The Transformation of Policy Ideas
Author(s): Jeffrey W. Legro
Published by: Midwest Political Science Association
Stable URL: http://www.jstor.org/stable/2669256
Accessed: 16/05/2011 13:50

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=mpsa.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.
The Transformation of Policy Ideas

Jeffrey W. Legro  University of Virginia

A wave of recent scholarship, embracing both rational choice and cultural approaches, highlights the importance of collective ideas in explaining politics. What this work has not adequately addressed, however, is why collective ideas sometimes radically change. Scholars typically attribute ideational change ad hoc to creative entrepreneurial agents or to exogenous shock. But these solutions fail to account for why some entrepreneurs and not others are successful in marketing their ideas or why similar shocks lead to change in one instance but continuity in another. Explaining these variations is aided by conceptualizing ideational change as a two-stage process involving collapse and consolidation. Collapse involves societal actors reaching an agreement that the old orthodoxy is inadequate. Consolidation requires social coordination on a replacement set of ideas. Both steps involve collective ideation problems where actors may have difficulty coordinating a change in dominant ideas or have incentives to shirk facilitating efficient adaptation. The outcome of these two steps is in part endogenous to ideational structure itself. Collective ideas generate expectations, which in particular ways interact with experienced events, either inhibiting change or allowing societies to transform their own dominant notions. The approach is illustrated with examples involving collectively held causal ideas in a variety of foreign policy domains.

Across a variety of literatures, scholars have taken a renewed interest in the role of collectively held ideas and beliefs in politics. Rational choice scholars have turned to collective ideas both to explain how societies make decisions under uncertainty and to resolve multiple equilibria problems (Kreps 1990; Ferejohn 1991; Garrett and Weingast 1993). A second group of scholars working more closely in a sociological/cultural framework have rediscovered and refined the explanatory power of ideas in a variety of issue areas including macroeconomic policy, trade, industrial strategy, national security, and internal security (Hall 1989; Sikkink 1991; Dobbin 1994; Katzenstein 1996a,b). But what both groups have not explained adequately is why enduring dominant ideas sometimes radically shift. What is necessary, but missing, is an explanation of discontinuous change in collective ideas. Explaining such change is a concern well beyond the study of politics as seen in the debates in the history and philosophy of science literature (Andersson 1994). In this article, however, I address the problem of ideational change with particular reference to policy concepts, especially foreign policy concepts held by societies we know as nation states. Why do collective ideas sometimes undergo dramatic transformation, yet at other points show basic continuity?

The array of extant answers to this question is usefully examined within the agent-structure debate that has shaped economics, sociology, and political science. While each of the three main positions in this debate—agentic, structuralist, and structurationist—all provide insights, each addresses the issue of change problematically. The standard solution to their pitfalls is to invoke “exogenous shock” as the source of change—but this factor is only recognized post hoc because a change has occurred. Instead what is needed is a better-specified notion of how events relate to the causal mechanisms of the three positions, either in support of ideational continuity or as a source of change.

Jeffrey W. Legro is Associate Professor of Government and Foreign Affairs, University of Virginia, 232 Cabell Hall, Charlottesville, VA 22901 (legro@virginia.edu).

For helpful comments on earlier drafts, my thanks to Dale Copeland, Peter Katzenstein, Robert Keohane, John Owen, Alexander Wendt, the participants of the Workshop on Rationalism and Constructivism at Duke University, and the anonymous reviewers of AJPS. I am grateful to the John M. Olin Institute for Strategic Studies, Harvard University, and the University of Virginia for financial support.

1 The aim is to explain relatively rapid transformations of ideational structure, not gradual alterations within the logic of a dominant idea or small changes that over long periods fundamentally alter the original structure.


©2000 by the Midwest Political Science Association
This task requires an appropriate conceptualization of ideational change. As opposed to a single phenomenon, such change involves two distinct analytical stages (which in practice are often difficult to disentangle). First, societal actors must somehow concur, explicitly or tacitly, that the old ideational structure is inadequate thus causing its collapse. Second, actors must consolidate some new replacement set of ideas, lest they return to the old simply as a default mechanism. Both of these steps are challenged by collective ideation problems that make coordination difficult and/or give individuals incentives to shirk efforts to challenge dominant beliefs. At a basic level, collective ideational change is constituted by a shift in the ideas of individuals and some process by which individual notions aggregate to form a new dominant social idea. As discussed below, however, while change in collective ideas may involve individual cognitive change, it must involve aggregation. Even though human cognition is inevitably involved, this article focuses mainly on the aggregation part of the puzzle: why do societies effectively transcend collapse and consolidation barriers in some cases, leading to change, but in others do not—thus reproducing the dominant orthodoxy?

The answer, I argue, depends on the role of ideational structure in its own transformation—specifically the interaction of ideationally shaped societal expectations and the consequences of experienced events. When events generate consequences for societies that deviate from their collective expectations and the consequences are starkly undesirable, the ideational and aggregation components necessary for change become more likely. By shaping expectations, collective ideas have an irreducible role in change. Thus similar societies with different dominant ideas will evolve differently in reaction to the same experiences or environmental pressures. This model of ideational change brings together individual and structural elements of change in ways distinct from purely rational adaptation or psychological learning formulations. Furthermore it addresses two outstanding anomalies in social life: why similar shocks cause ideational change in some instances but not others and when it is an idea entrepreneurs are likely to succeed.

In what follows, I first discuss what I mean by collective ideas. The second section considers the literature on ideational change and its weaknesses. I then develop an explanation of change to address those gaps. The logic of this approach is illustrated with examples from a number of foreign policy areas such as national security, trade, nuclear weapons safety, and grand strategy. Finally I discuss the implications of the argument.

**On Collective Ideas**

“Collective ideas” refers to concepts or beliefs held by groups (i.e., states). These ideas are social and holistic—they are not simply individual conceptions that are shared or added together. Collective ideas have an intersubjective existence that stands above individual minds and is typically embodied in symbols, discourse, and institutions. Individuals and their interaction naturally influence collective ideas, but they also must confront them as “fact” (Durkheim 1964; Searle 1995, 23–26). Collective ideas, therefore, cannot be reduced to individual ideas, belief systems, cognition, or psychology, even if such phenomena related to the human mind may often be a critical part of collective change.

