavior. He argues that rules, principles and categorizations evolve to help us deal with this problem of self-control, because they 'lump together' a large set of short run rewards and allow us to choose between them as a set and the long-run rewards as a set. Thus, abstaining from *this* cigarette (here and now) is encoded as abstaining from *smoking* in general, which makes it easier to forego *this* cigarette here and now. The categorization is, of course, buttressed by the causal rule: 'Smoking leads to more smoking (in the future).' Thus, rules and principles are functionally explained by the need to implement self-control by countering the effects of hyperbolic discounting.

Counterpoint 3.2.1. If we understand thinking as a form of behavior, adaptive thinking as a form of controlled processing and habitual patterns of thought as the 'comfortable' behaviors one must try to correct, it becomes possible to tap into the self-control literature for insight into the design of interventions aimed at developing more adaptive forms of thinking. Both the approaches of Baumeister and of Ainslie lead us to think that designing self-training programs based on the articulation of personal rules and self-enforced adherence to these rules can function to increase cognitive control (albeit for different reasons, which makes it interesting to articulate

experiments that help to tease apart the two effects). However, personal rules are *precisely* the stuff that automatisms are made of (something that Ainslie cautions us about): the path *away* from automatisms (which we habitually fall into) travels across automatisms (which we use to build up the self-control needed to escape automatisms). What to do? The key is to realize that not all automatisms are 'created equal' and that some rules and principles are more adaptive than others: 'Make everything as simple as possible but not simpler' (Einstein) is a thought heuristic that can discipline the imagination without inducing the rote, closed-ended automatic thinking that other rules ('obey the rules of Bayesian kinematics when reasoning about the unknown') can lead us to. We can think, then, of constructing classes and of rules for thinking that trade off optimally between their proclivity to build up self-control (and enhance our ability to escape automatism) and their proclivity to lead to closed-ended patterns of thought (which is precisely what we want to avoid). Thus:

RQ3.2.1.1. What are 'optimal rules for thinking', according to the criteria enunciated above? What effects do they have on the adaptivity of cognition?

RQ3.2.1.2. How can these rules be optimally implemented? Does explicit training in the use of these rules lead to lasting enhancements of the ability of subjects to control their own minds? Can these rules be implicitly trained? What are the differences between implicit and explicit training procedures for these rules?

3.3. Failures of Thought Control. Preliminary work on thought control [Wegner, 1994] suggests that our ability to control our own thoughts is limited, if not non-existent: the mind appears to be an 'uncontrollable organ' – as Tom Schelling has put it. We do not seem to be able to *not* think – on command, of particular objects (a 'white bear' in the Wegner experiments): subjects exhorted to not think of a white bear are *more* likely to think of a white bear than subjects primed on white bears but not explicitly asked to not think of a white bear. The finding resonates with other works of Wegner and his co-workers attempting to highlight the inescapable automaticity of action and thought and the illusory nature of the experience of 'will'. Of course, the work of Baumeister and his co-workers cited above suggests that the will is difficult to exercise, but not impossible: and the conflation of difficulty with impossibility should be in question.

Counterpoint 3.3.1. The experimental program on thought control relies on a metaphor of thought control as a simple feedback process comprising an executive function (the producer of thoughts) and a monitoring process (looking for the thoughts that must be suppressed) and argues that the very functioning of the monitoring process interferes destructively with that of the executive process (because it generates the very thoughts the executive process is trying to avoid). The injunction 'don't think of a white bear' produces more white bear thoughts than does its absence. There is, however, a way to not think of a white bear, and that is to strenuously think of a non-black non-bear (like a black dog or green basilisk). Thus, it may be expected that the *objective function* of the monitoring process can be designed to enhance the thought control abilities of individuals. In particular, in addition to examining the reasons for individual, developmental and contextual differences in thought control abilities, we can ask:

RQ3.3.1.1. Will thought control ability be enhanced by changes in the basic rules by which the monitoring process is set up? Will changing the 'objective function' of the monitor lead to improvements in the performance of the overall feedback system?

RQ3.1.1.2. Can optimal objective functions for the monitoring process be articulated and classified?

RQ3.1.1.3. Can the ability to design optimal objective functions for the monitoring system be trained and/or enhanced?

