BEYOND *B*EHAVIORAL *E*CONOMICS *N*UDGE (*BEN)*:

*B*OUNDED *R*ATIONAL *A*DAPTIVE *N*UDGE (*BRAN)*

Behavioral Insight Group

Kennedy School of Government

Harvard University

4/11/2016

by

Riccardo Viale

(*Scuola Nazionale dell'Amministrazione, Rome*

and Herbert Simon Society)

DRAFT

1) Introduction: BEN and Variants of Libertarian Paternalism

An intervention is classified as a *nudge* when it is not a coercive measure, retains freedom of choice, is based on automatic and reflex responses, does not involve methods of direct persuasion, does not significantly alter economic incentives, and does revise the context of choice according to the discoveries of behavioural economics (Thaler and Sunstein, 2008). What is proposed is therefore a form of *libertarian paternalism* that has a dual valency. As *paternalism*, it aims to make up for citizens' irrational and self-harming tendencies by "gently nudging them" to decide rationally for their own good. In its *libertarian* form it aims to give the last word to the outcome of the conscious and deliberative processes of the individual citizen who can always choose to resist the *nudge*.

As we will see, some crucial problems emerge from this particular choice architecture. First, what rationality model is used by "nudgers" to intervene in the choices made by citizens. Second, there is the question of which phase of the citizens' decision-making process to intervene in, namely what sort of paternalism to adopt.

Thaler and Sunstein's thesis is that citizens are subject to many deviations of rationality that bring them to make a decision that is against their own interests. This widespread irrationality is provoked by the automatic judgement and decision-making behaviour enabled by heuristics. In this way, the individual is incapable of following a series of basic principles of rationality at a probabilistic and logical level. By doing so, he makes the wrong choice and goes against his interests.[1] Thaler and Sunstein's theses derive in part from the programme of behavioural

[1] The identification of biases and formal errors of judgement is not a recent phenomenon. Adam Smith (1759) highlighted both the phenomena of "loss aversion" and "hyperbolic discounting" as explanatory factors for human

economics commenced in the postwar period by Herbert Simon and continued in the 70s by Daniel Kahneman, Amos Tversky and their school. However, this approach is based on a distorted concept of *bounded rationality* focused on the study of the correspondence between judgement and decision-making performance in tests and laboratory simulations and canonical models of probabilistic and deductive rationality. It therefore highlights a constant and systematic irrational misalignment between behaviour and norm. In this sense the acronym BEN or Behavioural Economics Nudge can be said to correspond to the initials of Benjamin, namely Benjamin Franklin, the champion of formal decision-making rationality (his Moral Algebra), the ideal of rationality for behavioural economics. This approach discounts a series of weaknesses in various analytical dimensions of an epistemic, epistemological, methodological and ethical nature. As is clear from the Simonian framework and from subsequent developments (Gigerenzer and Selten, eds., 2001 ; Viale, 2012) there are four attributes of the concept of *bounded rationality* that contrast with the current approach of behavioural economics that inspired *Nudge* theory. This is not a rationality that is interested in questions of form, but rather in the question of ecological adaptation to the environment of choice and problem-solving. It centres on the procedural and realistic attributes of rationality and not on instrumental and conventionalist ones. It recognises that the complexity of the real environment of choice brings the player face to face with decisions in conditions of uncertainty (for example, "ill structured problems" like financial markets or political forecasting) rather than risk (for example, "well structured problems", like dice throwing or chess). Lastly, it appears sceptical regarding the ability to divide the mind into separate systems of the intuitive-unconscious type and the analytical-conscious type (Viale, 2012; Macchi, Bagassi and Viale, eds., 2016). This rules out the specific argument in favour of the libertarian factor present in Nudge theory.

If it seems correct to assert that the behavioural economics underlying Nudge theory has a normative and non-adaptive conception of rationality, which can only be correctly applied in conditions of risk and not uncertainty, and which is based on empirical data derived from abstract tests that overlook that pragmatic aspects typical of everyday decisions, then we can ask what might be the contribution of adaptive behavioural sciences, namely having removed these underlying defects, to the day-to-day lives of individuals. In short, what we are trying to understand is the form of *libertarian paternalism* that adaptive behavioural sciences can offer.

There are various levels of possible behavioural intervention in citizen choice that are characteristic of paternalism. In the first place we need to understand whether or not any kind of paternalism can be completely libertarian. Thaler and Sunstein's thesis is that every choice architecture must allow the individual to make a conscious decision to accept or reject the proposed *nudge*. For example, being subject to weakness of will and the *status quo bias*, citizens, in some countries like France, are automatically enrolled by default as organ donors although they can consciously decide to "opt out". It has been shown that this method leads to a considerable increase in the number of donations. Using the language of System 1 and System 2 (Sloman, 1996; Kahneman, 2011), the citizen automatically opts for a default state using the automatic and intuitive System 1, but he can use the analytical and reflective System 2 to evaluate and decide whether to accept this state. It becomes hard to sustain the libertarian attribute when the nudge intervention is aimed primarily at exploiting our System 1 and the deliberative "opt out" clause becomes secondary and unpleasant. If, as Sen (2009) and Bobbio (2009) assert, no moral end or question of justice can be founded on Machiavellian grounds, namely using a process that contradicts the principle itself,

---

behaviour. Niccolò Macchiavelli also drew attention to the decision-making power of both the "endowment effect" and "loss aversion" (1532). Lastly, David Hume (1739) underlined the danger of "present bias" and the myopic nature of human judgement.

then it is legitimate to ask how much liberty of choice is protected by an approach based on non-deliberative automated processes. The possibility of deliberative choice exists in abstract terms, but the choice architecture is biased towards the maintenance of the default state due to the prevalence of our System 1 over System 2.[2] Paradoxically, the libertarian attribute would be more justified if, as other authors like Hammond (1996), Dennett (1987; 1991), Frankish (2004; 2009), (Frankish and Evans eds., 2009), (Kruglanski and Gigerenzer, 2007) affirm using cognitive continuum theory, the mind cannot be separated into two systems, but rather each decision presents mixed characteristics of an analytical and intuitive type and therefore each decision can be deemed deliberative, to a varying degree, and as such is relatively conscious and free.

At any rate the attribution of libertarian to paternalism alters depending on what stage of the decision-making process it is applied to. Clearly, together the terms form an oxymoron. In a paternalistic dimension, there is always a varying degree of heteronomy that reduces the space for autonomous decision-making. However, this reduction of autonomy differs depending on the type of paternalism corresponding to different stages of the decision-making process. We could identify three key stages by working backwards, from the downstream end back upstream.

At the furthest point downstream we have the choice architecture of a state of wellbeing. This set of interventions could be called *Hedonic Paternalism*. Nudge theory, generally, constructs choice as a default state. As we saw earlier, various criticisms can be put forward to contest the non-libertarian dimension of this paternalism, or in other words the tendency of *nudgers* to turn into preachers and technocrats of reason.

Conversely, and further upstream, there are nudges designed to reinforce the capacity for reasoning and judgement, thereby leading to the choice of what solution to adopt. In this case, we can talk of *Cognitive Paternalism*. This paternalism is completely libertarian given that it strengthens the deliberative capacity of the individual. For example, faced with a complex problem, a simplification is proposed by highlighting only the relevant variables and drawing attention to the underlying structure. Support for using a combination of *lexicographical heuristics* and *satisficing* is also a highly effective way of reducing the computational burden when choosing between alternatives. Faced with the medium and long-term effects of one's own choices, the proposal to highlight and simulate the relationship between an individual's own choices and their effects on his wellbeing over the medium term is also another form of paternalism that is acceptable from a libertarian point of view. For example, asking anyone taking out a loan to try and anticipate its effects on their objective financial solvency when that person finds themselves having to pay increasing monthly instalments for the coming years is certainly an acceptable paternalism in terms of deliberative freedom. So are proposals that increase environmental feedback for personal choices or that help by providing appropriate "warnings" to neutralise misunderstandings or errors in computational calculation. This type of paternalism must, however, also include the advice of bounded and adaptive rationality. In many contexts, especially where there is uncertainty, it is better that *nudges* do not target an increase in pointless (in the sense of hopeless) computational skills when dealing with forecasting problems where the alternatives are not known and there is no possibility of estimating the relative probabilities. Better in this case to encourage decisions based on simple, fast heuristics, some of which, but not all, are based on the intuitive component of the human mind (Gigerenzer, Todd et al., 1999; Gigerenzer, 2015).

[2] According to authors like Sudgen (2009) it is not possible for a deliberation to be free of reflex automatisms or to be a pure expression of System 2.

Lastly, upstream we have the most important paternalistic intervention that a government can make to improve the decision-making processes of its citizens: we can call this *Educational Paternalism*, namely the attempt to give anyone who wants to catch fish a good fishing rod. Behavioural economists do not believe much in economic education. They use a significant analogy that likens *bias* to optical illusion. In the same way that you can't see certain shapes in experiments using the ambiguous shapes of gestalt psychology, so it's not possible to avoid making certain mistakes in probabilistic and logical judgement. The use of education to strengthen metacognitive and "de-biasing" skills is, to them, a hopeless task. As Richard Thaler affirmed in *Nature*: "Our ability to de-bias people is quite limited" (Bond, 2009, p.1191). Hence the need for *hedonic* and *cognitive-paternalism*. Now, the problem of training in statistical thinking and logical reasoning, both of which are essential for economic and social decisions, has been examined in some past and recent studies (Fong, Krantz, and Nisbett, 1986; Nisbett, 2009; Gigerenzer 2014). These have ascertained whether it is possible to teach the statistics of natural frequencies, like the law of large numbers, and whether it is possible to reduce the propensity for confirmation bias in real-life type settings. This type of training is very useful in risk situations. In those dominated by uncertainty, where the principle of optimality does not hold, it is better to teach some simple fast heuristics and to pay attention to how the problems are framed.