The potential disconnect between individual ideas and collective ideas is evident in a number of situations where an individual might accept, adhere to, and even promote collective beliefs she does not personally believe. For example, individuals may accept an orthodoxy for society that few personally support in order to avoid a debilitating standoff. A dominant orthodoxy may also be accepted for reasons of social conformity. Even when a majority of individuals do not accept the dominant orthodoxy, they may not realize such widespread sentiment exists and be unwilling to pay the costs of publicly deviating from a perceived accepted group norm (Kuran 1995). Collective views may also endure despite their tenuous nature because individuals hope they will have a self-confirming aspect to them, that if everyone believes them, they might be true. Thus an army command or coaching staff may privately see little hope of victory, but may foster the group belief in victory so as to increase the chance of it happening (Gilbert 1996, 195–216).

Clearly the phrase “collective idea” does not imply a monolithic homogeneous entity that all societal actors internalize and advocate. It is unlikely that all individuals or societal groups will share or agree with the dominant orthodoxy in any particular issue area; one can always find opposing individuals and subgroups. Differing attitudes towards traits (e.g., language or religion) that characterize communities are common. Nonetheless, to act coherently, groups (e.g., states in foreign policy) require dominant themes for reasons described both by economic and sociological approaches. In economic terms, a dominant set of ideas helps to avoid problems linked to preference cycling, transactions costs, and free riding (North 1981). In social-psychological terms, a collective identity requires collective ideas (Tajfel 1981).

Because collective idea is a broad term that can refer to a variety of phenomena, I limit my focus further to the
concepts societies consider appropriate to achieve their interests in specific issue areas. This general formulation is found in various guises. Goldstein focuses on “beliefs about the efficacy of particular strategies for obtaining objectives” (1993, 9). Hall usefully distinguishes the notion of “policy paradigm” as a “framework of ideas and standards that specifies not only the goals of policy and the kind of instruments that can be used to attain them, but also the very nature of the problems they are meant to be addressing” (1993, 279). Ellingson uses the term “discourse” similarly to refer to a “set of arguments organized around a specific diagnosis of and solution to some social problem” (1995, 107). Finally, Dobbin refers to “industrial cultures” as “economic customs that structure...the means-ends designations” (1994, 5). Hence, it is possible and useful to talk about a collective idea in reference to an enduring dominant mode of societal thought on appropriate action in a particular issue domain, even if other often-contradictory views are sure to exist. Despite such endurance, abrupt change in dominant ideas does occur. Why?

Gaps in the Study of Collective Ideas

A necessary starting point for addressing ideational change is the array of extant explanations. These can be usefully grouped in terms of the agent-structure debate (Wendt 1987), which has produced three problematic approaches to ideational change. All three deserve attention because their limitations point to useful directions for conceptual improvements.

Agency

An agency approach explains collective ideas as the product of the views of some dominant or enterprising agent or an aggregation of many individuals’ views. From this perspective, change in collective ideas are a matter of understanding how the relative power, interests, or beliefs of individuals or smaller groups within society shift, ultimately affecting collective outcomes. Any notion of ideational structure as being relevant is subsumed by a bottom-up actor-centered momentum. “Collective” here is just a label for what is in the heads of societal actors.

A focus on agency is an important contemporary theme in work on social ideas from a number of different perspectives. One view comes from rational choice theorists who aim to explain collective ideational phenomena through the choices of optimizing self-interested actors (Kreps 1990; Young 1996). Closely related are interest group and social activist approaches which highlight the influence of subgroups seeking change (Frieden, 1988; McAdam, McCarthy, and Zald 1996). Another view from psychology considers the personality traits or cognitive structures of individuals in explaining collective ideas (Jervis 1976). The central problem with these three variants is that by attributing change to agency they can account for neither the collective properties of ideas nor ideational inertia since the aggregation of heterogeneous individual notions would seem to suggest perpetual transformation (cf. Arrow 1951). These problems are evident in the literatures on games, social movements, and the psychology of decision making.

In terms of game theory, an individualistic approach has not, for the purposes of understanding the problem at hand, resolved the quandary that strategic interaction among societal subgroups can generate no, multiple, or unstable equilibria. Theorists have usefully highlighted the way that preexisting structure can “induce” an equilibrium (Shepsle 1989, 135). For example, a number of scholars have turned to collective ideas to explain the consolidation of particular focal points (Schelling 1960; Ferejohn 1991; Garret and Weingast 1993). In doing so, however, they move beyond properties of agents to properties of structure which remain exogenous, yet demand explanation. What constitutes one structure (set of beliefs) as the decisive focal point at one time but not at another?

One resolution to this question is to attribute the dominance of a particular concept to the desires or ideas of powerful actors (Shepsle 1989, 137). The problem, however, is that determining power independent of what actually occurred is often elusive: ex-ante power (unless defined tautologically as “influence”) is not always clearly related to ex-post results and rarely so when change occurs. This issue is clear in the social movement literature that explicitly aims to understand how seemingly weak interest groups are able to prevail by instrumentally generating and spreading persuasive ideas. Social movement theory has moved away from an instrumental agency focus to consider how certain weak ideas become powerful because of the way they are framed vis-à-vis broader ideational structures (McAdam, McCarthy, and Zald 1996). What social movement scholars have not resolved is how to recognize a successful frame without knowing the outcome.

---

2Such concepts can be distinguished from a second type which “are generally more homogeneous, affective, impermeable to experience or contradiction, and leave little scope for individual variations” such as many religious beliefs (Moscovici 1998, 226).
Mackie’s (1996) provocative study of footbinding (in China) and female genital mutilation (FGM; in North Africa) seems to answer this complaint by suggesting that a necessary condition for change is the Pareto superiority of the replacement set of beliefs. This argument, a type of functionalism, suggests that if a new set of ideas can provide the benefits of the old at the same or even lower cost, it will be adopted. Thus he argues that footbinding and FGM, functionally tied to the goal of successful marriage, are vulnerable to change (i.e., they are suboptimal) because marriage contracts can still work without physical maiming. But the problem is societies do not always react to what outside observers take to be optimal. Indeed ideational change has only occurred in one of Mackie’s cases—footbinding in China. In explaining the absence of change in FGM in North Africa, Mackie suggests that agents of change were more skilled in the China case. Here, as is so frequently seen in the social movement literature, the success of change agents is linked to their unexplained ability to persuade. I argue below that there are other systematic conditions that help determine whether certain ideas are socially persuasive or not.