4. Summary and the Way Forward.

The experimental record in cognitive psychology provides a powerful guide to counter-integrative thinking proclivities, as a well as a powerful map for a new wave of research that can be expected to produce new interventions and selection mechanisms for the productions of adaptive, successful patterns of thinking:

1. Most individuals crave certainty and stability and attempt to achieve them through narrowing of the perceptual and cognitive field, entrenchment of existing but possibly outdated models and patterns of thought, ego-centric processing of information about others and maladaptive conflation of understanding and belief (Figure 1);
2. These tendencies at the perceptual and cognitive levels are buttressed by higher level cognitive processes that exhibit such high levels of automaticity that they seem hard-wired, are resistant to attempts at thought control and change (which is a costly activity) and are resistant to attempts at change through explicit training (Figure 2); however,
3. There are significant individual, developmental and situational differences in thinking patterns of studies exemplifying counter-integrative patterns of thought, and a careful research program on identifying these differences can lead to powerful mechanisms for selection, intervention and training of adaptive, actionable thought patterns (Figures 3,4); moreover,
4. Cognitive skill *can* be transferred across different tasks and the ability to control cognitive processes can be developed through targeted interventions and training programs, making it probable that training programs aimed precisely at the transfer of cognitive skill and the development of transferable cognitive skills can be developed. Since the existence of transferable cognitive skill is a key component of adaptive thinking, successful demonstrations of cognitive skill transfer entail also the successful demonstration of cognitive adaptation (Figure 5). Of course, the 'optimal path' to the transfer of cognitive skill may not be a cognitive (explicit) one, and the interaction between the implicit and explicit dimensions of reasoning will have to form the basis of fresh new research;
5. There is no evidence to suggest that meta-cognitive skills (the ability to change patterns of thinking and patterns of thinking about patterns of thinking) cannot be changed, but this is because meta-cognitive processes have not been seriously investigated in ways that heed both the rich philosophical and spiritual traditions of the east and the West *and* the means of investigating, representing and intervening of

modern social and cognitive psychology. Indeed, such an exploration is now

With Cognitive Psychology: The Way Forward I

The Way Forward I Individual Differences

Both cognitive tendencies

- 'seizing and freezing
- 'imperfect suspension of belief
- 'egocentric perception of others
- 'ego-preserving editing & distortion of belief & perception

...and meta-cognitive effects

- 'automaticity of cognitive processes'
- 'selfregulation failure
- 'imperfect thought control

... are corroborated by experimental results exhibiting significant individual differences, which creates an opportunity to find and study individuals who are not prone to these effects:

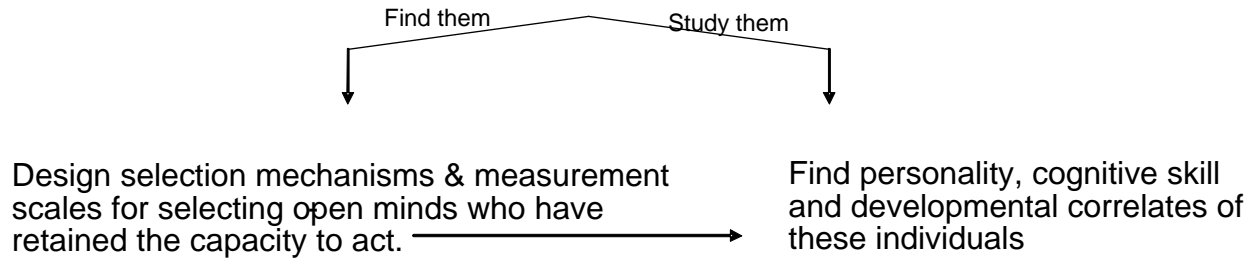


Figure 3.

Main Works Cited (Reproduced *Infra* for Limited Distribution).

Bargh, J.A. & Chartrand, T.L. (1999). The Unbearable Automaticity of Being. *American Psychologist*. (Vol. 54, No. 7, pp. 462-479). Washington, DC : American Psychological Association.

Bargh, J.A. & Ferguson, M.J. (2000). Beyond Behaviorism: On the Automaticity of Higher Mental Processes. *Psychological Bulletin*. (Vol. 126, No 6, pp. 925-945). Washington, DC: American Psychological Association.

Barnett, S.M. and Ceci, S.J. (2002). When and Where Do We Apply What We Learn? A Taxonomy of Far Transfer. *Psychological Bulletin*. (Vol. 128, No. 4, pp 612-637). Washington, DC: American Psychological Association.

