Behavioral rationality as a foundation for public policy studies

Bryan D. Jones

Department of Government, University of Texas at Austin, United States

Received 22 August 2016; received in revised form 11 December 2016; accepted 18 January 2017

Available online 28 January 2017

Abstract

Cognitive processing applies to a single human being, yet most key social processes are organized in collectives, and connecting them is not a simple task. I review the basic ways in which social and behavioral scientists have tried to link the individual with the system, illustrating the progress that has been made in this endeavor. We are closer than we ever have been to producing a behavioral model that integrates cognitive science yet does not produce a confusing overabundance of findings that have systems level implications. I conclude that we may well be at a tuning point in which the fully rational model of human choice, currently used as the microfoundation for economics and the study of political institutions, is replaced by a more robust behavioral model of choice that relies on developments in the cognitive sciences.

© 2017 Published by Elsevier B.V.

Keywords: Public policy; Decision-making; Information processing; Public administration

1. Introduction

In this paper, I address a (and perhaps the) central theoretical issue concerning how human cognitive processes relate to collective human action. How do the cognitive capacities of individuals relate to macro-phenomena such as the path of public policy activities in a nation or the performance of an economy? The issue is one of scale, or level of analysis. Cognitive processing applies to a single human being, yet most key social processes are organized in collectives. In order to study collective action within the organizations or systems of organizations that characterize modern societies, one must link key elements of individual cognitive processing and collective activities in governments, economies, and other societal level systems.

Here I will concentrate on developments in public policy process studies and linkages to behavioral models of decision-making. The state of the field is such that our understandings of the micro-macro linkages remain incomplete. As a consequence, I cannot definitively detail these connections, but I will summarize where we are in the project. Moreover, it is likely that the connections will never be entirely specified, for two reasons. First, humans are complex beings, as are the organizations and social systems they inhabit. As a consequence linkages are likely to be incomplete. Second, scholarship on cognitive systems and on policy processes will doubtless provide new lines of inquiry in the future. Nevertheless a modern summary of the issues is surely in order.

Policy processes is the study of how public policy is actually made, as distinguished from policy recommendations or policy evaluation, which focus more on the connections between the design of policies and their impacts. As a consequence, it relies on behavioral science for models of human individual choice in the process. But I will also touch on developments in behavioral economics and political behavior, as they are also relevant to the linkage problem in the policy sciences. I conclude that we may well be at
a tuning point in which the fully rational model of human choice, currently used as the microfoundation for economics and the study of political institutions, is replaced by a more robust behavioral model of choice that relies on developments in the cognitive sciences.

Generally behavioral sciences study the internal activities of individuals. Social sciences focus more on the interactions of people in collectives—and in particular in modern societies, how complex organizations interact with members, on the one hand, and with other complex organizations, on the other. We generally think of psychology as well as parts of sociology and political science as behavioral sciences, whereas much of sociology, political science, and economics are social sciences.

Political science consists of two broad sub-fields—the study of political behavior (basically a behavioral science) and the study of institutions (clearly a social science). Students of political behavior study the relation of attitudes to politically relevant behavior such as voting behavior, among other topics. The study of political behavior began in the 1920s at the University of Chicago, and subsequently grew as the field of political science grew.

In economics, the foundation of analysis rested on the assumption of fully rational actors and increasingly abstract and mathematical theories relying on maximizing based on the assumption of complete, or comprehensive, rationality. But that foundation was increasingly challenged by behavioral economics, in large part based on the behavioral science of psychology, beginning in the 1970s (Thaler, 2015).

The study of political institutions traced a radically different path. Emerging from an essentially descriptive and legal foundation, the study of political institutions linked closely to a micro-framework of bounded rationality. Herbert Simon was the key figure in this development; he studied public administration and political science at the University of Chicago in the 1930s, just as behavioral studies were emerging there. Simon insisted that the study of public organizations be based in the study of decision-making, and that any model of decision-making take account of what he termed the “raw material” of those decisions—human nature. Later under the rubrics of neo-institutionalism and social choice, political scientists and economists studied political institutions using a rational framework, but the approach never dominated the field and in the study of policy processes, bounded rationality has always served as the fundamental micro-foundation.

### 2. The problem of integrating cognitive science into policy studies

The two social science disciplines most intimately connected with the study of public policy—economics and political science—have both had trouble developing a consistent convincing integration of theories and findings from the cognitive sciences into an appropriate underpinning for the organizational dynamics that must characterize the study of these disciplines. Integrating cognitive studies into organizational systems of human action requires a delicate balance. If the approach simplifies too radically the internal organizational dynamics, then the consequences are to rely entirely on the stream of information and demands from the external environment. But too much detail drawn from modern cognitive science leads to a messy and imperfect linkage between the cognitive components of outputs and those driven by straightforward organizational adaptation. As Bendor (2010: 45) put it, “macrofields such as political science must be more ruthless than related microfields about microassumptions”. In a word, the integration of cognitive science into any part of social science is a difficult problem.

There are only two paths to solving the problem. On the one hand, scholars can make bare-bones assumptions about human decision-making. The classic approach, comprehensive rationality, is to assume that individuals have tastes or preferences, and that these preferences are fixed, at least for the purposes of analysis. Individuals make decisions based on the best match between their preferences and the alternatives before them. If there is a cognitive basis in the approach, it is Skinnerian psychology, in which humans respond to positive and negative stimuli in an optimal fashion and cognition is more an epiphenomenon than part of a causal chain from information to action. As a consequence, there is no need for any examination of the cognitive states of mind. On the other hand, scholars can introduce elements of the cognitive and emotional architecture of decision-makers into the model. But which elements? This is the crux of the problem, and it is a difficult problem indeed.