The move from individual to collective also challenges the psychological approach to social ideas. Scholars have made great progress in explaining key elements of individual cognition, but in choice situations involving individuals with different beliefs, how these individual traits aggregate to explain collective ideas is less clear. One example comes from Khong (1992) who argues that state decisions are molded by analogies that perform a variety of cognitive functions. Khong contends that American policymakers were predisposed to intervene in Vietnam due to the Korean and Munich analogies evoked by the situation. In Khong’s account, analogies relate to schema held by individual leaders. But why certain analogies resonate for the collective (i.e., across many different individuals with different schemas) at a certain time demands explanation. Khong attributes such agreement to a generational effect based on lessons learned from crucial political events. But this moves the explanation to intersubjective beliefs and the ad hoc way they do or do not change in relation to shocks, a topic we return to below.3 In sum, a one-sided focus on either rational or psychological agency begs the question of how heterogeneous individual preferences and/or beliefs aggregate.

3 Of course individuals with different ideas may be profoundly affected by some event that leads them to separately arrive at accepting the same idea (e.g., Schuman and Rieger 1992). But this is likely to be a social phenomenon as well since one would expect differing anomie individuals to draw different lessons from the same event.

A second view sees ideas themselves as a type of structure. Here agents are vessels filled from the structural trough—they are cultural mannequins. As one scholar notes, “most of the time, in most places, most people simply habitually reproduce the prevailing pattern of social life” (Hays 1994, 63). It is widely recognized that much of what humans learn is not through individual trial and error, but instead through socialization and imitation (Tomasello et al. 1993). From this structural culture view, extant collective and individual ideas are constituted by preexisting ideas. But from this view, there is no reason why collective ideas, as an inertial structure that lacks the capacity for cognitive evaluation, should ever change.

This is a common position in the recent literature that has focused on the role of collectively held ideas, beliefs, and norms to explain why societies continue to behave in traditional ways despite changes in their environments. For example, Katzenstein (1996b) and Berger (1998) argue that Japan and Germany are characterized by different dominant ideas about how to manage foreign affairs that emerged from a process of political contestation following World War II. Subsequently these ideas have guided their distinct policies (given their power position in the international arena) towards national security. Even as the international situation has changed dramatically in the wake of the Cold War, Katzenstein (1996b, 204 and following) and Berger (1998, 198–201) anticipate basic continuity from these two countries. To be sure, neither scholar expects such ideas to last forever; both are insightful on the ongoing political contestation over dominant ideas. But overall their analyses highlight the influence of a continuous ideational structure, not why it changes.4

Some argue that ideational structures contain their own logic of transformation (Levi-Strauss 1963). A structure might have (1) a preexisting formula for change, (2) contradictions that induce instability and change, or (3) encounters with unforeseen consequences that eventually undermine them.5 But in each of these instances either ideational transformation has not occurred or there seems to be a need to reference elements outside of structure. In the first case, ideas are not transformed, they are enacted, since the preexisting orthodoxy

4 This same bias exists in my own work. See Legro (1995, 24–25).

5 Both Sewell (1992) and Archer (1996) usefully speak to the structural level dynamics that may facilitate change. But neither of these authors argues for a strictly structural account as I use that term here.
included its own terms of alteration. In the second and third instances, factors external to the ideas are clearly important in specifying why an ideational structure with internal contradictions changes at one time and not another. If the contradiction is only realized in particular circumstances, we must ask what they are and who or what recognizes them and enacts change. There appears to be an inescapable need to reference environmental conditions, human thought, or both. As self-reflective creatures, people can respond to new information and alter their own desires and beliefs providing the impetus for change. An autonomous structure argument is therefore insufficient.

**Structuration**

The third approach accords neither structure nor agency autonomy. Instead this structurationist account sees the two entities as relational—they are “mutually constituted or codetermined entities.” The mutual constitution among agents and structures both produces and reproduces those entities (Giddens 1979). The key issue of course is whether reproduction (stasis) or new production (change) will occur. Since the two basic elements, structure and agency, cannot be distinguished, little can be said about their causal interaction or when one or the other will dominate. “Cultural change thus becomes an inmanent but indeterminate possibility, equally likely or unlikely at any given moment and therefore unpredictable and inexplicable” (Archer 1996, xxv).

There is a solution to this pitfall, but one that has taken a questionable form. Authors of different stripes seem to recognize that while structures and agents are ontologically inseparable, one can deal with them analytically as if they are distinct. Wendt argues for “bracketing”—“taking social structures and agents in turn as temporarily given in order to examine the explanatory effects of the other” (1987, 364–365). Taylor (1989, 118) agrees suggesting analytical separation is central to understanding the causal interaction of structure and agency and explaining social change. Over time, the relative effects of agency and structure (even as they are mutually constituted) can be disentangled.

While useful, this approach has evolved with a distinct bias. Scholars tend to divvy up explanations of social stasis and change in a set fashion: structure is used to explain constancy of social phenomena whereas agency is cited as the source of change (Taylor 1989, 121–122). Structure is about the perseverance of a particular configuration of parts and their constraining or constituting effects. It is about the enduring, not the random or ephemeral, elements of historical circumstances and/or social relations. In contrast, agents (especially human ones) are seen as the calculating and creative entrepreneurs ever-willing to innovate and transform the existing order. But this division of labor slights the way structural conditions facilitate change and the way agents thwart transformation.