2) Architecture of Mind and the Impossibility of Hedonic Libertarian Paternalism

The Nudge theory has two features. It is paternalist because it nudges people to make decisions that are in their own interest. It is libertarian because people are able to ignore the nudge and decide differently. In other words, the theory is paternalist because it uses the System 1 of cognition to drive the individual automatically towards a choice that is good for him or her, and it is libertarian because it relies on the System 2 of cognition to consciously analyse and evaluate the choice with the possibility of opting out. The possibility of making a conscious analysis of the nudged choice (above all if there is a default state) and, if it is not acceptable, to reject it and to make a difference choice underlies the libertarian justification of the choice architecture. The individual is not hemmed in by the specific choice architecture and obliged to choose what the behavioural technocrats want, but rather he has substantial not just formal freedom to reject the nudges. This is a fundamental political and philosophical point because if this were not the case then there would be no basis or justification for the libertarian attribute. Instead the situation would be one of authoritarian and non-liberal manipulation of citizens' choices using behavioural techniques. At any rate, even if this possibility of conscious choice were demonstrated *a posteriori*, Sen's anti-Machiavellian criticism (2009) remains, namely that it is not possible to achieve liberal goals of conscious individual wellbeing using means that are based on the opposite principles, whether paternalistic or technocratic, because they use unconscious mechanisms of choice.

Let us take a closer look at whether Thaler and Sunstein's libertarian claims are backed up by today's knowledge of architecture of the mind. As we saw earlier, the claim of libertarian paternalism will be analysed first in its "hedonic" variant, rather than in the "cognitive" one. Indeed, if it can be shown that the default states created in the choice architecture are not easily modified and, on the contrary, that the aim of this architecture is that they cannot be modified, then any libertarian claim would collapse. By way of example, think of the default states of welfare programmes, like "Save more tomorrow", or of the opting-out version of organ donation where . If in these cases, situations of choice are produced that are difficult to reverse owing to the "status quo

bias" and the effects of inertia and weakness of will, on the basis of which the Nudge was justified and constructed, then this cannot be justified as a libertarian intervention. The same can be said of other Nudges, like placing products with higher fat and calorie content on the top shelves in supermarkets, so that by being less visible and therefore less attractive to consumers they are set aside in favour of products with a lower fat and calorie content. Here too, it is difficult to argue the possible intervention of deliberative processes that can reverse a perceptive and cognitive-type propensity that focuses attention and analysis on the most visible part of the shelves.

In the first place, we should ask ourselves what characteristics of the mind are foreseen, either implicitly or explicitly, by Nudge theory in order to justify the attribute of libertarian, and how are they backed up by recent advances in our knowledge of the architecture of the mind.

a) In the first place, the phenomenon of biases cannot be likened to the optical illusions studied by gestalt psychology which are incorrigible and impenetrable. Instead, it should be possible to alter the outcome of the cognitive process responsible for the bias in order to take a normatively correct conscious decision. This cannot happen if, once the bias has been triggered by the nudge, it is no longer reversible using deliberative-type processes. Thaler and Sunstein's position appears to favour an analogy between bias and optical illusion when it introduces Part 1 "Humans and Econs" with Shepard's example of two tables (Thaler and Sunstein, 2008, pp. 17-19) that seem to have a different ratio of length to width. If this were really the case, there would be very weak justification for the libertarian attribute based on the repairing intervention of the analytic and conscious part of our mind. The Fodorian modular impenetrability of biases, like optical illusions, would not allow any correction. Human decision-making would therefore be trapped without escape in the bias induced by the nudge. This prevalence of manipulation to which an individual is subject would make the paternalism seem authoritarian not libertarian. As we saw earlier in Thaler's declaration in *Nature*, this seems to be the contradictory attitude underlying Nudge theory. On the contrary in my opinion references cited earlier appear to show that biases can be corrected and that they can also be prevented through training. This is the role of cognitive paternalism and of educational-paternalism. Moreover, as we will also see later, many of these biases have been detected in artificial linguistic and pragmatic conditions that do not match the real-life contexts of choice.

b) If we allow that, unlike optical illusions, biases are corrigible then we must assume that, in order to satisfy the libertarian attribute of the Nudge Theory, the mind must be structured into two systems or process types, of which one is automatic and associative and responsible for the biases, and the other is analytic and conscious and able to correct them. This is the position taken by Thaler and Sunstein (2008, pp. 19-22) when they state that our brains are made up of two systems, one automatic (System 1) and the other reflective (System 2). There are numerous differing positions on dual-process or dual-system theories and on their relative properties (for a complete list of theories see Stanovich, 1999; for a complete list of properties see Evans, 2008). This is not the place to analyse them. Instead, I would like to focus on just two hypotheses which might question the role of System 2 as an analytical corrector of bias, thereby weakening the libertarian thesis. In the first place, there is a growing programme of research that denies duality of mind and therefore puts forward a single-process account of dual-process phenomena. The two main anti-dualistic approaches are that of the "cognitive continuum" put forward by Hammond (1996) and Cleerman and Jimenez (2002; Osman, 2004), and the "rule-based processing unified theory of decision-making" argued by Kruglanski and Gigerenzer (2011). According to the dynamic graded continuum (DGC) different types of reasoning are dependent on the representations from which participants reason (Osman, p. 1002). The same underlying

production rule can generate a variety of responses as a result of the features of the task that an individual considers relevant (Osman, p. 1003). The properties of reasoning (control, awareness, speed, etc.) vary in degree depending on the structural features of the tasks that induce different cognitive activity (Hammond, 1996). The cognitive continuum ranges from intuition to analysis. The more well-structured a task is, the more analytically induced will be the decision-making mode. Conversely, with an ill-structured task decision making is likely to be intuition-induced (Hammond, 1996). These theories seem to fit better than dual-process theories to a set of criteria like the Criterion S, the individual differences in cognitive ability, the dissociation between implicit and explicit processing, as analysed in a series of tasks as the selection task, the conjunction problem, and belief bias in syllogistic reasoning (Osman, 2004, pp. 996-1005). For our purposes here, namely to examine the libertarian attribute of Nudge theory, what matters is whether in this architecture of mind, there can be correction after the bias has been generated. From what can be deduced from some of the key monist hypotheses, the type of processing appears to be triggered by the structure of the problem. If a problem is opaque and unnatural in its logical structure or if it is linguistically and pragmatically confused and deviant or if it is unfamiliar, namely the quality of representations is poor, this will stimulate the intuitive, fast and implicit part of the mind responsible for biases. In this case, contrary to what is proposed by dual-process theories, there is no possibility of correction by the analytical component which is only stimulated by well-structured problems and the relative quality of representations.

In principle, the situation is different in dual-process or system theories. The question is which of the 23 different theories identified by Stanovich can be deemed acceptable (1999; 2004)? An initial selection can be made by considering the neuroscientific data that seem to show that there is no single System 1, with a single system of attributes, but rather multiple cognitive and neural systems (Evans and Stanovich, 2013, pp. 224-6). Therefore, it is preferable to use the terminology of dual types of processing. Furthermore, a minimal part of the paired properties usually attributed to the two processes are defining (Table 1). The majority are correlates that might or might not be present. In addition, there are at least two sets of theories: the *parallel-competitive* theories (Sloman, 2000) assume that Type 1 and 2 processing proceed in parallel, each having their say with conflict resolved if necessary.

Table 1. Clusters of Attributes Frequently Associated With Dual-Process and Dual-System Theories of Higher Cognition

Type 1 process (intuitive)	Type 2 process (reflective)
Defining features	
Does not require working memory	*Requires working memory*
Autonomous	*Cognitive decoupling: mental simulation*
Typical correlates	
Fast	Slow
High capacity	Capacity limited
Parallel	Serial
Nonconscious	Conscious
Biased responses	Normative responses
Contextualized	Abstract
Automatic	Controlled
Associative	Rule-based
Experience-based decision making	Consequential decision making
Independent of cognitive ability	Correlated with cognitive ability
System 1 (old mind)	System 2 (new mind)
Evolved early	Evolved late
Similar to animal cognition	Distinctively human
Implicit knowledge	Explicit knowledge
Basic emotions	Complex emotions

Note. Italicized attributes are the proposed defining characteristics in the current article. Authors proposing two systems include the features attributed to Type 1 and 2 processing but may also include the additional features named.

discontinued and discouraged the use of the labels System 1 and 2 (e.g., Evans, 2010a; Stanovich, 2004, 2011).

Both Evans (2008, 2010a) and Stanovich (2004, 2011) have discussed how terms such as *System 1* or *heuristic* *system* are really misnomers because t what is being referred to is a singular syst the term *System 1* should be plural becau set of systems in the brain. Stanovich (example, noted the wide diversity

Table 2. A Glossary of Dual-Process Terminologies Used in This Article

Term	Definition
Dual processes	The assumption by many theorists that cognitive tasks evoke two forms of processing that contribute to observed behavior. Unless otherwise indicated, the term refers in this article to dual-type theories.
Dual types	Terminology that implies that the dual processes are qualitatively distinct. Type 1 processes are (broadly) intuitive and Type 2 processes reflective (see Table 1).
Dual systems	It is common in the literature to use the terms *System 1* and *System 2* to refer to the Type 1 and 2 distinction. Some but not all authors associate these with an evolutionary distinction. The current authors now prefer to avoid this terminology as it suggests (falsely) that the two types of processes are located in just two specific cognitive or neurological systems.
Modes of processing	Modes of processing are forms of Type 2 thinking that may differ on a continuum. Individual differences on such continua are often assessed with thinking-disposition measures.
The autonomous set of systems (TASS)	The proposal that there are multiple Type 1 systems of different kinds, including modular, habitual, and automated forms of

Criterion S refers to situations where individuals are led to respond in a manner consistent with Type 1 but then come to realize an alternative responding consistent with Type 2. In contrast there are the *default-interventionist* theories (Kahneman and Federick, 2002; Kahneman, 2011; Evans and Stanovich, 2013; Evans, 2008). They assume that fast Type 1 processing generates intuitive default responses on which subsequent reflective Type 2 processing may or may not intervene.