Even though the assumptions of the preference-based model are demonstratively unreal, nevertheless the approach reveals all sorts of problems in achieving collective choice. Collective choices represent the outcomes of a policymaking process, so that the study of them is critical for policy studies. Some of these issues are actually solved by dropping a fixed preference model and integrating elements of human cognitive architecture. To take one very important example, in the study of cooperation, early thinking led to skepticism about cooperation in the case of collective goods, because rational actors would become free riders. Only with coercion could cooperation be forged. However more intensive studies in the laboratory and in the field showed that with communication even un-coerced actors can cooperate to solve collective action problems (Ostrom, 1990, 1999). People indeed might not cooperate, but they could forge norms to foster such cooperation. Without an appreciation of these aspects of human nature, this could not be understood. Ostrom (1998) terms this kind of phenomenon “better than rational”.

There are other reasons having nothing to do with the inherent difficulties of balancing realism in a behavioral model and generalizability to systems-level behavior that have hindered the development of a robust cognitive infrastructure for policy sciences. Both political science
and economics have developed structures that interfere with the development of an integrated framework. In economics, the resistance has been fundamental: the discipline’s commitment to a model of human behavior that is rational rather than behavioral. The pious assumption (and it is nothing more) of fully rational action on the part of economic actors is convenient for developing mathematical models of economies, and serves as a normative guide to making decisions (that is, how one should make economic decisions). In many ways, the system works well, but it fails to explain major events—economic bubbles, for example (Shiller, 2000).

In political science, the problem is very different: the division of the field into behavioral and institutional studies. Behavioral studies are indeed focused on the psychological dynamics of the individual actor, and rely on cognitive sciences work in many fields. But it generally limits itself to studies of mass behavior such as voting choice, and has contributed little to the development of a behavioral base for organizational dynamics that dominate the study of public policy. Indeed, what insights we have in the design and implementation of government programs have come more directly from public policy and public administration studies or from behavioral economics, a relatively new field, than from behavioral political science, whose roots can be traced back to the 1920s. Institutional studies in political science, on the other hand, have tended to be focused on the exact nature of governing arrangements and their incentives for political actors—a focus on strategic dynamics generated by institutional rules and arrangements generally disassociated from the theories and findings of behavioral political science, behavioral economics, or the cognitive sciences more generally. The key assumption is that formal rules generate incentives that actors with fixed preferences respond to in an optimal manner.

To summarize: in political science, behavioral studies, by focusing on cognitive and emotional aspects of individuals, have tended to ignore political institutions and policy-making, and hence have provided little in the way of a behavioral theory of policymaking. Institutional studies have tended to center on assumptions of fixed preferences and optimal action in response to incentives, hence providing little direction toward integrating findings from cognitive science or even behavioral political science into a behavioral theory of institutions and policy.

3. Bounded rationality as a foundation for the study of public policy: a brief history

What micro-macro linkage occurs in political science emerged in the study of policy processes. Bounded rationality arose in the study of public administration (there was no separate field of public policy before the 1970s) as a ‘negative’ model of decision-making in public organizations. Developed by Herbert Simon in Administrative Behavior (1947), bounded rationality was highly critical of the emerging fascination in economics with “comprehensively rational” decision-makers—those who were capable of making choices through comparisons of alternatives and maximizing the expected utility from those choices. He later admitted that the approach articulated in that book consisted largely of “residual categories” and that “the positive characterization of the process of choice is very incomplete” (Simon, 1976: xxix). Simon used the term “behavioral rationality” to denote this positive model of choice, contrasting with the negative model, bounded rationality.

But Simon wanted to understand organizations, and his Administrative Behavior insisted that any theory of administration (or, indeed, any sensible study of administration) be based on an understanding of the individual decision-making process. As he later commented, “the book’s aim is to show how organizations can be understood in terms of their decisional processes” (Simon, 1996: vii; see also Simon, 1992). The study of public administration before Administrative Behavior dealt almost exclusively with the organizational or inter-organizational levels of analysis, with little attention to micro-models of human decision-making. Public administration understandably also contained a strong normative element, focusing on the elements of good administration and how to encourage it. Simon thought that such normative questions should be separated analytically from empirical issues so that a science of administration could be built.

Simon’s work was perhaps more controversial in public administration because of this latter point than his decisional analysis. But that left somewhat underdeveloped the level of analysis conundrum: if one is to understand the connection between individual decisional processes and the operation of organizations and their interactions within systems of organizations, how is that link to be forged? This is not an easy leap, and as a consequence the full flowering of what is termed “behavioral organization theory” required both a positive model of how humans made decisions and a theory of how those decisions were related to organizational outputs.

As early as 1959, Simon and other scholars had developed a fairly complete model of behavioral rationality (Bendor, 2010; Jones, 2002, 2003). The model was premised on the notion of intention: Simon postulated that generally decision-makers were goal-directed and desired to maximize outcomes, but because of their cognitive constitutions they were incapable of doing so. They made mistakes (when applied to the objective standards of full maximization), and not infrequently human decision-makers made these mistakes repeatedly.

The date of 1959 is not arbitrary. It marked the confluence of five important streams of development in behavioral science that transformed these sciences (unfortunately the impact was less intense in the social sciences). These include:
The articulation of an outline of a positive model of behavioral choice (Simon, 1957). The positive approach included an alternative approach to the expected utility model, “satisficing”, in which decision-makers stop searching for solutions when they find an alternative that satisfies a pre-set but adjustable aspiration level rather than continuing to search for a yet-better alternative.

The emergence of a new cognitive psychology based on “getting inside the heads” of subjects in which Newell and Simon (1958) asked subjects to talk through how they were solving problems. This approach led to a firm psychological foundation for the notion of “cognitive shortcuts” or “heuristics” as critical for the search for solutions in human problem solving.