Work by Snyder (1991) and Kupchan (1994) illustrates this solution and exposes its related problems. Both authors depict ideational change in foreign policy as a product of elites instrumentally manipulating public images and ideas in order to build support for their policies. At a later time, however, elites become constrained by the structure of collective ideas they themselves have developed. Such ideas become institutionalized and are difficult to change, even if elites have not internalized their own message, because of collective action problems in affecting change and the political costs of reneging on one’s position. Strategic ideas, in effect, take on a life of their own. The gap in these otherwise insightful analyses is that they give little attention to when and where the inertial imperatives of preexisting ideational structure versus the transformational power of agents will decide outcomes. Elites are able to shape societal beliefs at one point but at a later point are no longer able to manipulate them. But why were they able to change beliefs in the first place if ideational structures capture actors? When can extant collective ideas be expected to constrain actors and when can actors reconfigure ideational structure itself?

The most common response to such questions is to invoke “exogenous shock” as the motor of change (e.g., Higgs 1987; Goldstone 1991; Khong 1992, 35; Berger 1998, 12–15, 207). These arguments typically portray a *stasis-shock-change* pattern of conceptual development, one that is familiar in the literature on institutional change as well (e.g., see Olson 1982). While crises are often related to change, exogenous shock remains an indeterminate explanation. Similar shocks seem to have different effects: some lead to change, some do not. Why? We have a poor understanding of the causal mechanisms that link shock to change. This is important because some scholars argue that actors confronting a crisis are likely, not to innovate, but to retreat to even greater reliance on preexisting patterns (Staw, Sandelands, and Dutton 1981). The implicit answer is that certain events somehow show extant beliefs to be wrongheaded and that change is therefore unavoidable. But it is not clear which events are likely to do this. Comparable wars or economic crises often seem to have differential effects.

One way to explain this difference is to heed what Sahlins (1991, 43–44) calls the structure of meaning that defines an event as significant. As an analogy, similar earthquakes can have radically different implications.
based on the nature of the buildings (or lack thereof) in the area where they occur. Likewise similar events can be understood by, and effect, societies characterized by differing “mental buildings” (i.e., collective ideas) in different ways. What is needed is an approach that incorporates the impact of events, but in the context of extant societal thinking. It is the interplay between these two that opens up a space for reasoning agents to perceive the need or opportunity for change and to successfully organize to overcome extant constraints. In what follows, I develop a synthetic framework that recognizes the recursive nature of these relationships—one that gives codetermining weight to collective ideas, but also allows for the influence of events and the instrumental efforts of calculating agents. What I try to do then is to provide micro-foundations to a structural approach to ideational change (but this task might also be viewed as establishing the structural conditions of an agency-based explanation). Overall the aim is to specify better the likely conditions in which we can expect collective ideational transformation or continuity.

Explaining Change in Collective Ideas

Change in collective ideas can involve both shifts in the notions of individuals that constitute societies and the social aggregation of such heterogeneous views into an enduring collective perspective. Because there is no physical group mind, aggregate conceptual change must necessarily relate in some manner to the thinking and/or actions of individuals or subgroups. Appropriately, substantial work has been done on why individuals change their ideas (Tetlock 1998a). But individual attitude change is not a sufficient or even necessary cause for explaining the particular timing and circumstances of change in collective ideas. As noted above, even when the majority of individuals privately believe otherwise (suggesting the need for change), the extant collective orthodoxy may endure for a variety of reasons including individual ignorance of others’ own reassessment of what is proper and desirable, fear of social ostracism for challenging group beliefs, or a failure to agree on a new ideational structure. Conversely, even when individual ideas remain constant, collective outcomes may change simply based on differing aggregation mechanisms. Similar to Arrow’s (1951) logic on preference formation, changing the order of idea articulation can change the collective outcome.

In what follows, therefore, I focus more on the social aspects than the individual cognitive aspect of collective idea change. Change in societal-level beliefs requires “collective ideation”: the formation of group ideas implies a need for coordination and/or collective action to facilitate change. Individual actors may lack motive to allocate cognitive and other resources to alter their understanding and/or to incur the burden of acting on such personal beliefs when doing so exceeds the expected benefits. Motivated actors, moreover, may find it difficult to ascertain the specific ideas of the large numbers of individuals involved or to coordinate their efforts.

These collective ideation problems can appear in each of the two ideal stages that together constitute ideational change. The first stage involves the collapse of the reigning extant consensus where actors are able to collectively agree that the old orthodoxy is inadequate and should be replaced. This is no simple task because such enduring ideas are likely to be institutionally embedded and have a cohort of hard core supporters. Second, ideational change also requires the consolidation of a new ideational structure, again inviting aggregation problems. Actors may not be able to coordinate because of a lack of information on what outcome is acceptable to others and/or because they disagree on which set of beliefs is most desirable. Reform or revolutionary movements that may bond in opposition to something frequently fall out when that something disappears and conflict breaks out over a replacement (Burns 1996, 350–351). Thus, even when ideational collapse occurs, failure to reach a consensus on a replacement could still produce continuity as society reflexively reembraces the old orthodoxy. What then are the factors or conditions that allow some societies to overcome collapse and consolidation and effect ideational change, while at other times ideational continuity prevails?

Collapse

Collapse of an extant orthodoxy seems driven by the interaction of (1) collective expectations (generated by collective ideas) and (2) the experienced consequences of critical events. First is the distinction between fulfilled expectations and unfulfilled expectations. Idealational prescriptions carry a set of social expectations of what should or should not result from group action. These expectations become salient vis-à-vis events. When expectations of what should happen are not matched by the consequences of experienced events, there is pressure for collective reflection and reassessment (Levy 1994, 305). One might expect cognitive biases (e.g., dissonance reduction, selective attention to evidence, attributional pathologies) to mitigate the effects of such discrepant information (Jervis 1976, especially 143–145, 288–315; Eagly
and Chaiken 1993, 559–625; Tetlock 1998b). But such inertial human mental habits are partly contained by the public nature of significant gaps between expectations and results. One can more easily rationalize personal contradiction-prone excuses. This becomes more difficult to do publicly, especially when at least some others have different (and critical) views. Unfulfilled expectations, however, are only part of the picture.