One difficulty with parallel-competitive forms of dual-process theories is that Type 1 processing is much quicker than Type 2 processing. As Evans and Stanovich write "…the fast horse must wait for the slow horse to arrive before any potential conflict can be resolved" (2013, p.237). In these cases it is likely that the analytic correction arrives often out of time. ".., the associative system always has its opinion heard and, because of its speed and efficiency, often precedes and neutralizes the rule-based response" (Sloman, 2002, p.391). If this is the case the Type 1 processing induced by Nudge will be dominant and the justification of libertarian paternalism would become very weak.

c) According the default-interventionist theory the analytical intervention is subsequent to the Type 1 processing. In order to justify the libertarian requisite the Type 2 corrective processes must be impermeable to the biasing influence of Type 1 process. According Evans (2009) the Type 1, also dubbed as *autonomous* processes, are not only those that control the behaviour directly, without need for any kind of controlled attention. It includes also the *preattentive* processes that are those that supply content into working memory. The content includes perceptual processing and retrieving stored memories and beliefs (fig. 1).

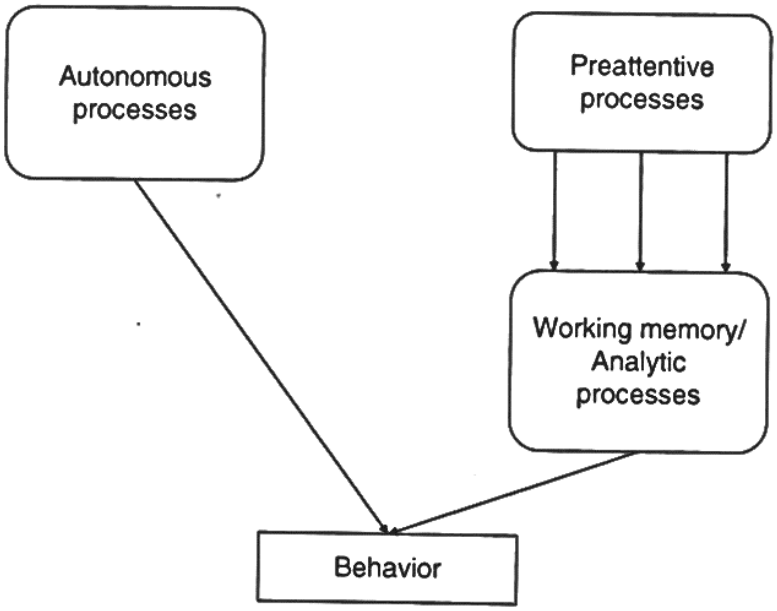


Fig. 2.1 A simple representation of the difference between autonomous and preattentive type 1 processes.

Therefore the Type 2 analytic processes manipulate through working memory explicit representations supplied by Type 1 preattentive processes and exert conscious volitional control on behaviour. The preattentive processes are described as attentional, linguistic and pragmatic. They are distinct from autonomous processes that recruit implicit knowledge of an associative nature. Instead they rather concern the implicit processing of explicit knowledge. Their pragmatic function is manifest in contextualizing our thought, retrieving stored memories and beliefs that are relevant to the current context. These processes can be the cause of cognitive biases if they fail to encode logically relevant information, or encodes irrelevant information. Stanovich (1999) terms it the *fundamental computation bias,* that is the tendency to automatically contextualize problems. The bias prevents the individuals from reasoning about a task according to its logical properties. They rely on cues from its context which are interpreted in relation to real-life situations (Evans and Stanovich, 2013). The subsequent analytic processing may cause biased responding (Evans, 2009). Since the contextualization has biasing effects and the analytical processes rely on the preattentive processes they are not impermeable to Type 1 implicit biasing counter normative influence. The "framing effect" is one of the best-known examples. This phenomenon is particularly relevant in the choice architecture of Nudge theory. The goal of nudgers is to shape environment of choice that frames in the desired way the representation of the choice and the subsequent response of the citizens. The frame affects not only the Type 1 autonomous processes but also the preattentive processes that are the necessary

premise of the analytical ones. Therefore the analytical processes are born biased by the nudge and can't express their corrective power.
Moreover, there are many examples of the permeability of Type 2 processes from Type 1 autonomous processes. For example in social psychology implicit association as racial prejudice may influence the explicit association based on propositional evaluation against the racial prejudice as that shown in self-report (Gawronski and Bodenhausen, 2006; Smith and Collins, 2009). When the associative system works unnoticed by the individual, the information arising from that system may exert an unwanted biasing influence on the Type 2 analytical processing (Smith and Collins, 2009). If people respond according an implicit attitude or implicit stereotype they must generally lack awareness that these behaviours are in conflict with their explicit attitudes and beliefs (Evans, 2009). Moreover, when the impulsive system's behavioural motivation is strong, individuals may act in ways that contradict their stated values or long term rational goals. For example cheating on diet or on gambling. The permeability of Type 2 analytical processes is proved also by Embodied Cognition perspective (Semin and Smith, 2008). Behaviour often shapes cognition. Over time physical actions can become associated with specific states or stimuli, eventually causing the cognitions to be automatically triggered by the physical action, even when inappropriate (Smith and Collins, 2009). All these data seem to prove that the corrective action of analytical processes to neutralize the biased effect of nudge is not justified. The Type 2 analytical processes are permeable to the biasing effect of Type 1 autonomous processes and in situation of framing effect to the biasing contextualization of the Type 1 preattentive processes.

d) If the correction of Type 2 analytical processes impermeable to the biased influences of Type 1 autonomous processes was possible in theory, the question that would then arise would be whether this corrective activity is effective at a temporal level. In other words, there should not be too great an asymmetry between the speed of automatic biased decision-making and the slowness of correction through analytic processing. Indeed, if analytic processing is very slow it runs the risk of not being effective because it might become more difficult to make the choice reversible. If the individual, after making the biased choice generated by the autonomous processes, had already activated the status quo bias in relation to the choice, weakness of will and inertia would not allow the analytical process of correction to be activated. On the contrary, if analytic processing might be activated very rapidly, then the problem would be whether the individual has sufficient motivation to put it into practice. The environment of choice that led to the creation of the bias tends to entrap the individual in a situation that does not motivate the use of analytic processing. People perform analytic processing when they have the motivation, time, and cognitive capacity allowing for more effortful processing. This involves the active, effortful scrutiny of all relevant information, requiring cognitive capacity (Smith and Collins, 2009). When people are low in capacity and motivation they will not engage in much elaboration so judgements will be based mostly on salient peripheral cues and not on the central structure of the argument (Petty and Cacioppo, 1986). Peripheral cues are for example, the attractiveness of the message source or the situation linked to the message.

e) Lastly, if analytic processing is activated in good time to correct the results of autonomous processes, in order for this to justify attributing the libertarian epithet to Thaler and Sunstein's hedonic paternalism its role as a conscious and rational correction should be certain. In other words, it should not be in turn a source of unintended biases and errors. On the contrary biases don't arise solely from the use of Type 1 autonomous processes. Rule-based and analytic

processing can also be biased by the perceiver's motives, by priming, by mood and other factors that make certain cognitive structures less accessible. For example research in social psychology demonstrates that intentional efforts to correct bias may even lead to further bias. Wegener and Petty (1997) describe how people use intuitive theories of bias in an effort to correct judgements that they believe might be biased. Of course to the extent that such intuitive theories of bias are inaccurate such efforts of correction might fail to remove bias in judgement, or even exacerbate them (Smith and Collins, 2009, p. 207). One example comes from the effect of disclosure by the financial advisers (Sah, Cain, and Lowenstein, 2013). In order to accurately discount the biasing influences behind advice the advisee generates a theory on adviser behaviour that predicts both the impact of the conflict of interest on the advice and, the impact of disclosure on that advice. Without a correct advice-discounting theory advisees have difficulty using disclosure effectively: they may ignore it, they may discount it but insufficiently so, or they may discount the advice too much. In some cases, disclosure might actually increase trust in the advice in an irrational way. In fact Type 2 processing can lead often to errors. Evans (2007) and Stanovich (2009) give a wide overview of cognitive biases caused by Type 2 processing and of the shallow, error-prone Type 2 thinking (fig. e and table 2). Since humans are cognitive misers they follow the rule to engage the brain only when all else fails and usually not even then (Krueger and Funder, 2004). Cognitive misers have 3 rules: default to Type1 processes whenever possible; when that is not possible and analytic processing is necessary default to serial associative cognition with focal bias (as in the case of matching bias in selection task); when the reflective mind wants to start cognitive simulation by decoupling, not complete it, that is override failure (Stanovich, 2009, p. 69). Moreover if the analytic processes wants to override the Autonomous processes it is possible they don't succeed because the mindware is not available or it is contaminated. What is the mindware? It includes rules, procedures, and strategies that can be retrieved by the type 2 analytic Processes and used to transform decoupled representations to override the Autonomous mind. It is mainly the product of past learning experiences. If a rule is not learned or is not well learned or is not appropriately applied is the cause of an override failure. For example the knowledge of the gambler fallacy for a pathological gambler (Stanovich, 2009, p. 73). But not all mindware is helpful for overriding autonomous processes. Some acquired mindware can be the direct cause of errors and biases. For example the egocentric thinking, the myside perspective, the evaluation-disabling memes (the memes against critical thinking and for consensual dogmatic faith) or false lay psychological theories (e.g. the personal immunity to bias or the personal knowledge of the causes of own actions, etc..)