The emergence of the study of artificial intelligence (AI)—the notion that computers could be used to model thought (now a much more controversial idea in cognitive science, but a breakthrough at the time) (Newell, Shaw, & Simon, 1957; Newell & Simon, 1956). As Simon (1996: 27) latter put the matter, AI provided alternate methods to solving complex problems “in the form of heuristic search (selective search using rules of thumb)” finding “decision that are ‘good enough’, that ‘satisfice’.

The publication of Charles Lindblom’s (1959) classic study of incomplete search in policy choice based on incomplete understanding of complex problems and local, incremental search for solutions involving “successive limited comparisons”. This paper presaged a whole research program in the study of public budgeting (Bendor, 2010: 23–28). I will return to this program later in this essay.

The maturing of behavioral organization theory. Simon (1947) emphasized how organizations compensated for the mismatch between the cognitive abilities of a single human and the magnitude of the problems that needed to be addressed. March and Simon (1958) began to develop a more robust union between the elements of organizations that compensated for human weaknesses and reinforced them through, for example, out-of-date heuristics incorporated into bureaucratic routines. Clearly both are relevant.

4. Behavioral organization theory as the key

In the first edition of Administrative Behavior, the author states clearly, and reinforces the statement in later editions, that “decision-making is the heart of administration”, and that “administrative theory must be derived from the logic and psychology of human choice” (Simon, 1997: ix, xi). But this is a much messier task that building a theory of administration based on the maximization principles of homo economicus. Findings from psychology and biology offer a bewildering array of potentials for linking individual decisions and collective ones. Assumptions are necessary.

A place to start is by observing that, while organizations expand human capacities, they also canalize many facets of human cognitive architecture. By ‘canalization’ I mean that behavior is pushed into canal-like systems of similar repetitive actions. Like humans, organizations must process information from their environments. They generally expand capacities for processing information by specialization; for example, the US Congress employs a decentralized committee structure that both allows for specialization by policy area and for parallel processing of the streams of information. To process this information, each subunit must prioritize incoming stimuli by importance, and the central processing system (in the case of Congress, the chamber leaders) must similarly prioritize, because every enacted policy requires some central action that involves serial processing of even the information filtered by the committee system. As a consequence, committees, as well as the chamber itself, have limited attention spans. Limited attention is a key facet of human cognitive capacity, and is reflected in organizations.

I have proposed that the key assumption of behavioral organization theory be centered in exactly this causal canalization of human cognitive capacities into features of the organization, and that this relation is not simply metaphorical—although it is that. The relation is causal; the connection between organizational agendas and human attention is based not on organizational limits but human ones (Jones, 2001: Chapter 6). The cognitive architectures of humans that causes canalization of behavior in human action is directly and causally related to similar processes in organizations. This implies a research approach that isolates enduring components of the organization that can be demonstratively linked to limitations in human capacities in processing information.

This linkage must be forged at the empirical level, with empirical generalizations and inductive reasoning. I will pursue this avenue later in this essay. But it must also be pursued at the formal level—by foregoing the assumptions of comprehensive rationality imbedded in economics with a focus on the enduring problems that affect organizations within the assumptions of behavioral rationality. Several scholars are pursuing this route today. Jonathan Bendor (2010) and Thomas Hammond (1986) have independently and collectively (Bendor & Hammond, 2010) produced a body of work that has gone a long way to providing this general framework. However, Hammond (1986, 1990) has repeatedly stressed the independent effects of structure beyond what could be explained by tenets of behavioral organization theory, suggesting that there is more to the story than the simple one of organizations overcoming human limitations or canalizing them.

5. Studies in bounded rationality: earlier studies focusing on solution search

Early students of decision-making in government relied heavily on bounded rationality as an underpinning for their
mostly observational work. Often they relied directly on Simon’s insights, but it was independently clear that the data derived from observational studies conflicted with ideas of good decision-making developed in economics and in public administration. Problems were complex and changing, oftentimes involving conflicting interpretations of goals; moreover it was clear that human capacities for dealing with such complexities were limited.

In addressing what they were observing, these scholars focused on the solution much more than on the definition of the problem. Lindblom’s (1959) classic decision-making paper both critiqued what he called the “root” or “rational-comprehensive” approach to decision-making and offered his notions of a more sensible approach—that is, based on human capacities and problem complexities. He noted that the root approach, involving ends-means analysis, was only possible when there was agreement on goals and the operationalization of those goals. Because of problem complexity and divergence of values, most choice in public administration involved a “branch” approach based on “successive limited comparisons” and incrementalism, in which only minor shifts from the existing policy were considered.

In the end the focus on solution search in the local region of existing policy solutions as a system of policy choice (Lindblom, 1959: 88) failed, both as an adequate description of policymaking and as a theoretical concept capable of bridging the “levels of analysis” issue. But it succeeded as a critique of rational-comprehensive decision-making (Bendor, 2010). As we shall see, it failed because it was not an adequate description of policy change—there were too many large-scale policy changes for a ‘successive limited comparisons’ approach to serve as an adequate description. But by expanding the model to incorporate problem search as well as solution search, and by using the simple notion that the focus of decision-making attention is critical, I show that Lindblom’s model, and those of other bounded rationalists, particularly Aaron Wildavsky, remain relevant within a broader framework.

6. The study of budgets

Theories of public budgeting emerged from what we would call today an adaptive systems framework (Miller 
& Page, 2007). Any system interacts with its environment in a way that reflects both the nature of the environment and the internal constraints of the system itself. As Simon (1996; see also Jones, 2001) emphasizes, human organizations are continually adapting both to external flows of information and the internal structure of the organization, including the cognitive capacities of the human actors who occupy positions in the organization.