A second critical factor involves consequences and the distinction between success and failure. Extant collective ideas contain not only a notion of appropriate action, but also a portrayal of what consequences are a success (or socially-approved) versus a failure or (socially-stigmatized). This distinction is important because a range of studies indicate that failure as opposed to its opposite success is associated with a change in collective mindset (Perrow 1984; Arnold 1990, 51; Hall 1993, 280; Dobbin 1994). The standard proposition is that success contributes to policy continuity whereas failure leads to change.

But this blanket thesis needs to be modified because it neglects the key interaction with social expectations. When failure is accurately expected by ideational beliefs it is unlikely to result in the transformation of those beliefs. Furthermore, there is a difference between unexpected failure and unexpected success. Both instances involve unfulfilled expectations, but in the former case cognitive reflection is more likely. People are more sensitive to losing something that they expect to have than they are to gaining something they did not expect to have (Kahneman and Tversky 1979). The former inspires more intense preferences than the latter and likewise more intense scrutiny.

This asymmetry is relevant to collective ideas because it has social implications. Unfulfilled collective expectations that result in undesired consequences (losses) help to spark the type of social interaction that can resolve collective ideation problems. This interaction gives actors more incentive to reveal and act on their ideas, thus making it more likely (that the often unexplained impetus for) tipping towards a critical mass or threshold of collective action in favor of change is initiated. Undesirable consequences raise the salience of an event, encouraging individuals to rally around it as an issue deserving attention. These events allow individuals to confidently overcome “do you see what I see” concerns resulting from the pressures of social conformity. Failure necessarily generates a social tendency to investigate and assign blame. People want to find out what went wrong and why. Nonevents tend to be ignored. Unexpected successes do not draw critical attention for reasons captured in the folk saying, “if it ain’t broke, don’t fix it.” Failures, in contrast, give critics of the dominant orthodoxy intersubjective, socially potent, evidence to argue for the replacement of extant ideas. In such circumstances, the social silencing of critics is diffused and consensus on collapse easier to achieve.

Hence, we should expect situations involving the combination of unmet expectations and undesired consequences to be more likely to facilitate collective ideation than those where expectations are fulfilled or desired consequences occur. In part such conditions do so because they are likely to make individual’s ideas prone to change, but more importantly they do so because they make social consensus on change more likely. The more significant the contradiction between expectations and situations, the more likely societies will face contestation regarding supposedly efficient and appropriate means, and the more probable it is collapse will take place.

Consolidation

The consolidation of a new dominant orthodoxy is more open-ended than the collapse stage. For good reason, scholars often depict the consolidation of a new ideational orthodoxy as a highly contingent political battle (see Hall 1989; Katzenstein 1996b). But again, attention to ideational structure may help to narrow this indeterminacy and shed light on conditions that facilitate new consensus. Like collapse, consolidation is importantly affected by the structure of ideas vis-à-vis events. As noted above, ideational structures are rarely monolithic. Usually they contain a hierarchy of ideational elements with one creed dominant and at least one main challenger. The result is often a binary axis in the ideational structure (Levi-Strauss 1966; Kane 1991, 56). Hence new ideas do not emerge full-blown out of the ether, but instead take shape within extant structures.

The character of the opposing idea and its fit with events can importantly influence the likelihood of change. Since there are always ideas available the key issue is their social plausibility: does the opposing idea have a social base—a cohort of advocates—prior to an event that confirms the opposing idea’s, but not the

---

6For example, military victory might objectively be coded as success, but whether a society sees it as such depends on the ideational framework through which it is understood—e.g., U.S. disillusionment after WWI.

7Sometimes such demands will be thwarted by the claim that a given idea did not receive a fair test. Jervis (1976, 275n) argues this is likely when the original values are deeply rooted, when a policy is implemented half-heartedly, or there are unique circumstances in the original case that would not affect future cases.
dominant idea’s expectations? Such a dynamic gives its advocates persuasion power: they not only criticize, but offer a socially salient solution that will serve as a new focal point for collective ideation. Moreover, as Jervis writes, “success is apt to consolidate the power of those who advocated the policy, defeat to undermine it and strengthen the hand of those who had different views” (1976, 238). Any initial success, whether it is a product of the ideas or not, will help to solidify the new orthodoxy, while unexpected negative results will lead to continuation of the consolidation struggle or even a default return to the old. This ideational level dynamic, in conjunction with events, gives voice and power to previously marginalized concepts and their supporters. It will not always tell us exactly which new view is likely to emerge, but it does provide clues and a more specific notion of how consolidation is likely to be resolved by directing our attention to the oppositional idea as the most likely new dominant idea. In situations where there are no alternatives, default to the old orthodoxy is more likely; where there are several alternatives with equal social salience, the outcome will be contingent on political struggles.

In sum, ideational change is usefully examined as a product of collapse and consolidation. Situations involving the combination of unmet expectations and undesired consequences are likely to facilitate collapse while those where expectations are fulfilled and/or desired consequences occur favor ideational reproduction. Consolidation of a new structure requires the existence of a viable oppositional idea, the prescriptions of which correlate with socially desired results.

Illustrations

Key aspects of this view of ideational change and inertia are represented as ideal types in Table 1. The vertical axis refers to the social expectations resulting from ideas and behavior. Collective ideas generate a net of social expectations about what should occur and what is desirable. These form a baseline for distinguishing particular events as socially notable: what is a crisis depends on the lens through which it is viewed. Group (i.e., state) behavior can either adhere to or violate ideational prescriptions. Ideational rules are usually respected, but this need not always be the case: other factors besides ideas also can affect behavior and for various adventent and inadvertent reasons these may do so in particular situations. When adhering to ideational prescriptions, the collective expectation is that events will match; that prescribed action will bring desirable consequences and prescribed action undesirable consequences. The horizontal axis reflects the actual outcomes societies experience vis-à-vis events. Events can have consequences that are either socially desirable and considered a success or undesirable and seen as a failure. The interaction between expectations and consequences is central to the possibility of ideational stasis or change. The table suggests four ideal types each with a specific logic of continuity or change. In what follows I further clarify the asymmetric logic of ideational change and provide illustrations of the different types. Each example is drawn from a “least likely” foreign policy area: one not typically associated with ideationally driven decision making and action.