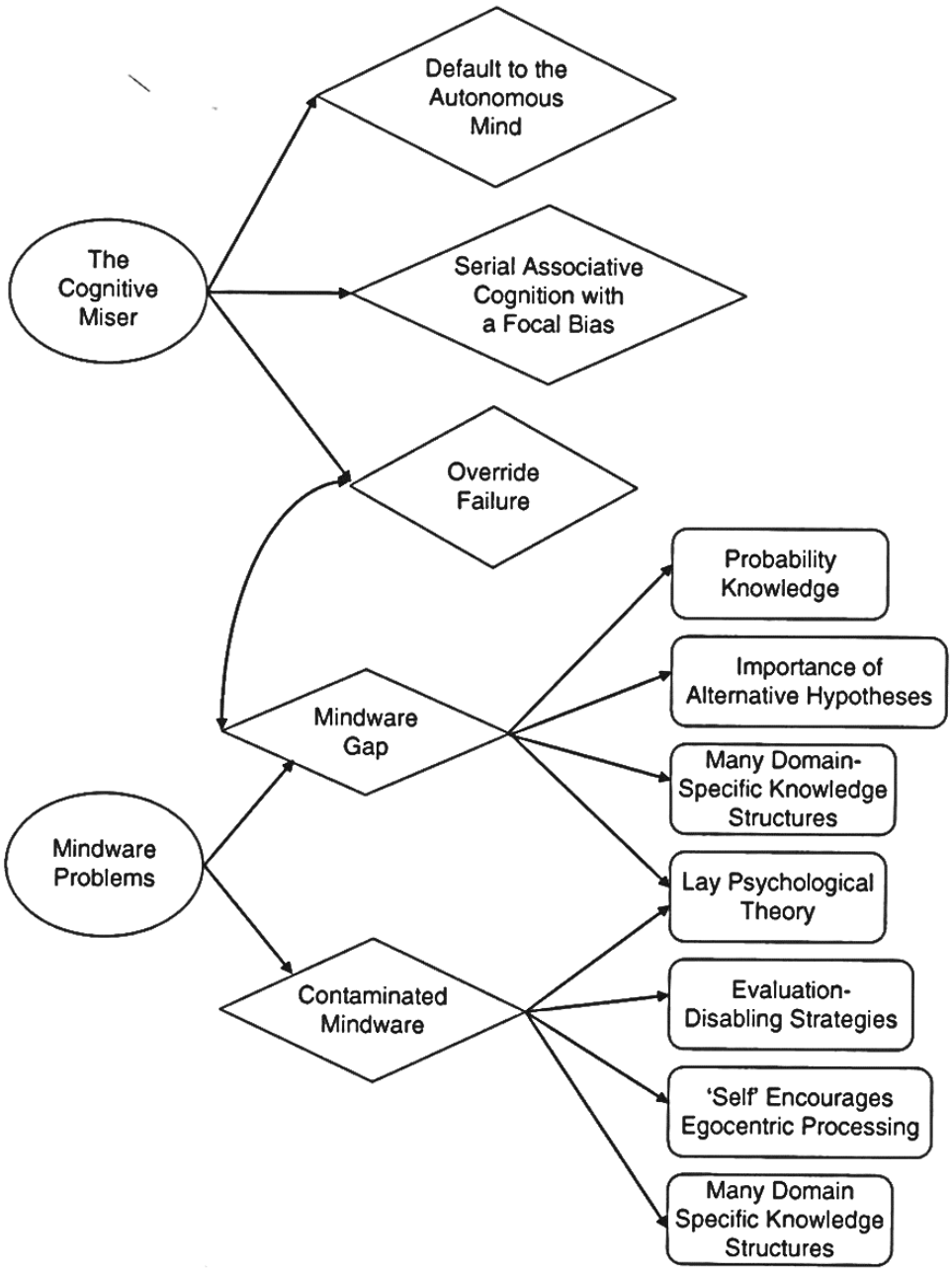


Fig. 3.4 A basic taxonomy of thinking errors.

Table 3.1 A basic taxonomy of rational thinking errors

	The Cognitive Miser			Mindware Gaps (MG)		MG & CM	Contaminated Mindware (CM)	
Tasks, Effects, and Processing Styles	**Default to the Autonomous Mind**	**Focal Bias**	**Override Failure**	**Probability Knowledge**	**Alternative Thinking**	**Lay Psychological Theory**	**Evaluation Disabling Strategies**	**Self and Egocentric Processing**
Vividness effects	X							
Affect substitution	X							
Impulsively associative thinking	X							
Framing effects		X						
Anchoring effects		X						
Belief bias			X					
Denominator neglect			X					
Outcome bias			X					
Hindsight bias ('curse of knowledge' effects)			X					
Self-control problems			X					
Noncausal baserates				X				
Gambler's fallacy				X				
Bias blind spot						X		
Causal baserates			X	X				
Conjunction errors	X			X				
Ignoring P(D/~H)		X			X			
Four card selection task	X	X			X			
Myside processing		X						X
Affective forecasting errors		X				X		
Confirmation bias		X			X		X	
Overconfidence effects		X						X
Probability matching		X		X				
Pseudoscientific beliefs					X		X	
Evaluability effects		X				X		

Lieberman (2009) gives an explanation based on neural imaging. The reflective system (Type 2 processing), mainly the temporal lobe incorporating hippocampus, can have a disturbing effect on the activity of the reflexive system (Type 1 processing), mainly basal ganglia, when people learn complex task or make judgements.

To conclude a lot of empirical evidence seems to prove that most of the people are not able to deliberate the correction and overriding of biases caused by the dominance of the autonomous Type 1 processes. There are different styles of thinking or in other worlds different modes of thinking (corresponding to the dimension of the *reflective mind* according Stanovich, 2009). These remarks are also evident in the "geography of thinking" that highlights cognitive differences in adults, but not in children, between the analytical American and holistic Far Asian modes of thinking (Nisbett, 2003; Nisbett, and Masuda, 2006; Viale and Osherson, 2000; 2006; Norenzayan, 2006; Viale, Andler and Hirchfeld, eds., 2006). Few people have the *need for cognition*, that is a cognitive mode and a style of thinking, reflecting the extent to which individuals are inclined towards effortful cognitive activities and high elaboration. Few people need to structure relevant situations in meaningful, integrated ways and need to understand and make reasonable the experiential world. But even among these individuals the Type 2 analytic processes (that is the reflective and algorithmic mind according Stanovich, (2009) are not the guarantee of correction, because in many cases there is overriding failure.

In conclusion since the possibility of the deliberate analytic correction of the induced biases is neither in theory universal the libertarian label to hedonic paternalism is not ethically justified.

Whereas there is evidence that Type 2 processes are responsible of biases there is also growing evidence that Type 1 processes can lead to superior decisions and that conscious reflection can sometimes impair the quality of judgements. Recent years have seen an increased emphasis on the idea that more affective, intuitive judgements are often useful and accurate rather than being condemned because they are not narrowly "rational". Work on fast, effective judgements by domain experts (Klein, 1999); Wilson and Schooler's work (1991) showing that intuitive attitude judgements can be more adaptive for the perceiver than more deliberate attitude judgements; and in particular the work on fast and frugal heuristics of Gigerenzer and ABC group (Gigerenzer, Todd, and ABC group, 1999; Gigerenzer and Selten eds., 2001) are completely changing the landscape of the studies on decision making and rationality.

3) Instrumentalist Features of BEN

Nudge theory is based on behavioural economics. Behavioral economists who decades ago defined their critical contribution to the neoclassical mainstream a purely descriptive enterprise (Thaler, 1991) now advocate using behavioural concepts for prescriptive policy purposes (Thaler and Sunstein, 2008). This evolution is not justified because a) the descriptive behavioural enterprise seems not to fulfil the realist desiderata of a true empirical endeavour to substitute the as-if approach of neoclassical economics; b) the prescriptive behavioural enterprise is biased by the conventionalist nature of the descriptive side and often it is not capable of truly offer prescriptions that increase the wellbeing of the citizens.

When Herbert Simon began his attempt to change empirically the economics his methodological and epistemological coordinates were realist (Simon, Egidi, Viale, and Marris, 1992; Simon, 2009). His main critical target was the instrumentalist as-if approach of Milton Friedman (1953). A descriptive enterprise in economics had to overcome the unbounded rationality assumptions of neoclassical economics as unbounded self-interest, unbounded willpower and unbounded computational capacity. The behavioral economics programme initiated by Simon had the goal of replacing these a-priori assumptions with more realistic ones. How much psychological realism has been brought into economics by behavioural economists? Unfortunately very little because there are barriers to psychological realism that are common to neoclassical economics and that are the son of the shared reliance on Friedman's as-if principle (Berg and Gigerenzer, 2010). All relevant behavioural theories suffer of the same shortcomings of neoclassical economics: assuming that risky choice always emerge from a process of weighting and averaging of all relevant pieces of information; the decision maker knows the objectively feasible action set; the decision maker know the list of outcomes associates with lotteries or the probabilities of the known outcomes (Berg and Gigerenzer, 2010). The shift from neoclassical economics to behavioural economics and in particular, after the impact of Allais Paradox, from *expected utility theory* to *prospect theory*, appeared to be based on the introduction of more transformations with additional parameters to square the basic operation of probability-weighted averaging with observed choices over lotteries (Berg and Gigerenzer, 2010). Weighting-and-adding objective function is used as-if it is a model of mind. But it's not. It is a fictional mind, a valid instrument to make a posteriori inferences through the introduction of suitable parameters in order to reach a better R-squared

The same methodological model is observed in many other behavioural theories (Berg and Gigerenzer, 2010). For example the Fehr and Schmidt *social preference model* (1999) recognize the insight that people care about others' payoffs. Therefore they modify the utility function with addition of at least two additional free parameters. People are assumed not to maximize a utility function depending only on their own payoffs but a behavioural or other-gathering utility function. To do it decision maker assigns benefits and costs to each element of the choice space based on weighted sum of the intrinsic benefits of own payoffs together with the psychic benefits of being ahead of others and psychic costs of falling behind others. The decision maker will select the action with the largest utility score based on weighted summation. Another as-if model is the Laibson's (1997) *model of impulsiveness in consumption*, a psychological bias that over-weight the present over the future. He puts more weight on the present by reducing weight on all future acts of consumption. In other words he reduces the weight of all terms in the weighted sum of utilities except for the term representing utility of current consumption. The unrealistic pretension is evident: the decision maker after an exhaustive search of all possible acts of consumption compute the weighted sum of utility terms for each act and choose the one with highest weighted utility score. The deviation between the value that recovers the neoclassical version and the new parameter that reduces the weight on the future is considered empirical confirmation of the model.