The serious examination of government budgets as a major component of decision-making began in the 1940s and 1950s. Indeed, Simon’s insights on the limits of rationality began with his observations of budgeting while working in city government. He observed then that few of the strictures of “good” decision-making were followed because problems are complex and cognitive abilities limited. For example often “the goals of an organization cannot be connected operationally with action” leading to the substitution of simplified shortcuts, or heuristics, for goals (Simon, 1979: 353). A major reason involved disagreements on how the generally-agreed upon goals were to be translated into action, resulting in a great deal of tension.

Early models of budgeting emphasized internal organizational dynamics. While they did not ignore changes in the external environment, they certainly focused on the interchanges among various components of the budget-making system. Wildavsky’s (1964) incrementalist model, developed from his incisive observations of budget construction in the US federal government, centers on the limited rationalities of decision-makers and the emergence of heuristic rules for building budgets—the two most important being “base” (start at last year’s budget) and “fair share” (make adjustments from the base based on well-understood changes in the demand structure for services by the various participants). It is not that the early students of budgeting ignored external forces; indeed, they saw them as highly complex and uncertain. As a consequence, incremental budgeting became both an empirical and a normative theory of decision-making because making major changes in the face of a highly uncertain and complex world was a dangerous strategy, especially given the cognitive constraints of decision-makers (Lindblom, 1959).

The first efforts at assessing the incrementalist model used a regression framework with budgetary change within policy categories as the dependent variable (Davis, Dempster, & Wildavsky, 1966, 1974). As is often the case, this operationalization of Wildavsky’s verbal model incompletely captured its richness (Bendor, 2010, Chapter 1). While this system worked reasonably well, the investigators had to insert dichotomous variables to indicate differences in “budgetary eras” to achieve satisfactory statistical fits. This implied that the incremental model could not explain all of budget politics, and was hence incomplete.

John Padgett transformed our view of the incrementalist model in a pathbreaking paper in 1980 in which he showed that the incrementalist model implied that a distribution of first budget differences would be Normal.1 But Padgett’s data on annual budget requests in the President’s budget submission to Congress were far from Normal. He went on to develop a decision-making model, the serial processing model, that relied on limited successive comparisons that involved a decision-maker working through a trial-and-error process until an external constraint was encountered, which ended the search.

Padgett’s paper exposed a serious flaw in the incrementalist model, and offered an approach to addressing the internal dynamics of budget systems. But it did not grapple

1 Heterogeneous budget estimates, that is, cross-category estimates, imply Student’s t distributions.
with the impacts on a budgetary system of external flows of information. Nor did the serial processing model supplant the incrementalist model as a standard of internal decision-making behavior. It is fair to say that the full implications of Padgett’s stochastic process approach were not appreciated until a decade and a half later, when Jones, Baumgartner and True recognized the relevance for Padgett’s approach for the punctuated equilibrium model they had developed (Jones, Baumgartner, & True, 1998; True, Jones, & Baumgartner, 1999).

6.1. Did the boundedly rational models developed to examine decision-making in government fail?

By the time of Padgett’s papers, a consensus had developed among students of the process of budgeting that the early models developed by Lindblom and Wildavsky had failed. This is by no means the case. As Bendor (2015) argues, Lindblom had two major purposes. One was to document the failure of the comprehensive-rational model as both a guide to public sector decision-making and as an adequate model of how decisions are actually made. The second was to provide a positive model of choice that would serve as an adequate guide to practical decisions: the method of successive limited comparisons, with its focus on limited search in the area around the existing policy solution, leading to disjointed incrementalism. The first of the two, Bendor argues, was successful, especially in the applied areas of public administration. The second objective was not as successful, at least judged by its use in the subsequent literature. If Lindblom’s prescriptive model failed, it may be time to re-evaluate it—along the lines begun by Lindblom (1979) himself by adding a shift to strategic action when incrementalism failed. Empirical findings from the recent budget literature have reinforced the notion that smaller incremental changes are superior to larger ones, at least in the sense that they are longer lasting (Breunig & Koski, 2012).

Throwing out the model because it fails in an aggregate data test, whether based on regression or empirical estimates of probability distributions, was premature. The initial models were based on observations of the decision-making process, and it is entirely possible that most budget decisions are made with simple heuristics that are interrupted occasionally by major adjustments. We know, however, that these adjustments cannot be explained by ‘eras’ of budgeting; they have a much less predictable pattern.

The most powerful case against incrementalism comes from stochastic process studies initiated by Padgett and since used to examine budget authority in the US (Jones & Baumgartner, 2005) and elsewhere. However, it is not simple to link a stochastic distribution of budget changes to a particular decision-making process. The level of analysis issue raises its head again. Budgets come about through inter-organizational processes, whereas decisional models such as Lindblom’s or Wildavsky’s apply to individuals or at most groups of individuals. Padgett (1980) showed how a boundedly rational search process compatible with successive limited comparisons could lead to non-Normal budget changes by adding an external constraint (such as available revenue) that had to be met.

My colleagues and I developed a different line of thought (True et al., 1999; Jones & Baumgartner, 2005). Based on our earlier work on punctuated equilibrium in policymaking (Baumgartner & Jones, 2009; first edition 1993), we postulated two distinct decision mechanisms in the budget process. One was typical incremental decision-making, dominated by negative feedback processes in which the range of movement is limited by the potential reactions against it. That is, in situations in which there is no agreement on goals, or agreement on goals but not on how to achieve them, or where goals involved difficult trade-offs, change was limited. The second was a process characterized by positive feedback, in which a policy change led to demands for even more changes. This happened when the focus of attention in a decision-making system shifted, perhaps because of a crisis, and it became clear that the prevailing understanding of a problem was flawed (Jones, 1994). In these circumstances, heuristics governing more typical situations, such as base and fair share, break down as attention focus shifts from the internal dynamics of mutual partisan adjustment (Lindblom, 1959) to addressing an externally-generated problem. Such shifts can account for the non-Normal budget distributions that undermined incrementalism.