Type 1: “Old Faithful”

Continuity is likely when actors adhere to ideational guidelines and the consequences are socially desirable. This would be a confirmation of extant beliefs and hence would help reinforce or reproduce them. Societal actors would find little reason to reassess the prevailing orthodoxy and would have few tools in convincing others that the orthodoxy should be changed. Equally important, societies do not have an incentive to see if such a belief is objectively correct. After all, the desired outcome might result for reasons unanticipated by ideational prescriptions. For example, a belief that a spirit looks after one’s general well-being may be bolstered by an absence of undesired car incidents, but the actual causes may have more to do with driving lessons or minimal road traffic.

In contemporary politics, a similar example of the dynamic of Type 1 is evident in general deterrence theory. A societal belief (as found perhaps in the United States, Switzerland, and Singapore) that the general threat to impose costs on a potential attacker prevents aggression would be supported by an absence of attacks. Hence the normal absence of attacks upholds that belief and few policymakers question the validity of deterrence. The problem, of course, is that the absence of attack may have little to do with the casual mechanisms (the threat to impose costs that changes the intentions of the would-be aggressor) asserted by the theory (Lebow and Stein

---

8 There are parallels here with the external vs. internal sources of acceptance/rejection of scientific theories in the philosophy of science literature. See Andersson (1994).

9 For a more in-depth analysis of one idea, see Legro (2000).

10 When an approach produces a desirable outcome people “satisfice” and ignore the possibility of even better alternatives (Simon 1957).
1989). States may not be attacked for a variety of other reasons that might constrain would-be aggressors such as domestic politics, moral considerations, lack of incentive, international stigmatization, and so on. Despite this possibility of unfounded belief, ideational continuity in this instance is likely.

**Type 2: “Do something!”**

Transformation is most likely when societies adhere to ideational prescriptions, but the actual outcome contradicts expectations with stark failure. This combination challenges dominant beliefs. The extant collective idea has led the society in a misguided direction—that is, prescribed behavior does not generate anticipated outcomes but instead highly undesirable ones. This opens a window of self-critique and reflection in society, and it gives opponents of the dominant orthodoxy leverage to make their case and rally support. Likewise, for critics to successfully fight (and hence consider fighting) the accepted collective wisdom, strong intersubjective evidence (to convince others) and motivation (to bear the costs) for doing so has to be present. Otherwise social sanctions and collective action problems can inhibit action. Thus, the more significant the contradiction between expectations and consequences, the more severe the consequences, the more likely societies will face widespread discord, the more likely thresholds of collective action will be met, and the more probable is collective reorientation (assuming a viable replacement idea).

Goldstein’s (1993) study of American trade policy illustrates this dynamic. Goldstein’s account tracks the shift from protectionism to liberalization as the dominant conceptual orthodoxy regarding trade policy. She argues that American trade policy after World War I—as witnessed in the Fordney-McCumber and Smoot-Hawley tariffs—was the product of a post-civil war tradition. These tariffs were suited neither to America’s position in the global economy nor the economic efficiency of U.S. producers, both of which suggested that a more open economy was more appropriate. Nonetheless, the protectionist orthodoxy that held that America could expand exports without changing its own market openness prevailed, even while most economists believed that liberalization was needed. But the economists were allegedly poor advocates of their ideas and the hegemony of protectionist ideas was potent (Goldstein 1993, 126, 134–135).

The question then is what changed this orthodoxy? Goldstein argues that the Great Depression acted as a shock that broke the logjam favoring protectionism. But why it was the Great Depression that had this effect demonstrates the causal mechanisms of Type 2: actual consequences contradicted the expectations of ideationally driven behavior with dramatic undesirable consequences. “By the early 1930s, both the economic model and the policy process associated with high tariffs were in disrepute. This national mood created a policy window, allowing those critical of high-tariff policy to restructure tariff-making institutions to facilitate tariff reform. The shock of the Great Depression then created an opportunity for political entrepreneurs to restructure trade institution” (Goldstein 1993, 139). Equally important, there was a readily available opposing replacement idea—the liberal doctrine favored by economists. This idea gained salience because it had a solution for the problems just experienced, and it met with initial success after World War II. Goldstein (1993, 239n) contends these results were not a simple rational adaptation to circumstances or “an enlightenment triumph” but instead a shift in collective mentality brought on by the interaction of ideas and circumstance and cemented by the availability of a socially plausible alternative.

**Type 3: “If it ain’t broke”**

Continuity is probable even when ideational mandates (and societal expectations) are violated but the results are desirable. This case challenges the legitimacy of the

### Table 1: Ideas and Events

<table>
<thead>
<tr>
<th>Prescribed by Ideas</th>
<th>Success</th>
<th>Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type 1:</strong> “Old faithful.”</td>
<td>Continuity</td>
<td>Type 2: “Do something!”</td>
</tr>
<tr>
<td><strong>Type 3:</strong> “If it ain’t broke.”</td>
<td>Continuity</td>
<td>Type 4: “Told you so.”</td>
</tr>
</tbody>
</table>

**Behavior**

**Proscribed by Ideas**
extant collective idea, but it does so weakly. Because the consequences are favorable, there is less likely to be a public debate on “what went wrong.” People tend to pay attention to unanticipated events when the consequences are disagreeable. When they are not, there is no experience to motivate the uncertain costly action of change, and collective ideation and action are difficult. The continuity of this outcome with past ideationally shaped behavior makes it far less salient than in Type 2 above. An incident becomes an event or shock exactly because it disrupts the extant order, which in this case does not happen.

Sagan’s (1993) study of U.S. nuclear weapons safety suggests that even when behavior is contradictory to collective ideas, if the results are desirable, ideas often change little. Sagan describes a military culture that strongly emphasized operational safety, but at the same time prioritized military readiness (which might suffer from too many safety restrictions) and encouraged the cutting of corners. A variety of military accidents occurred that contradicted the injunction on safety, but because none of these resulted in actual nuclear disasters, little organizational change took place.