The instrumentalist methodology of behavioural economics uses the addition and managing of free parameters to improve the realism of the models. In so doing it improves the within-sample fit and improves the R-squared. Most of the philosophers of science both in the realist tradition (e.g. Hacking, 1983) and in the antirealist tradition (e.g. van Fraassen, 1980) agrees on the empirical adequacy by successful prediction, particularly of novel facts, as the first principle in deciding between competing hypotheses (Viale, 2013). A large number of

free parameters allow the model to fit many sets of data without proving to generate successful out-of-sample prediction. On the contrary the most challenging test of a theory is in prediction using a single set of fixed parameters. Something that few models of behavioural economics dare to do.

4) Rationality in a "Large World"

The real life problems are inside a complex environment. They are typically ill-defined problems; that is, the goals are not definite; we don't know what counts as an alternative and how many alternatives there are; it's unclear what the consequences might be and how to estimate their probabilities and utilities. This environment might be called also as Large World (Savage, 1954) and it is characterized by uncertainty. Small Worlds instead are in principle predictable and without surprises and they are characterized by the knowledge of all relevant variables, their consequences and probabilities. The conditions of small world are the requirements of Neoclassical Rationality as Simon stressed in his Noble Lecture (1979, p. 500). In these worlds the problems may be well-defined but they can be also computationally intractable. As it is well known an example of a computational tractable problem is the dice game or the roulette game. Instead well-defined problem as chess game is computationally intractable. In any case the real world is most of the time large and these conditions of knowledge are rarely met. Since they are rarely met the normative rational requirements of neoclassical economics are unjustified and the application of their theories can lead easily to a disaster (Stiglitz, 2010). Unfortunately behavioural economics whereas it criticizes the descriptive side of neoclassical economics, without really proposing an alternative realist model of decision making, it retains the normative one. In fact the heuristic and biases programme is developed to cope with what it is called human irrational behaviour, characterized by biases and formal errors caused by psychological mechanisms as the heuristics. Thaler (1991, p.138) writes very clearly about:

A demonstration that human choices often violate the axioms of rationality does not necessarily imply any criticism of the axioms of rational choice as a normative idea. Rather, the research is imply intended to show that for descriptive purposes, alternative models are sometimes necessary.

In a large world the axioms of rationality can't be applied. Therefore they can't be considered feasible normative canons of rationality. Moreover as Jonathan Cohen (1981) properly writes in his seminal article:

However, nothing in the existing literature on cognitive reasoning, or in any possible future results of human experimental enquiry could have bleak implications for human rationality, in the sense of implications that establish a faulty competence. (p.152)

Consequently the label of irrational behaviour attributed to biases and errors is not justified. In a large world the rationality must be judged in relation to the proper adaptation of the choices and problem solutions to a given environment. Rationality can't be formal but only *Ecological.* Counterintuitively the formal rationality should be considered pathological because it doesn't supply canons or reasoning to properly adapt the behaviour to the

environment. On the contrary libertarian paternalists believe that people suffer from systematic reasoning errors due to their cognitive limitation in achieving the normative standard of formal rationality. And they claim that these errors imply serious costs for human wellbeing. Actually there are no data that prove that this is the case. For example a systematic review of hundreds of framing studies could not find a single one showing that framing effects incur real costs in terms of health or wealth (Arkes, Gigerenzer, and Hertwig 2015).

There is also another reason that weakens the libertarian paternalist claim. The norms that are evaluated in the experiments "are syntactical, unconditional to the semantics (the content) and the pragmatics (the intentions)" (Engel and Gigerenzer, 2006, p.9). In other words some errors and biases are discovered in artificial experimental tests that are far from the pragmatic dimension of the decision making in everyday life and in some cases presents noise effects from the point of view of the pragmatics of the language and signalling theory. There are many examples of this kind of bias. Here I focus briefly on *confirmation bias* and *framing effect*.

In the first, as empirically verified using Wason's famous 4-card task (1966), there is a systematic tendency for individuals to find cases that confirm their own hypotheses rather than checking them rigorously. The original test was based on letters and numbers instead of using arguments taken from real life: the rule to be verified was "If there is an A on one side of the card, then there must be a 2 on the other side".[3] It was therefore completely abstract. Indeed the vast majority tended not to check it but to confirm it.[4] However, when realistic contents were inserted instead of letters, and in particular if permission rules were used, such as "If someone is drinking alcohol, then they must be at least 21 years old" (Griggs and Cox, 1982) or "If a letter is sealed then it has a 50-lire stamp on it" (Johnson Laird, Legrenzi and Legrenzi, 1972), then the *confirmation bias* was reduced and people were able to choose the cards that checked the truth of the rule.

One of the best known phenomena of behavioural economics is the *framing effect*. This effect undermines one of the cornerstones of economic rationality, namely the "invariance principle" (Tversky and Kahneman, 1986, p. 253). This phenomenon is one of main justifications of *libertarian paternalism*: given that it is impossible to avoid it and given that individuals' choices and their preferences are conditioned by how problems are framed, it is necessary to decide for them. This effect is based on a principle that if applied at a formal level cannot grasp the pragmatic dimension and the *conversational implicatures* (Grice, 1989) underpinning some of the questions through which it was discovered. For example, as was shown by Gigerenzer (2015), in the case of a person affected by serious heart disease who has to decide whether to undergo high-risk surgery, his reaction to the advice given by his doctor varies if the same alternative is expressed using two different frames:

- Five years after surgery, 90% of patients are alive
- Five years after surgery, 10% of patients are dead.

Faced with the same semantic contents, he will make a pragmatic inference that will prompt him to interpret the first frame in favour of surgery, and the second against it. In other words, it will activate a *conversational implicature* that will lead him to infer that if the doctor makes the first assertion, he is recommending surgery, while if he states the second, he is advising against it. The same is true of

[3] Individuals were presented with 4 cards which were marked with an A, a D, a 2 and a 7 respectively on the visible side. On the hidden side, there might be letters or numbers.

[4] Most subjects almost never chose the card with a 7 on it which, if it had an A on the other side, would have falsified the rule. Many correctly chose the card marked with an A which, if it had a 7 on the other side might have disproved the rule. In general, however, this choice was guided by the expectation that the hidden side would have a 2 on it.

the famous "Asian Disease Problem". Faced with the two options of the gain-framed message, for example:

- If Programme A is adopted, 200 persons will be saved
- If Programme B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved

the choice will tend towards Programme A (low-risk option) capable of saving 200 people (but there is no mention of the fact that 400 will die). An analysis that refers to the ecological rationality tells us instead that people refer to the frame characteristics to obtain relevant additional information that help to recommend the response to give. Namely the fact of not mentioning that 400 people will also die prompts individuals to view the option as pragmatically preferable to the other (high-risk option) – a one-third probability that 600 people will be saved and a two-thirds probability that no one will be saved – even though the two options are formally the same.

The previous analysis of the framing effects is in terms of communicating reference points and signalling recommendations. What it is important to realize is that descriptive invariance is not a reasonable yardstick for rational behaviour.

Finally Thaler and Sunstein (2008) argued that people "fail to make forecasts that are consistent with Bayes' rule". On the contrary there are many data that show that people learn probability from experience and they make judgements consistent with Bayes' rule. There is a Bayesian model of cognition (Griffith, Kemp, and Tenenbaum, 2008) that shows how perceptual processes, language understanding, and categorization are consistent with Bayesian model. And there are many cognitive scientists (Chater and Oaksford, 2008) that claim that people's judgements are consistent with Bayes'rule. Why do the experiments to whom Thaler and Sunstein are referring seem to prove the contrary? Because as Gigerenzer highlights (2015) the experiments, for example the "engineers and lawyer problem" and the "cab problem", are not properly designed. In the first case the random sampling was not properly communicated. When random sampling was experienced by the people that could randomly draw description out of a urn, their neglect of base rate disappeared (Baratgin and Noveck, 2000). In the second case if information is presented as the outcome of learning from experience, known as natural frequencies, and not as conditional probabilities, the proportion of people reasoning by Bayes'rule increases a lot (Gigerenzer and Hoffrage, 1995).

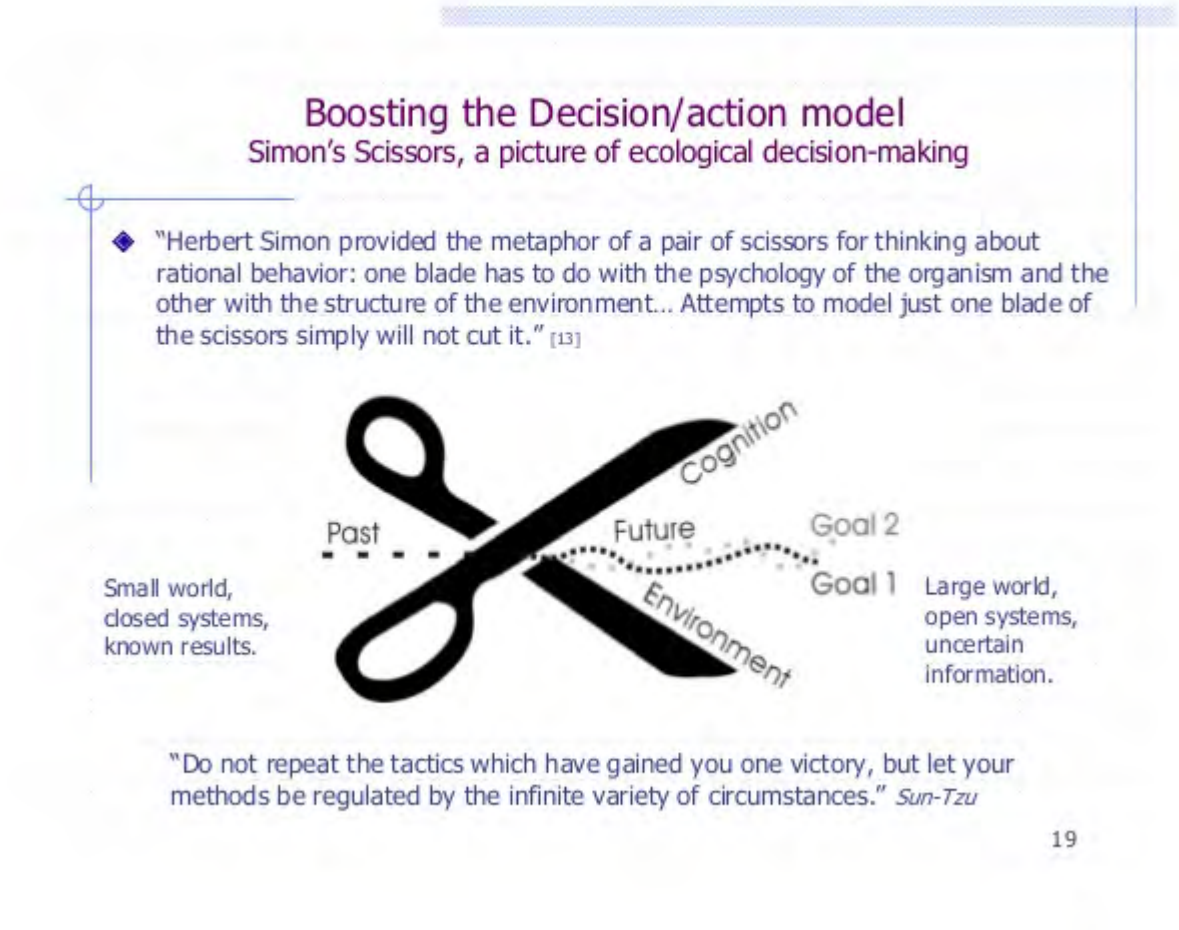
5) Conclusion: the proposal of BRAN (Bounded Rational Adaptive Nudge)