That budget heuristics break down in limited circumstances does not make them irrelevant. In truth, modern budget theory has relied too heavily on budgetary data and not enough on process observation, so that we don’t know how extensive or exactly where in the complex budget process the breakdown is. We are reasonably sure that the observed budget distributions are not due to formal changes in budget accounting, as the observed distributions across time in a number of countries are similar in stochastic form (Jones et al., 2009). Bendor (2015, 2010) reminds us that the use of solution search heuristics in decision-making does not lead only to an incrementalist approach, and it is the use of heuristic solution search that was the central claim of the early bounded rationalists. Problem space redefinition obviously cannot lead to local search alone, as the incrementalists had implied (Bendor, 2010: Chapter 3), but that by no means rules out heuristic search as a mechanism for attaching solutions to the new problem understanding.

6.2. Path dependence and increasing returns to scale

Incrementalism has appeared through a different route in policy studies, one not related to decision-making but

---

2 The stochastic process approach used by Jones, Baumgartner, and True met considerable resistance among budget scholars, and their early studies were set up as time series regressions until 1999—almost 20 years after Padget’s breakthrough paper.
that often "lock in" tends to characterize incremental adjustment. Policies create beneficiaries of the stabilizing negative feedback that characterizes them. And the decision-making rooted in negative feedback for a brief period of time. The link between work of Brian Arthur (1994). The basic idea is that firms (and even industries) can experience increasing returns to scale, in which each marginal investment yields a larger return rather than the standard decreasing return. This is destabilizing because without some limits theoretically increasing returns could consume all the investment within a system, and in a very short time. So at some point stability must be restored within the industry and a more normal pattern of decreasing returns established.

Paul Pierson (2000) brought the notion of increasing returns and path dependency into the study of political science, integrating it within the framework of historical institutionalism. In the theory of punctuated equilibrium in policymaking, discussed above, bursts of positive feedback destabilize a policy subsystem and its incremental decision-making rooted in negative feedback for a brief period of time. The link between work of Arthur and Pierson and that in punctuated equilibrium is clear, but there is no microfoundation in the former. Clearly, however, in policymaking systems incrementalism must underpin the Arthur/Pierson notion of path dependency. And the decisional dynamics of budgeting and policy punctuations must also characterize them.

In either case, feedback from earlier policies is an important source of the stabilizing negative feedback that characterizes incremental adjustment. Policies create beneficiaries that often "lock in" support for the policies and make them hard to change (Campbell, 2003; Mettler, 2005; Pierson, 1993; Soss, 1999).

6.3. Integrating the problem space into models of decision-making

The realization that incremental decision-making was not adequate was associated with a broader understanding in the policy sciences: focusing on solution search alone was not enough to explain policy outcomes in decision-making systems. A full explanation would need to involve how problems are determined. It became increasingly clear in policy process studies that defining problems and searching for solutions are subject to different organizational dynamics.

This is the major lesson of the so-called "garbage can" theory of organizational decision-making (Cohen, March, & Olsen, 1972). The authors describe some organizations as capable of being viewed as collections of solutions seeking problems and decision situations in which they can be adopted. The connections between solutions and problems thus was attention-driven: only when the organization devoted collective attention to a problem could a solution be connected to the problem. Kingdon (1984) expanded the idea to governmental policy-making systems, and the idea that solutions and problems are not sequentially linked became a bedrock proposition in the study of how policymaking organizations set policy agendas. One must explain the unification of two agendas.

This organizational dynamic has an individual decision-making analog. Herbert Simon and his collaborator, cognitive scientist Allen Newell (Newell and Simon 1972), in their studies of human problem-solving had discovered something quite similar at the individual level. In particular, people resisted reexamining the problem-space—they treated it as set and relied on heuristic decision rules in searching for solutions ("trial and error" being most common).

Humans have a set of pre-packaged solutions to problems that are based in human biology and cultural transmission, so that when they encounter a familiar problem they respond virtually automatically. The same is true about policy making and implementing organizations. Following the basic rule of behavioral organization theory that organizations canalize human cognitive and emotional characteristics, organizations encode these pre-packaged solutions into bureaucratic decision rules and routines. These systems have been mapped for both budgeting (Crecine, 1969) and service delivery systems (Jones, 1980, 1985).

Problem search is enormously expensive compared to adopting a previously acquired solution automatically. But this relies on the decision-maker facing a recurring set of problems that are easily categorized and connected to pre-set solutions for a class of problems. So any organization faces a preparation-deliberation trade-off (Jones, 2001: 58). Prepared solutions lessen decision costs and they work most of the time. But if the problem-space changes, or is re-defined through the political process, then the pre-packaged solutions won't necessarily work. If the problem-space is redefined, two sources of uncertainty emerge: problem space re-definition and solution search based on a new understanding of the problem. As a consequence, response time increases and disjoint and episodic policy solutions are more likely as out-of-date heuristics are discarded and new ones developed (Jones, 1994).

Disjoint and episodic policy responses are fundamental to punctuated equilibrium in policy process studies (Baumgartner & Jones, 2009). Punctuated policies occur when problem-spaces are redefined. When problem-spaces are redefined, then new heuristics for solution search and implementation are necessary. While these occurrences are rare, they are disproportionately important in the policy process.
7. The allocation of attention

I’ve shown how the early models of decision-making based on bounded rationality seemed to have failed in the study of budgets; however, that apparent failure occurred because the existing models examined only the solution space, treating the problem space as fixed and hence subject to fixed heuristics. Then I discussed how the addition of problem-space analysis and the integration of a stochastic process approach into budget studies resolved the dilemma.