Sagan assesses the proposition that organizations learn by trial and error (in a rationally adaptive model) by studying whether the United States altered its military procedures based on experience in two cases: (1) a comparison of the nuclear alerts in the 1962 Cuban Missile Crisis and the 1973 Yom Kippur war and (2) a series of false warning incidents that occurred within NORAD in 1979 and 1980. He finds that very little actual learning takes place in terms of searching for the causes of the accidents and then altering procedures (and ideas) to correct them. In the first case, the incidents of the 1962 crisis were subsumed under the broader interpretation that “deterrence works” (Sagan 1993, 224). In terms of the false warnings, Sagan maintains that it was hard to know what to make of them: they were a failure in terms of preventing false warning, but a success in terms of avoiding undesired nuclear attack. What is interesting is how often the incidents were suppressed or “operator error”—the fluke fault of an errant individual—and not the underlying systematic problems—was held to blame. Thus, lacking a disaster, defenders of the extent orthodoxy were able to highlight success, rather than the failure, and little learning or change occurred.

Sagan concludes: “The failures of organizational learning suggest that such displays of confidence have been internalized. What is perhaps most disturbing, is the degree to which a set of internal organizational myths have developed over time: that false warning problems have been solved; that there is no chance of an unauthorized use of a nuclear weapon; that nuclear alerts are not dangerous; and that public fears about accidental nuclear war are therefore deeply misguided” (1993, 284). In the absence of socially undesirable results, change in “myths” is difficult.

**Type 4. “Told you so”**

Continuity is likely in situations where societies ignore ideational mandates, and the results are undesirable. As a discrepant costly outcome, it draws attention and becomes socially salient. In the usual effort to assign blame, however, the dominant orthodoxy’s defenders are likely to reinforce their position by warning of the penalties of deviation and need to adhere to prescribed behavior. Critics, in contrast, are left with a more difficult task in effecting widespread change. This combination reinforces the extant ideational structure.

An example of the causal mechanisms in Type 4 is evident in German foreign policy and its relationship to diplomacy during the 1991 Yugoslavian crisis, specifically Germany’s unilateral recognition of Croatia. Since World War II, Germany has maintained a belief in multilateralism and action through international institutions as the most efficacious foreign policy (Berger 1998). During the summer of 1991, when Croatia and Slovenia declared independence from Yugoslavia and the Yugoslavian Army used force to prevent the secession, the European Community (EC) formulated a collective policy of supporting a united Yugoslavia to encourage peace. Throughout the summer and fall, Germany advocated recognizing Croatia and Slovenia as independent countries, but could not carry the weight of the EC with it. By the end of the year, others began to see merit in the German position, but wanted to make sure that the new states met certain conditions regarding minority rights, governmental and legal structure, and education. Nonetheless, before these issues were resolved, Germany on December 23, 1991 unilaterally recognized Slovenia and Croatia and in doing so violated its long-standing norm of multilateralism. In this instance, domestic political pressures and bargaining dynamics induced German leaders to deviate from the country’s post war foreign policy tradition (Krieger 1994; Maull 1995/1996, 102–105; Crawford 1996).12

---

11Blaming individuals for failures is a similar to crediting idea entrepreneurs for any changes that do occur. Both ignore the forest for the tree. See Perrow 1984, 9, 23–30.

12Arguing against this common interpretation is a book by the Foreign Ministry official who was in charge of German policy toward Yugoslav from 1991 to 1995. See Libal 1997, 148–156.
The result, Krieger notes, was that “Germany got its fingers burned” (1994, 38). The decision contributed to the spiral towards conflict in Yugoslavia and to the difficulties the EC has had in formulating a common foreign and security policy. It also went against the preferences of Germany’s major allies, including France, Britain, and the United States, and in so doing sparked widespread international criticism. Maull argues these consequences were “something of a trauma for German diplomacy. The principle conclusion drawn from the experience in Bonn was that henceforth Germany should avoid such strong-arm tactics, and should keep in step with its major partners” (1995/1996, 109). In short, the experience reinforced extant collective thinking.

**Implications**

Ideational change has been the stepchild of the recent wave of work on collective ideas in politics. In addressing this issue, I have highlighted a neglected, but nonetheless critical, factor: the role of ideational structure itself, which in conjunction with experienced events, encourages actors to overcome collective ideation challenges in some cases, but not others. This argument speaks to two central questions in social analysis: (1) when can individuals seeking change overturn long-held societal concepts and beliefs? and (2) when are societies likely to alter their modes of thinking and action to fit circumstances? These two issues, often addressed separately in different literatures, are intertwined. Both relate to the endogenous role of ideas in their own transformation.

First, a fundamental approach to social change focuses on agency—the instrumental and creative efforts of various entrepreneurs who are able to persuade and inspire others to innovate. But what makes such agents likely to succeed? One can always find individuals promoting a variety of ideas, and the successful idea will always have an entrepreneur behind it. What allows some ideas to become dominant in guiding societal calculations and action? Rather than feature the resources, strategies, persuasiveness, or charisma of individuals, the argument here suggests the success of agents depends importantly on ideational structure and its interaction with events which in particular circumstances—when expectations are contradicted with undesirable consequences—allows for dramatic change while in many other instances works as a significant constraint.

This is not to suggest that human traits or efforts are irrelevant to change. The human mind, after all, actually processes the contradiction between expectations and circumstances. Hence understanding individual belief change is an important agenda that both complements and informs the primarily social focus of this article. Entrepreneurial efforts of groups and individuals also are consequential. Ideational transformation depends on the existence of an opposing idea with social support, which implies the necessity of individuals creatively opposing the dominant orthodoxy until circumstances are conducing. But ideational stasis, and even transformation, cannot be reduced to individual efforts that ignore the preexisting structure of collective ideas and their interaction with events. In some respects this argument parallels the development of (rational choice) agency approaches that seek to take account of intersubjective social beliefs (e.g., Ferejohn 1991; Weingast 1996; Bates, de Figueiredo, and Weingast 1998). But the difference is that here collective ideas have a causal autonomy that is by definition minimized in rational choice approaches where ideas are referred to more as a residual factor—e.g., to resolve multiple equilibria problems.