We live in a large word where the canons of neoclassical rationality are unjustified both descriptively and normatively. Therefore the reasoning errors, fallacies and biases that libertarian paternalism are engaged to overcome most of the times are not irrationalities. Moreover the decision making models that behavioural economists have introduced most of the times are as-if instrumentalist tool to fit observed choice data. Adding parameters and transformations to ensure that a weighting-and-adding objective function could fit observed choice data is not a realist process model of decision making as one would expect in the bounded rationality tradition. Therefore libertarian paternalism relies in an unjustified descriptive reasoning theory. Lastly the hedonic paternalism is unjustified in using the attribute libertarian because of the same architecture of mind (System or Processes 1 and System or Processes 2) that libertarian paternalism refer to.

What kind of features a cognitive inspired policy making theory ought to have? What kind of nudges are feasible in order to help the citizens to fulfil their own wellbeing?

The proposal of a Bounded Rational Adaptive Nudge (BRAN[5]) has the following features.

a) Since we live in large word characterized ontologically by complexity, recursitivy and non linearity and epistemically by uncertainty the rationality of choices should be judged by their adaptivity and problem solving ability. In fact bounded rationality is not confined only to the constraints of computational power of human mind. As in the scissors metaphor (fig.3) of Simon (1990) the rationality should be judged by the matching or mismatching of the relation mind-environment or in other words choice-task structure.



What kind of reasoning processes are able to match the environmental tasks and solve the problems? This is an empirical question that has been faced some years ago by some cognitive scientists, as Herbert Simon, Vernon Smith, Richard Selten and in particular more directly Gerd Gigerenzer and the Abc group (Gigerenzer, Todd, and the Abc Group, 1999). The adaptive

[5] BRAN has a double meaning as an acronym but also as something that symbolizes simplicity and frugality.

toolbox of formalized heuristics is the result of these empirical investigation. In a number of problems simple heuristics were more accurate than standard statistical methods that have the same or more information. The results became known as less-is-more effect. There is a point where more is not better, but harmful. There is an inverse-U-shaped relation between level of accuracy and amount of information, computation, or time (Gigerenzer and Gassmaier, 2011, p.453). For example "starting in the late 1990s it was shown for the first time that relying on one good reason (and ignoring the rest) can lead to higher predictive accuracy than achieved by a linear multiple regression" (Gigerenzer and Gassmaier, 2011, p.453). Herbert Simon himself spoke, in his appraisal to the volume of Gigerenzer, Todd and the Abc group (1999), of a "revolution in cognitive science, striking a great blow for sanity in the approach to human rationality". The tool box is composed by many heuristics (table 3 and 4).

Table 1-1: Twelve Well-Studied Heuristics With Evidence of Use in the Adaptive Toolbox of Humans

Heuristic	Definition	Ecologically rational if:	Surprising findings (examp
Recognition heuristic (Goldstein & Gigerenzer, 2002; chapter 5)	If one of two alternatives is recognized, infer that it has the higher value on the criterion.	Recognition validity > .5	Less-is-more effect if $\alpha > \beta$; systematic forgetting can be beneficial (chapter 6)
Fluency heuristic (Schooler & Hertwig, 2005; chapter 6)	If both alternatives are recognized but one is recognized faster, infer that it has the higher value on the criterion.	Fluency validity > .5	Less-is-more effect; systematic forgetting can be beneficial
Take-the-best (Gigerenzer & Goldstein, 1996; chapter 2)	To infer which of two alternatives has the higher value: (a) search through cues in order of validity; (b) stop search as soon as a cue discriminates; (c) choose the alternative this cue favors.	Cue validities vary, high redundancy	Often predicts more accurately than multiple regression (Czerlinski, Gigerenzer, & Goldstein, 1999), neural networks, exemplar models, decision tree algorithms
Tallying (unit-weight linear model; Dawes, 1979)	To estimate a criterion, do not estimate weights but simply count the number of positive cues.	Cue validities vary little, low redundancy (Hogarth & Karelaia, 2005a, 2006b)	Often predicts as accurately as better than multiple regressio (Czerlinski et al., 1999)
Satisficing (Simon, 1955a; Todd & Miller, 1999; chapter 18)	Search through alternatives and choose the first one that exceeds your aspiration level.	Distributions of available options and other costs and benefits of search are unknown	Aspiration levels can lead to substantially better choice tha chance, even if they are arbitr (e.g., Bruss, 2000)
One-bounce rule (Hey, 1982)	Continue searching (e.g., for prices) as long as options improve; at the first downturn, stop search and take the previous best option.	Improvements come in streaks	Taking search costs into consideration in this rule doe not improve performance

Table 1-1: Twelve Well-Studied Heuristics With Evidence of Use in the Adaptive Toolbox of Humans

Heuristic	Definition	Ecologically rational if:	Surprising findings (examp
Gaze heuristic (Gigerenzer, 2007; McBeath, Shaffer, & Kaiser, 1995)	To catch a ball, fix your gaze on it, start running, and adjust your running speed so that the angle of gaze remains constant.	The ball is coming down from overhead	Balls will be caught while running, possibly on a curve path
1/*N* rule (DeMiguel, Garlappi, & Uppal, 2009)	Allocate resources equally to each of *N* alternatives.	High unpredictability, small learning sample, large *N*	Can outperform optimal asset allocation portfolios
Default heuristic (Johnson & Goldstein, 2003; chapter 16)	If there is a default, follow it.	Values of those who set defaults match those of the decision maker; consequences of a choice are hard to foresee	Explains why advertising has little effect on organ donor registration; predicts behavior when trait and preference theories fail
Tit-for-tat (Axelrod, 1984)	Cooperate first and then imitate your partner's last behavior.	The other players also play tit-for-tat	Can lead to a higher payoff than "rational" strategies (e.g. by backward induction)
Imitate the majority (Boyd & Richerson, 2005)	Determine the behavior followed by the majority of people in your group and imitate it.	Environment is stable or only changes slowly; info search is costly or time consuming	A driving force in bonding, group identification, and moral behavior
Imitate the successful (Boyd & Richerson, 2005)	Determine the most successful person and imitate his or her behavior.	Individual learning is slow; info search is costly or time consuming	A driving force in cultural evolution

Note. For formal definitions and conditions concerning ecological rationality and surprising findings, see references indicated and related chapters in this book.

They have been tested successfully against statistical algorithms of rationality not in the easy task of fitting closed sample of data but in the much harder task of prediction. They have proved to be both a better description of decision making and a better prescription on how to decide. Obviously the adaptive success of any given heuristic depends from particular given environment. In which environments will a given heuristic succeed, and in which will it fail? Todd et al. (2011) have identified a number of environmental structure variables:

1) Uncertainty: how well a criterion can be predicted
2) Redundancy: the correlation between cues
3) Sample size: number of observations (relative to number of cues)
4) Variability in weights: the distribution of the cue weights.

How do we assess the adaptive success in ecological rationality? Gigerenzer and Gassmaier (2011, p. 457) write:

The study of ecological rationality results in comparative statement of the kind "strategy X is more accurate (frugal, fast) than Y in environment E"…

What kind of implications to policy making come from ecological rationality? Is it possible an ecological rationality inspired libertarian paternalism and a bounded rational adaptive nudge?

Following the previous distinction of Hedonic, Cognitive and Educational paternalisms I can exclude the first because of the unjustified attribution of the label libertarian[6]. A real libertarian paternalism is aimed to supply the cognitive tools to the people to process better the information and to improve their deliberate problem solving in the large world. In other words to increase their ecological rationality. Therefore the only justified libertarian paternalisms seem to be the cognitive and the educational ones.

What characterized better the BRAN approach is the design of environments of choice that increase the correct utilization of the tool box of heuristics together with a proper education on what and when utilize them. How is it possible to design ecology rational environment? By designing environment that nudges the utilization of a proper suitable heuristic. It is possible for example to design environment that exploit the so called social intelligence by relying on heuristics designed for social information. *Imitate-the-successful* heuristic, for instance, speeds up learning of cue orders and can find orders that excel take-the-best's validity orders (Hertwig and Herzog, 2009). Other heuristics include *imitation heuristics*, *tit-for-tat*, the *social-circle heuristic*, and *averaging the judgements of the others to exploit the "wisdom of crowds"* (Gigerenzer and Gassmaier, 2011).