The big changes in policies that occur quickly are invariably associated with shifts in the allocation of collective attention in a policymaking system. These policy changes tend to occur within policy areas rather than across the board. Everything cannot get done at once. Even in periods of many major policy changes, such as during the Great Society, it is necessary to process policies serially. As a consequence, large changes cannot be explained by the parameter shifts that Davis et al. (1966, 1974) used in their attempts to salvage the incrementalist approach. But a single distribution of budget changes across policy areas, such as Padgett (1980) and Jones and Baumgartner (2005) display, shows clearly that in fact most decision-making is incremental.

The lesson here is that behavioral rationality can account for both major and incremental policy changes by integrating shifts in attention at the system level. We term the prioritization of problems in government its agenda, and major changes in policy are presaged by shifts in the agenda.

Simon (1987) pointed out that the decision rule of maximizing subjective utility carried a lot of unspecified baggage, or givens. In particular, he noted that “In a substantive theory of rationality, there is no place for a variable like focus of attention. But in a procedural theory [that is, a behavioral theory] it may be very important to know under what circumstances certain aspects of reality will be heeded and others ignored” (Simon, 1987: 31). Indeed, Newell (1958: 13) argued that any rule of decision-making, comprehensively rational or not, is conditioned on “what information is entered into the system... the judgment law is quite secondary, and amounts to doing the obvious with the information finally selected”.

Agenda-setting is the organizational analog of individual attention allocation and hence represents the choice about what information or aspects of reality are incorporated into the judgment rule. This is critical whether that rule be subjective utility maximization or ‘satisficing’ or some other way of taking information and making a decision from it. Moreover there is no part of behavioral rationality that is less in doubt than the close relationship between attention allocation and emotional arousal. Attention is selective, is hard-wired into human cognitive architecture, and requires serial processing of information. As a consequence, it is just not credible to believe that the choice of information to bring into the choice situation is governed by some rational analysis of costs and benefits regarding what information should be incorporated, if we had such a theory. We do not. The economics of information tells us how much information we should collect to solve a problem based on marginal expected returns (search for information until the cost of more search equals the expected returns on that search). As a normative guide, this just not helpful in a crisis when passions are running high. In any case, such a rule gives us no indication of the content of such information that decision-makers should search for. As a consequence, it is not an empirical proposition of any use in explaining policy outcomes.

Any complete theory of policy-making must incorporate an analysis of problem-space construction—termed issue definition in the field—and the allocation of attention to problems (problem prioritization) affect policy choices (Jones & Baumgartner, 2005). Not surprisingly, behavioral rationality is far more critical in explaining these phenomena than the judgment rule.

7.1. Information processing

In recent years, some policy scholars have attempted a more integrative approach, by thinking of policymaking systems as adaptive systems and hence as information processors (Jones, 2001; Jones & Baumgartner, 2012; Jones, Workman, & Jochim, 2009; Workman, 2015). In essence, any policymaking system adjusts to two general sets of factors: information coming into the system from its environment, and the system’s internal organizational structure. Because of the necessity of model simplification to bridge the micro-macro divide, we postulated a summary measure of the internal decision-making structure of the system: stick-slip friction (Jones & Baumgartner, 2012). This type of friction is common in the study of earthquakes; Bak (1997) highlighted a similar dynamic (1997). In a dynamical system characterized by stick-slip friction and input with a steady input stream, sudden, large-scale changes will occur sporadically. Adjustment to linear inputs will not necessarily be linear in such systems.

We proposed two sources of the friction in the system, which we termed cognitive friction and institutional friction (Jones & Baumgartner, 2005; Jones, Sulkin, & Larsen, 2003). Institutional friction was well-known to political scientists, who observe a substantial amount of what is called “gridlock” and a seeming inability to act even in the face of convincing information that a policy problem ought to be addressed. Cognitive friction captures a major element of behavioral rationality: the tendency to continue with an earlier set of decision rules. Partly this comes from the selective nature of attention: we don’t become aware of important information until long beyond when we should have acted. We argued that the system would react as the information accumulated, but would do so disproportionately by ignoring important aspects of the environment
until major changes were long overdue. Then the system would scramble to adjust. Finally we showed via simulations and empirical analysis how an error accumulation model, in which deviations from an ideal policy course of action would lead to dramatic policy shifts. 3

As I noted above, because the system adjusts to changes in its environment in a piecemeal fashion, much of the system may proceed undisturbed according to the prevailing set of decision rules. So any distribution of policy outputs across time will consist of incremental adjustments and large changes; few moderate changes occur. We hypothesized that because of this dynamic, the distribution of policy changes would approximate the Gutenberg-Richter Law, an empirical generalization that the size of earthquakes would be distributed as a power function. We found convincing evidence for that law in the study of budgets for the right tail (when budget changes are positive), in many budgeting systems. But the left side, representing budget cutbacks, more resembled an exponential decline, basically less severe (Jones et al., 2009).

We proposed the term disproportionate information processing to capture the pattern of under- and over-reactions that characterized policymaking systems (Jones, 2001; Jones & Baumgartner, 2005). Incoming stimuli cannot account for the distribution of policy outputs because of the internal dynamics of the policymaking system. Behavioral rationality has a very large role to play in this approach. Although it is not easy to rule out the hypothesis that the system of rules that must be satisfied to enact policy is the only source of friction, hence allowing for a rational explanation for the disproportionality, it is unlikely. First, observations and case studies of governmental decision-making offer scant support for a rational decision-making process, as we have reviewed in this paper. Second, friction apparently occurs in governments with different structural forms of government. Yet the US formal system, for example, based on divided and separated power, should exhibit considerably higher levels of friction than parliamentary systems.

The power of disproportionate information processing as a summary concept is that it is capable of linking micro cognitive processes of decision-makers to system-level outputs. In particular, we are able to behavioral rationality to patterns of policy outputs via the error accumulation model. Point predictions are not possible, but distributional ones are. That is, the model leads directly to the thesis that changes in policy outputs will be characterized by high kurtosis, with slender but very tall central peaks and fat tails. And it offers one potential avenue for connecting the cognitive capacities of human decision-makers to systemic policy outputs.