This autonomy is apparent in the different ways collective ideas can set the broader parameters that make instrumental efforts of actors intent on change likely to succeed or fail. First collective ideas determine social (and political) expectations which are not easily derivable from (and may constitute) the beliefs individuals hold (Eagly and Chaiken 1993, 628–632, 642–646; DiMaggio 1997). Second, and most critical, it is the interaction of those expectations with specific consequences that sparks the type of social “tipping” that can overcome collective ideation problems that might otherwise frustrate collapse and consolidation. The “Old faithful,” “If it ain't broke,” and “Told you so” types in Table 1 reflect the powerful tendencies toward continuity in social life where the barriers to collapse and consolidation are formidable. In contrast, the Type 2 “Do something!” nexus of ideationally driven behavior and failure provides the volatile conditions most likely to spark change and overcome any extant aggregation barriers impeding collapse.

As well as telling us something about when the old orthodoxy is likely to give way, ideational structure may also give clues to the vexing issue of what new ideas will become dominant during consolidation. Such structures seem to frequently take a binary form with a dominant idea that guides societal action and an alternative idea that exists in opposition—be it containment- engagement, free trade-protectionism, isolationism-internationalism, Keynesianism-monetarism, or offense-defense. When an old orthodoxy collapses, an oppositional idea with preexisting social support that appears to coincide with socially desired results (whether it actually caused such results is irrelevant) is likely to be the new focal
point of consolidation and institutionalization. To the extent these conditions are not met, or ideational structure has multiple viable oppositional ideas, the consolidation phase is more likely to be characterized by the indeterminate political contestation noted by other observers.

The explanation of change here also addresses the key question of when societies can be expected to adapt to the demands of their environments. Many posit that societies rationally react to external stimuli and collective ideas are adopted as simply devices that facilitate functional societal adjustment. In effect, mental phenomena, both individual and collective, are given little causal weight. Yet the argument here suggests the centrality of collective ideas. The point is not that brute circumstances are irrelevant to ideational change. Instead it is that they interact with prior collective ideas to influence future ideas and behavior.

This view has similarities with Bayesian learning which allows for rational learning based on subjective beliefs. But there appear to be several noteworthy differences. The first is that my argument seems relevant to understanding a range of political circumstances where there is not adequate data to form rational beliefs about the underlying state of the world in accordance with Bayesian reasoning. Under conditions of strong uncertainty, ambiguity, and problem complexity—when even subjective probability distributions are elusive—Bayesian updating is problematic. Yet the absence of data and problem complexity are key traits in many areas of politics (especially foreign policy) where certain strategies are sustained by beliefs about other strategies that are never experienced (i.e., counterfactual) or where learning takes place not on a repeated test basis involving frequent feedback, but instead according to small and biased samples (Jervis 1976, 235; Tetlock 1998a, 870–872). Facing such conditions of strong uncertainty, there seems to be increasing agreement across disciplines that humans turn to collective ideas to inform action and learn, not simply to the objective external world (Denzau and North 1994, 14–17; Eagly and Chaiken 1993, 628–632).

Another difference between the framework and a Bayesian approach is an asymmetry summarized in the distinction between Type 2 “Do something!” (action prescribed by the dominant ideas that has undesirable consequences) and Type 3 “If it ain’t broke” (action prescribed by ideas that has desirable consequences). These two situations are largely symmetrical in terms of their contradiction with expectations and the need to update beliefs. But they have asymmetrical social effects for reasons not well captured by rational learning theory. That is, societies will treat equivalent losses and gains differently: losses are much more likely than comparable gains to trigger aggregation processes that lead to change. This makes no sense from a rational learning perspective.

Finally, Bayesian learning applied to societies is based on a unitary actor perspective, and it necessarily misses the collective ideation problems associated with change in collective ideas. This collective aspect may reduce motivation for individuals to allocate cognitive and other resources to gather information and improve their individual understanding and/or to act on such individual images to change collective ideas. These dynamics are, by assumption, ignored in a Bayesian format. Thus societies composed of varying individuals and interest groups react differently to external stimuli than the Bayesian unitary actor.

This is not to deny the importance of environmental conditions—especially critical events/shocks. But I try to go beyond standard formulations by specifying the likely conditions of when such events are likely to cause change or reinforce continuity. The relevance of this move is apparent when considering one of the most ubiquitous current approaches to change: path dependent and punctuated equilibrium analyses (Gould and Eldredge 1977; David 1986). These models have provided insights both in the realm of natural science (plate tectonics, paleontology) and in the social sciences (e.g., “qwertronics”). Nonetheless applied to nontechnical social affairs they have had little to say about why a critical juncture is a juncture, and why one path is taken versus another after a critical juncture occurs (Higgs 1987, 70–72). Yet in social relations, a crisis or critical juncture is in part dependent on the preexisting orthodoxy in that area. This idea determines both what is to be expected (and hence what is inadequate), what an alternative path might be, and in combination with action and consequence, whether change is likely. Thus, path dependency analyses that are based on human action would benefit from attention to collective beliefs that govern societal thinking. The challenge that confronts a variety of perspectives on politics is to specify how it is that collective ideas and situations, and agents and structures, interact and influence each other. Absent this specification, we are left with residual and unsurprising explanations contingent on history or politics or elegant meta-theory that has

---

13 On the way that uncertainty is handled in economics—and the problems it can cause for substantive rationality accounts, see Denzau and North 1994.

14 Bates, de Figueirdo, and Weingast 1998, 249–250 propose a variant of this but also lean heavily toward individual instrumental action and charismatic or persuasive agency in deciding outcomes.
little power to illuminate empirical puzzles. The approach here is intended as a modest attempt to address that divide. I have focused on collectively held causal ideas mainly related to foreign policy. But the overall framework should be applicable to a broader range of ideational political phenomena involving not only national policies on issues other than foreign affairs, but also such topics as revolution, regime change, and social movements. In each of these areas, it appears useful to investigate the way that collective ideas play a role—even in their own transformation.

Final manuscript received October 18, 1999.

References


McAdam, Doug, John McCarthy, and Mayer Zald, ed. 1996. *Comparative Perspectives on Social Movements: Political Opportunities, Mobilizing Structures, and Cultural Framings.* Cambridge: Cambridge University Press.