Simplifying and structure complex choice is also a good challenge for BRAN. Beyond the valid proposals of Thaler and Sunstein (2008) exemplified by the example of paint store, there can be also a BRAN way to structure the complexity. For example trying to select environment that present high redundancy and variability in weights of their structure. High redundancy means structure where cues are highly connected. High variability means structure where there is great difference in weight between some cues and the others. In this structure when there is also high uncertainty it is likely that one-reason decision making as take-the-best heuristic is able to allow successful inferences that can be superior than those based on algorithms as classification and regression tree or conjoint analysis. In most of the choice linked to your wellbeing, as education, health, food, consumption goods, housing, and so on, you have to search for more than one cue. In these cases also you may follow a sequential heuristic that is based on one-reason decision making. An example is "elimination by aspects" lexicographic heuristics to nudge proper choices in large world. How? Structuring for example with a proper software the information given to the families, by fast-and-frugal trees in which is incorporated the lexicographic logic. Let's do the example of the choice of a school: the first question might be "in your opinion what is the most important feature of the schools in your town?". The answer might be "to be among the best five in the ranking in the quality of teaching". Here is the first selection of five. Then the second question might be "among these five schools what is the preferred aspect for choosing one of them?". The answer might be "The cost must be no more than 15 thousand dollars" . Here is the second selection of two schools A and B. The third question might be "between A and B what is the best feature to choose one of them?". The answer might be "the closeness to family house". Here is the final choice of A if it is closer than the other. Otherwise there might be another question asking another comparative feature. This is the typical non-compensatory strategy for choosing in an ecological rational way. In this strategy people order the cues relying on recall from mental sample. A person doesn't need to learn cue orders individually but instead can learn

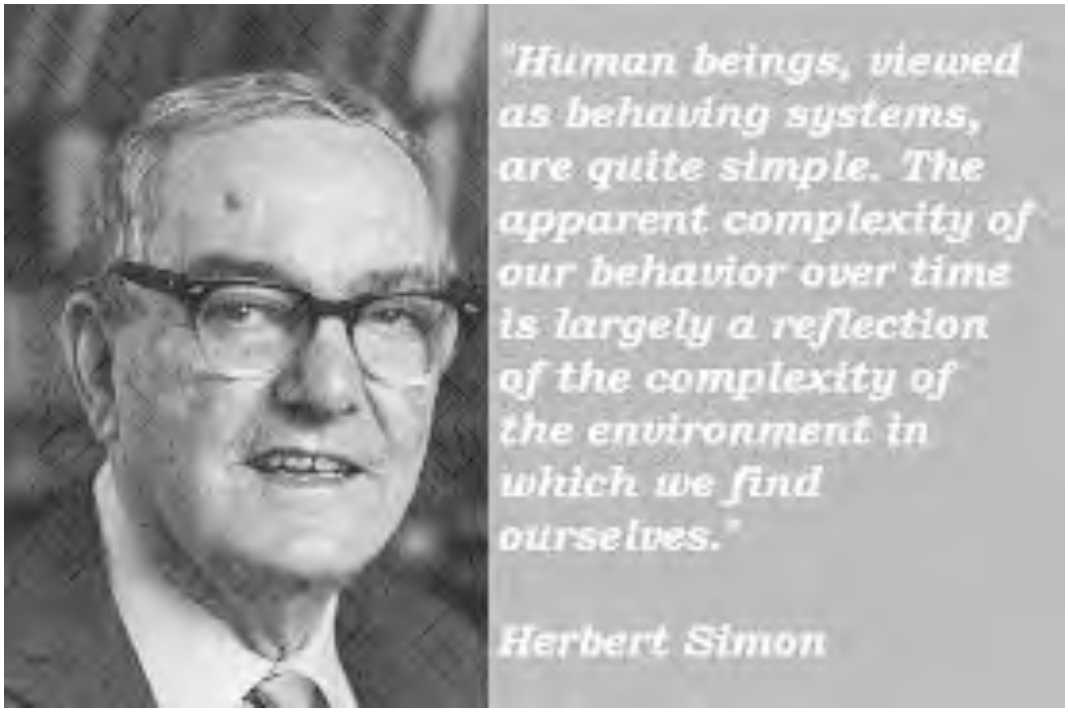
[6] Among the other things even the explanation given to the default effects by people's inertia seem empirically not well grounded. For example experiments by McKenzie, Liersch, and Finkelstein (2006) indicate that people accept the default not for inertia but because interpret it as a recommendation by policy maker.

from others, as through teaching and imitation (Gigerenzer and Gassmaier, 2011). This is an example of BRAN.

Another example BRAN is how to avoid errors in probabilistic choices. Before we have seen that if information is presented as the outcome of learning from experience, known as natural frequencies, and not as conditional probabilities, the proportion of people reasoning by Bayes'rule increases a lot (Gigerenzer and Hoffrage, 1995). The importance of nudging people by the natural frequency format to reason correctly in statistical task is crucial in many environment (Gigerenzer, 2014). In particular the frequency format improves the statistical and the Bayesian reasoning in many medical judgement to predict correctly the probability of a disease according a prior probability and a new evidence (supplied for example by a test with some false positives). The same argument can be applied to many public policies with dramatic future implications for the human life as natural disasters, terrorist attacks, micro criminalities, epidemic, but also more quiet social phenomena where people has prior probability and some new evidence, as the choice of the faculty for the sons in relation to labour market or the choice of an hospital for a surgical operation in relation to the success rate of similar medical institutions.

This topic is related to another important component of cognitive paternalism: how to increase the *knowledge of feedback* from our choices. One of the reason to increase the feedbacks is not only that we can learn from our error and not to fall another time in the same choice. It is also that we can improve inductively our theories of the world. That is, we can improve our prediction on future states of the world, for example our future choice of a party or of a school. In the experiments on Bayesian learning people learn probabilities from experience and are subsequently tested as to whether they make judgements consistent with Bayes' rule. Many time the test are successful. Therefore many cognitive scientists conclude that people's judgements are largely consistent with it (Chater and Oaksford, eds., 2008; Gigerenzer, 2015). This kind of test are the cognitive justification for an ecological rational role of the nudges that manage to increase the knowledge of the feedback from people choices.

It is possible also to design ecologically sound *mapping of the choice into future welfare*. For example when an individual has to make a choice of different mortgages or credit agreements it is possible to simulate future simple environments with few cues (for example the monthly rate) and ask him to imagine to be in that situation. In this case the attempt is to create a situational rationality dimension and to trigger embodied cognition aspects of the choice. This situation would allow him to understand better the future effects of his choice trying to make converge subjective present utility with future utility.



6) Bibliography

Arkes, H.R., Gigerenzer, G., and Hertwig, R. (2015). *Coherence cannot be a universal criterion for rational behaviour: a systematic and theoretical review.* Submitted manuscript.

Baratgin, J., and Noveck, I.A. (200). " Not only base rates are neglected in the engineer-lawyer problem: an investigation of reasoners' underutilization of complementarity". *Memory and Cognition,*28, 79-91.

Bobbio, N. (2009). *Eguaglianza e Libertà.* Torino: Einaudi.

Bond, M. (2009). " Risk School". *Nature 461(29)*: 1189-92.

Butchel, E., E. and Norenzayan, A. (2009). "Thinking across cultures: Implications for dual processes", in K. Frankish and J. Evans (eds), *In Two Minds: Dual Processes and Beyond.* Oxford: Oxford University Press.

Cader, R., S. Campbell, D. Watson (2005). "Cognitive Continuum Theory in nursing decision-making" *Journal of Advanced Nursing,* 49(4), 397-405.

Chater, N., and Oaksford, M. (2008). *The probabilistic mind: Prospects for Bayesian cognitive science.* Oxford: Oxford University Press.

Cleermans, A., and L., Jimenez (2002). "Implicit learning and consciousness: A graded, dynamic perspective". In R.M. French and A. Cleermans (eds). *Implicit learning and*

consciousness: an empirical, philosophical and computational consensus in the making (pp. 1-40). Hove, UK: Psychology Press.

Cohen, L.J. (1981). " Can human irrationality be experimentally demonstrated?", *Behavioral and Brain Sciences,* 4, 317-370.

Cosmides, L. and J. Tooby (1996). " Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgement under uncertainty". *Cognition 58*: 1-73

De Miguel, V., L. Garlappi, and R. Uppal (2009). "Optimal versus naïve diversification: how inefficient is the 1/N portfolio strategy? *Review of Financial Studies 22*: 1915-1953.

Dennett, D. (1987). *The intentional stance.* Cambridge Mass: MIT Press.

Dennett, D. (1991). *Consciousness Explained.* Boston: Back Bay Books – Little, Brown Company.

Dennett, D. (1996). *Kinds of minds: Towards an understanding of consciousness.* New York: Basic Books.

Engel, C., and Gigerenzer, G. (2009). "Law and Heuristics: An interdisciplinary venture". In G.Gigerenzer and C.Engel (eds). *Heuristics and the Law.* Canbridge Mass.: The MIT Press.

Evans, J.St.B.T. (2007). *Hypotheticaal thinking: Dual processes in reasoning and judgement.* Hove: Psychology Press.

Evans, St. B. T. J. (2008). "Dual-processing accounts of reasoning, judgement and social cognition". *Annual Review of Psychology.* 59, 255-278.

Evans, St. B. T. J. (2009). "How amny dual-processes theories do we need? One, two or many? in K. Frankish and J. Evans (eds), *In Two Minds: Dual Processes and Beyond.* Oxford: Oxford University Press.

Evans, St. B. T. J. and K. E. Stanovich (2013). "Dual-Process Theories of Higher Cognition: Advancing the Debate", in *Perspective on Psychological Science* 8(3) 223-241.

Fehr, E., and Schmidt, K. (1999). " A theory of fairness, competition and cooperation". *Quaterly Journal of Economics,* 114, 817-868.

Fong, G.T., D.T. Krantz, and R.E. Nisbett, (1986). " The effects of statistical training on thinking about everyday problems". *Cognitive Psychology* 18: 253-292.

Frankish, K. (2004). *Mind and Supermind.* Cambridge: Cambridge University Press

Frankish, K. (2009). " Systems and Levels: Dual-systems theories and the personal-subpersonal distinction" in K. Frankish and J. Evans (eds), *In Two Minds: Dual Processes and Beyond.* Oxford: Oxford University Press.