7.2. The second face of disproportionate information-processing

Disproportionate information processing implies that policymaking systems underreact to complex information streams. The error accumulation model implies that over-adjustment can occur, but only when the system has seriously under-responded previously. Yet positive feedback mechanisms that are central to complex systems (Erdi, 2008) would point to policymaking systems that can over-react via reinforcing complex feedback systems. Such feedback systems can lead to policy bubbles (Jones, Thomas, & Wolfe, 2014), in which a systematic over-investment in a policy instrument (that is, the method by which a policy’s goals are achieved) can occur. Another positive feedback mechanism that can lead to overreaction in policy processes is contagion, in which previously disconnected elements of a system become interconnected because of communication. Thomas (Cognitive Systems, this issue) is studying contagion through the use of agent-based simulation models. Contagion involves passing information signals from actor to actor, but it also invariably raises emotion strong enough to reinforce existing patterns of communication or even disrupt the normal hierarchical structures of organizations.

8. Contributions of other fields

While there has been substantial cross-fertilization among the social sciences on the issue of scaling up cognitive processes to the system level, progress has been slower than could be the case. I want to look briefly at two subfields that have contributed to the development of policy-centered models based on cognitive processes, and explore briefly why I think progress has not been as rapid as desired. The two fields are behavioral economics and political psychology.

8.1. Behavioral economics

Increasingly behavioral economics, based on experimental methods and direct observation of behavior rather than assumptions of rational economic behavior, has influenced thinking in economics. Because of the pre-eminent place that the fully rational model played in economics, much of behavioral literature in economics has produced a set of findings unified in its relevance for showing that the standard model is wrong. The stimulus for this work was provided by work by psychologists Kahneman and Tversky (1979), who explicitly addressed economic rationality in their work. Studies in behavioral economics became known as the “anomalies literature” and warranted a regular column in the Journal of Economic Perspectives, edited by the leading behavioral researcher, Richard Thaler (2015). All scientists trying (consciously or not) to upend prevailing understandings within a discipline run into considerable hostility, and few scientists were

---

3 The notion of error accumulation in budgeting was introduced by Pat Larkey (1979), and promptly ignored until we saw its relevance for the punctuated equilibrium model of budgeting and adopted it as a central feature.
more subject to this hostility than behavioral economists. One observer termed the adherence to the assumption of perfect rationality as a ‘cult’, and noted that “if the economy seems to be driven by behavior that isn’t quite rational, macroeconomists usually bend over backward to explain it as a failure of economic institutions rather than a result of human psychology” (Smith, 2016: 2). Yet behavioral economics has slowly and steadily chipped away at the rationality assumption.

The old issue of a reliable but not overly complicated model of what parts of the complications of human nature to use in explaining macro-phenomena emerges in economics as it did in political science. One problematic approach is a leap to the idea that humans are irrational in their economic behavior. Dan Ariely’s (2008) innovative research program, for example, perhaps jumps to quickly to that characterization.

A second approach, more generally relevant to the issue addressed here, are research programs in economics that try to link human cognitive and emotional capacities to macro-level phenomena. Akerlof and Shiller (2009) use five different aspects of human psychology to explain macroeconomic phenomena. While I applaud this ambitious move, the choice of the aspects—confidence, fairness, corruption and anti-social behavior, money illusion, and stories-strike me as somewhat strange. Some of the aspects are fundamental to human psychology, such as confidence and fairness. Others seem to rest on more fundamental psychological dynamics—money illusion being the most glaring. Stories—or narratives—can be based on more fundamental aspects of human psychology and integrated into a robust research program in belief systems (Jones, Shanahan, & McBeth, 2014).

Separately, Shiller (2014) has demonstrated behaviors in equity markets and in the US real estate market in which a purely rational model fails to explain macro-level outcomes. He received the Nobel Prize in economics for his work on the long-run predictability of equity markets, a finding at odds with the classical efficient market thesis that claimed markets were not predictable because they were so efficient in the utilization of information in decision-making). His empirical and theoretical contributions to the study of real estate have additionally contributed to our understanding of “non-rational” economic bubbles (Shiller, 2000).4

So we can see two forms of behavioral economics. One centers on experimental methods that illuminate aspects of economic decision-making where the potential for a divergence in behaviors predicted by the standard rational model exist. But scaling up these findings has proved difficult, at best explaining only limited aspects of macroeconomics. On the other hand, a second approach represented by Akerlof and Shiller, has mapped macro-level phenomena that seem inexplicable by standard approaches. But these approaches suffer from difficulties in “scaling down” to the cognitive and emotional elements responsible for the divergence, over-relying on plausibility arguments to support the connection.

Behavioral economics has also made a more practical contribution to public policy studies. Some quite simple “tweaks” or “nudges” on the way policy is delivered can yield better results, whether the policy is delivered in the private or public sectors (Thaler & Sunstein, 2008). Interestingly these devices mostly can be traced back to the initial formulation of bounded rationality by Simon, because they are mostly based on selective attention and the use of heuristics. Such nudges would not be needed under a rational model.

The major issue that I’ve addressed here, the linkage between the micro-processes of cognition and the macro-phenomena that social scientists study is quite in evidence in economics as it struggles to integrate behavioral approaches into the standard approaches. But some economists have begun the difficult task of unifying more solidly behavioral science with macroeconomics. Xavier Xavier Gabaix (2016: 1) is developing a “framework for analyzing how bounded rationality influences monetary and fiscal policy”. Compared to traditional models, he reports substantially different implications for policymaking. The major difficulty in such exercises remains choosing what aspects of behavioral rationality to choose to model, and the choice made by the researcher can lead to different results. Of course the proof in in the pudding—how close do the empirical results match the model’s prediction. One of the major differences between Gabaix’s models and traditional models is the major role that attention—and inattention—plays (7).