Baumeister, R.F. and Heatherton, T.F. (1996). Self-Regulations Failure: An Overview. (Vol. 7, No. 1, pp 1-15). *Psychological Inquiry*. Mahwah, NJ. Lawrence Erlbaum Associates.

Bugental, D.B. (2000). Acquisition of the Algorithms of Social Life: A Domain-Based Approach. *Psychological Bulletin*. (Vol. 126, No. 2, pp 187-219). Washington, DC: American Psychological Association.

Fiske, A.P. (1992). The Four Elementary Forms of Sociality: Framework for a Unified Theory of Social Relations. *Psychological Review*. (Vol. 99, No. 4, pp 689-723). Washington, DC: American Psychological Association.

Gallese, V., Keysers, C. and Rizzolatti, G. A Unifying View of the Basis of Social Cognition. *TRENDS in Cognitive Sciences*. (Vol. 8, No. 9, pp 396-403). New York, NY: Elsevier.

Greenwald, A.G. (1980). The Totalitarian Ego: Fabrication and Revision of Personal History. *American Psychologist*. (Vol. 35, No. 7, pp 603-618). Washington, DC: American Psychological Association.

Gilbert, D.T. (1991). How Mental Systems Believe. *American Psychologist*. (Vol. 46, No. 2, pp 107-119). Washington, DC: American Psychological Association.

Kruglanski, A.W. & Webster, D.M. (1996). Motivated Closing of the Mind: "Seizing" and "Freezing." *Psychological Review*. (Vol. 103, No. 2, pp. 263-283). Washington, DC: American Psychological Association.

Langer, E.J. (1989), *Minding matters: The Consequences of Mindlessness-Mindfulness*, *Advances in Experimental Social Psychology*, Vol. 22:137-173.

Lieberman, M.D. (2000). Intuition: A Social Cognitive Neuroscience Approach. *Psychological Bulletin*. (Vol. 126, No. 1, pp 109-137). Washington, DC: American Psychological Association.

Muraven, M. and Baumeister, R.F. (2000). Self-Regulation and Depletion of Limited Resources: Does Self-Control Resemble a Muscle? *Psychological Bulletin*. (Vol. 126, No. 2, pp 247-259). Washington, DC: American Psychological Association.

Nickerson, R.S. (1999). How We Know – and Sometimes Misjudge – What Others Know: Imputing One's Own Knowledge to Others. *Psychological Bulletin*. (Vol. 125, No. 6, pp 737-759). Washington, DC: American Psychological Association.

Wegner, D.M. (1994) Ironic Processes of Mental Control. *Psychological Review*. (Vol. 101, No. 1, pp 34-52). Washington, DC: American Psychological Association.

Wegner, D.M. & Wheatly, T. (1999). Apparent Mental Causation: Sources of the Experience of Will. *American Psychologist*. (Vol. 54, No. 7, pp 480-492). Washington, DC: American Psychological Association.

Other Cited References

Bauer, R. (2005), *Dilemmas, Paradoxes and Quagmires*, presentation to MBA Integrative Thinking Practicum.

Kuhn, T. (1990), *Rationality and Theory Choice*, in T. Kuhn, *The Road Since Structure*, Cambridge: MIT Press.

Lakatos, I. (1970), *Falsificationism and the Methodology of Scientific Research Programmes*, in I. Lakatos and A. Musgrave, eds. *Criticism and the Growth of Knowledge*, New York: Cambridge University Press.

Martin, R.L. (2004) personal communication.

Martin, R.L. (2002), *Integrative Thinking*, presentation to 1st year MBA Class.

Moldoveanu, M.C. (2002), *How to Survive Your Beliefs*, presentation to Young Presidents' Organization, Rotman School of Management.

Moldoveanu, M.C. and E.J. Langer, (1999), *Mindfulness*, in M. Runco and L. Ricci, eds., *Encyclopedia of Creativity*, San Diego: Academic Press.

Popper, K.R. (1959) *The Logic of Scientific Discovery*, London: Routledge.

Popper, K.R. (1999), *All Life is Problem Solving*, London: Routledge.

Rokeach, M. (1968), *Open and Closed Minds*, New Jersey: Erlbaum.

Singley, J. and J.R. Anderson (1989), *Cognitive Skill Transfer: The ACT* Approach*, New Jersey: Erlbaum.

Stanovich, K. and R. West, (2002), *Individual Differences in Reasoning: Implications for the rationality debate?*, *Behavioral and Brain Sciences*, Vol. 23:645-726.