Frankish, K. and J. Evans (2009). "The duality of Mind: an historical perspective" in J. Evans and K. Frankish (eds.), *In Two Minds.* Oxford: Oxford University Press.

Friedman, M. (1953). *Essays in Positive Economics.* Chicago: University of Chicago Press.

Gawronski, B., and Bodenhausen, G.V. (2006). "Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change". *Paychological Bulletin,* 132, 692-731.

Gigerenzer, G. (2014). *Risk savvy: how to make good decisions".* New York: Viking.

Gigerenzer, G. (2015)." On the Supposed Evidence for Libertarian Paternalism". *Rev.Phil.Psych.*DOI 10.1007/s 13164-015-0248-1

Gigerenzer, G. (2016). "Rationality without optimization: Bounded rationality" in L.Macchi, M. Bagassi, and R. Viale (eds.) *Human Rationality and the Cognitive Unconscious.* Cambridge Mass: The MIT Press.

Gigerenzer, G, and W. Gassmaier, (2011). "Heuristic Decision Making", *Annual Review of Psychology,* 62: 451-82.

Gigerenzer, G. and Hoffrage, U., (1995). "How to improve Bayesian reasoning without instructions: frequency formats", *Psychological review,* 102, 684-704.

Gigerenzer, G., P.M.Todd, and the ABC Research Group (1999). *Simple Heuristics that make us smart.* New York: Oxford University Press.

Gigerenzer, G., and Selten, R. (eds.) (2001). *Bounded rationality: The Adaptive Toolbox.* Cambridge Mass: The MIT Press.

Grice, P. (1989). *Studies in the way of words.* Cambridge Mass: Harvard University Press

Griggs, R.A. and J.R. Cox (1982), "The elusive thematic-materials effect in Wason's selection task" in *British Journal of Psychology, 73*: 407-420.

Hacking, I. (1983). *Representing and Intervening.* Cambridge: Cambridge University Press.

Hammond, K.R. (1996). *Human judgement and social policy.* Oxford: Oxford University Press

Hertwig, R., and Herzog, S.M. (2009). “Fast and Frugal Heuristics: Tools for of social rationality”. *Social Cognition,* 27, 661-98.

Hume, D. (1739). *A Treatise of Human Nature.* London: Penguin Classics

Johnson Laird, P., P. Legrenzi, and M. Sonino Legrenzi (1972). “Reasoning and a Sense of Reality”, in *British Journal of Psychology, 63*: 395-400.

Kahneman, D. (2011). *Thinking fast and slow.* London: Allen Lane.

Kahneman, D., and Federick, S. (2002). “Representiveness revisited: Attribute subsistution in intuitive judgement. In T. Gilovich, D. Griffin & D. Kahneman (eds.), *Heuristics and Biases: The psychology of intuitive judgment* (pp. 49-81). Cambridge: Cambridge University Press.

Klein, G. (1999). *Sources of Power.* Cambridge Mass: the Mit Press.

Krueger,J and Funder, D.C. (2004). “Towards a balanced social psychology: Causes, consequences and cures for the problem-seeking approach to social cognition and behaviour”, *Behavioral and Brain Sciences,* 27, 313-76.

Kruglanski, A.W. and G. Gigerenzer (2007). “Intuitive and deliberative judgements are based on common principles”, *Psychological Review,* 118, 97-109.

Laibson, D. (1997). “Golden eggs and hyperbolic discounting”, *Quaterly Journal of Economics,* 112, 2, 443-77.

Lieberman, M.D. (2099). “ What zombies can’t do: A social cognitive neuroscience approach to the irreducibility of reflective consciousness”. In J. Evans and K. Frankish (eds.), *In Two Minds.* Oxford: Oxford University Press.

Macchi, L. Bagassi, M., and Viale R. (eds.) (2016). *Human Rationality and the Cognitive Unconscious.* Cambridge Mass: The MIT Press.

Macchiavelli, N. (1532; ed. it. 1961). *Il Principe.* Torino: Einaudi.

McKenzie, C.R.M., Liersch, M.J., and Finkelstein, S.R. (2006). “Recommendations implicit in policy default”, *Psychological Sciences,* 17, 414-20.

Nisbett, R. (2003). *The Geography of thought: How Asians and Westerners think differently… and why,* New York: The Free Press.

Nisbett, R. (2009). *Intelligence and how to get it: why schools and cultures count.* New York: Norton.

Nisbett, R and Masuda, T. (2006). “Culture and Point of View”, in R. Viale, D. Andler, and L. Hirschfeld (eds.), *Biological and Cultural Bases of Human Inference,* Mahwah, NJ: Lawrence Erlbaum Associated, Publishers.

Norenzayan, A. (2006). "Cultural Variation in Reasoning" in R. Viale, D. Andler, and L. Hirschfeld (eds.), *Biological and Cultural Bases of Human Inference,* Mahwah, NJ: Lawrence Erlbaum Associated, Publishers.

Oliver, A. (2013). "From Nudging to Budging: Using Behavioural Economics to Inform Public Sector Policy". July: 1-16.

Osman, M. (2004), "An evaluation od dual-process theories of resoning". *Pshychonomic Bulletin and Review,* 11 (6), 988-1010.

Petty, R:E:, and Cacioppo, J.T. (1986). "The Elaboration Likelihood Model of persuasion. In L. Berkowitz (ed.) *Advances in experimental social psychology.* New York: Academic Press.

Savage, L.J. (1954). *The Foundations of Statistics.* New York: Dover. 2nd ed.

Sah, S., Cain, D:M:, Loewenstein, G. (2013). "Confessing one's sins but still committing them: transparency and the failure of disclosure". In A. Oliver (ed.) *Behavioural Public Policy.* Cambridge: Cambridge University Press.

Semin, G.R. and Smith, E.R. (eds.) (2008). *Embodied grounding: Social, cognitive, affective and neuroscientific approaches.* Cambridge: Cambridge University Press.

Sen, A. (2009). *The Idea of Justice.* London: Allen Lane

Simon, H. A. (1955). " A Behavioral Model of Rational Choice". Quaterly Journal of Economics., 69: 99-118.

Simon, H.A. (1985). "Human nature in politics: the dialogue of psychology with political science". *American Political Science Review* 79: 293-304.

Simon, H.A. (1979). "Rational Decision Making in Business Organization", *American Economics Review,*69, 493-513.

Simon, H.A. (1990). "Invariants of Human Behaviour", *Annual Review of Psychology,* 41, 1-19.

Simon, H.A. (2000). "Bounded rationality in social sciences: Today and tomorrow", *Mind&Society,* 1, 25-41.

Simon, H.A. (2009). *An Empirically Based Microeconomics.* Cambridge: Cambridge University Press.

Simon, H.A., Egidi, M., Viale, R., and Marris, R. (1992). *Economics Bounded Rationality and the Cognitive Revolution.* Aldershot: Elgar.

Sloman, S. (1996). " The empirical case for two systems of reasoning". *Psychological Bulletin, 119.* 3-22.

Sloman, S., A. (2002). "Two systems of reasoning". In T.Gilovich, D. Griffin, & D.Kahneman (eds.). *Heuristics and Biases: The Psychology of Intuitive Judgement* (pp. 379-398). Cambridge: Cambridge University Press.

Smith, A. (1759). *The Theory of Moral sentiments,* London: Penguin Classics 2010.

Smith, E.R., E. C. Collins (2009). "Dual-process theories: A social psychological perspective. in J. Evans and K. Frankish (eds.), *In Two Minds.* Oxford: Oxford University Press.

Stanovich, K., E. (1999). *Who is rational*? *Studies of individual differences in reasoning.* Mahwah, NJ: Erlbaum.

Stanovich, K.E. (2004). *The robot's rebellion: Finding meaning in the age of Darwin.* Chicago: University of Chicago Press.

Stanovich, K.E. (2009). "Distinguishing the reflective , alghoritmic and autonomous minds: Is it time for a tri-process theory? In J. Evans and K. Frankish (eds.), *In Two Minds.* Oxford: Oxford University Press.

Stiglitz, G. (2010). *Freefall: America, free Markets, and the Sinking of the Worls economy",* New York: Norton.

Sudgen, R. (2009) "On nudging a review of Nudge by Richard H. Thaler and Cass R. Sunstein". *International Journal of Economics of the Business" 16:* 365-373.

Thaler, R.H. (1991). *Quasi Rational Economics.* New York: Russell Sage Foundation.

Thaler, R and C.R.Sunstein. (2008). *Nudge: improving decisions about health, wealth, and happiness.* New Haven, CT: Yale University Press.

Todd, P.M., Gigerenzer, G., and ABC Group. (2011). *Ecological Rationality: Intelligence in the World.* New York: Oxford University Press.

Tversky, A. e D. Kahneman, (1986), "Rational choice and the framing of decisions". *Science 211:* 453-458.

Van Fraassen, B. (1980). *The Scientific Image*. Oxford: Oxford University Press.

Viale, R. (2012). *Methodological Cognitivism. Vol. I, Mind Rationality and Society.* Heidelberg: Springer.

Viale, R. (2013). *Methodological Cognitivism. Vol. II, Cognition, Science and Innovation.* Heidelberg: Springer.

Viale, R. and D. Osherson (2000). "The diversity principle and the little scientist hypothesis", *Foundations of Science 5(2)*, 239-253.

Viale, R. and D. Osherson (2006). "Cognitive Development, Culture, and Inductive Judgement" in Viale, R., D. Andler, L. Hirschfeld (eds.). *Biological and Cultural Bases of Human Inferences.* Mahwah: Erlbaum.

Viale, R., D. Andler, L. Hirschfeld (eds., 2006). *Biological and Cultural Bases of Human Inferences.* Mahwah: Erlbaum.

Wason, P.C. (1966), "Reasoning" in Foss, B., M. (ed.) *New Horizons in Psychology.* Harmondworth: Penguin.

Wegener, P.T. and Petty, R.E. (1997). "The flexible correction model: The role of naïve theories of bias in bias correction. In M.P. Zanna (ed.) *Advances in experimental social psychology.* San Diego, CA: Academic Press.

Wilson, T.D., and Scooler, J.W. (1991). "Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology,* 60, 181-92.