8.2. Contributions of political psychology

Political psychology developed from the study of voting behavior and public opinion. While it has been recognized as a separate area of specialization in political science for a quarter of a century or so, its roots go back to the 1950s. Most explanations of voting behavior back then were based in sociological analyses—age cohorts, group memberships, and region of the country, for example. But as polling techniques improved, political scientists began to integrate cognitive and emotional components of individual voters into their studies. Most of these studies focused on attitudes, but over time as the field developed scholars began to examine psychological dynamics that were likely to affect voting choice and other elements of political participation.

The field of political psychology is diverse and robust, but with few exceptions there has been little scholarship that has been explicitly linked to macro-level phenomena. The biggest general exception is voting patterns, where scholars have explored these linkages. While it is hard to

---

4 There really are formal theories based on rationality of why bubbles should occur. A barebones rational model of economic action leads to a very indirect and complex explanation of bubbles, whereas a behavioral approach is far more simple, intuitive, and direct.
characterize as large a subfield as political psychology, it is fair to say that most of the literature has emphasized the non-rational aspects of human behavior, shying away from behavioral rationality and sometimes viewing behavioral rationality as close to comprehensive rationality. One consequence has been little in the way of connectivity with public policy process studies, where there could be a productive interchange.

One area where there have been solid explorations of the role of behavioral rationality and policy choices is in the field of international relations. Most prominent among these are integrating Kahneman and Tversky’s prospect theory into the field (McDermott, 2004). A major premise of prospect theory is that the status quo point from which the decision starts matters. If people view themselves as sustaining earlier losses, they are more likely to take risky action—a laboratory finding with serious implications for interactions among world leaders.

A second important area within international relations involves the manipulating of attributes that characterize choice structure in simulations mimicking international crises (Mintz, Geva, Redd, & Carnes, 1997). These approaches have great potential for linking to macro-level phenomena, but so far they have mostly been used to illuminate specific decisions or crises by providing case studies of choices and making arguments about the plausibility of the specific underlying cognitive dynamics. While this approach has been extremely productive, it has not addressed our central issue systematically.

9. Summing up and bridging the gap

Integrating a better (in the sense of more scientific) model of human nature in the macro-level phenomena that the social sciences generally address is a key challenge in these disciplines in the future. The standard approach has been to assume comprehensive rationality and use the mathematics of optimization to study systemic processes. Great breakthroughs in our understandings of human societies have resulted, but there have been problematic limitations.

It is highly unsatisfying to base system-level models on a micro-model of behavior that is so demonstratively wrong. But it is more than simply wrong: it is misleading in key facets. There are two elements of any model of decision-making that is capable of yielding insights into such macro-level phenomena as general policy processes. The first is the rule: what is the rule for choosing among the potential alternatives in a choice situation. The second is the information fed into the decision rule: what parts of a complex environment are chosen as relevant. Much empirical evidence in psychology, behavioral economics, and political psychology indicate serious problems in this assumption—enough that the continued justification for continuing to use it is ringing increasingly hollow.

Perhaps even worse, the rule is disastrously incomplete. All choice situations are not equally structured, yet applying a rule to alternatives relies on a set of assumptions that transform all choices into fully structured ones. Alternatives (or solutions) are known, problems are well-understood, and the ends-means relationships are understood (at least as a probability distribution). As Simon (1987) noted, in the rational model these are “givens”. Yet few problems are that simple, and changes in the givens in a choice situation can change dramatically the choices that people make.

In my field of policy process studies, we’ve made great strides in understanding how policies are made and with what impact by integrating a few principles from behavioral understandings of cognitive processes. This allows simplification enough to begin to scale upward from the decision-making process to the policymaking system. These are at a minimum:

1. Actors are behaviorally rational—they intend to be goal oriented and strategic, but they make mistakes—mistakes that are not just a consequence of a complex and dynamic reality, but are also a consequence of human cognitive architecture.

2. Different dynamics can influence the problem space—understanding of the problem faced—and the solution space—the search and specification of alternatives to address that problem. Elements critical to the democratic process, such as public opinion and freedom of expression, are far more critical in problem specification than solution search. Often, but not always, experts are better at solution search (Baumgartner and Jones, 2015).

3. The role of selective attention is critical in both problem definition and solution search. No policy model is anywhere close to complete without an appreciation of the role of attention in structuring choice in complex environments—and this includes inattention.

4. Emotional arousal is a dual-edged sword—it is critical in shifting the focus of attention, but it is also an essential component of cognition and hence reason (Damasio, 1994). But it can interfere with even intended rationality in evaluating the best policy alternatives to achieve a goal. Schneider and Ingram (1993) show that our emotional evaluations of the target populations can strongly influence how much coercion is used in the policy instrument.

In my opinion, we are at a turning point in the way we approach micro-macro linkages—how much or how little of the cognitive and emotional constitutions of humans we bring to our understanding of the operation of macro-level processes. Few social scientists have any faith that the rational model can take us any farther than it has. But we have not thought deeply enough about what elements are necessary and which are expendable. There is much interesting but ad-hoc research in behavioral political science and economics, with interesting but disparate findings. Often these findings are useful for policy design and implementation, but it is high time that we harness this
capacity more directly on addressing how our policymaking systems actually run. For the most part we have failed to design experiments or collect observational data (the later is much underrated as a scientific tool) that is directed at the key questions of linking behavioral rationality with the dynamics of policy processes. It is that task that I hope will attract the attention of both social scientists and cognitive scientists; it is a big task that needs serious interdisciplinary work.

References


