Individual Adaptations to Cultural Contradictions:

Using Non-Monotonic Logic to Reconstruct Merton's Theory of Anomie

James D. Montgomery
Department of Sociology
University of Wisconsin – Madison
1180 Observatory Drive
Madison, WI 53706

jmontgom@ssc.wisc.edu

11 June 2004

I am grateful to Tom Fararo for helpful comments.

Individual Adaptations to Cultural Contradictions:
Using Non-Monotonic Logic to Reconstruct Merton's Theory of Anomie

Abstract. Merton's (1957) theory of anomie contains a classic sociological analysis of individual adaptations to cultural contradictions. The present paper uses non-monotonic logic to formalize Merton's micro-level analysis. I show that the adaptations of innovation, ritualism, and retreatism correspond to different extensions of a level default theory. My formalization makes explicit the importance of the self-concept in Merton's theory, and links the rebellion adaptation to a dynamic model of self-concept formation.

Introduction

Sociologists have often considered how individuals cope with "contradictions" in normative systems. In his theory of anomie, Merton (1957) offered a typology of adaptations based on the individual's acceptance or rejection of cultural goals and institutionalized means. But while Merton's theory has attained classic status (Cole 1975; Orrù 1990) and remains influential in subfields such as crime and deviance (Adler and Laufer 1995; Passas and Agnew 1997), many critics have emphasized the gaps and ambiguities in this theory (Dubin 1959; Harary 1966; Levine 1985; Messner 1988; Besnard 1990; Sztompka 1998; Featherstone and Deflem 2003). If Merton's theory is to remain relevant for contemporary sociology – more than a museum-piece to be admired by students or a straw-man to be thrashed by theorists – then it needs to be reconstructed in a more rigorous fashion.

Merton's own use of terminology from deontic logic – the branch of logic addressing the normative concepts of *prescription*, *permission*, and *proscription* (cf. Merton 1957, p. 133) – might immediately suggest the use of formal logic to model decision-making within his theory. Of course, given that classical logic does not tolerate contradictions (the proposition Q and its negation $\neg Q$ cannot both be true) and that standard deontic logic does not permit moral dilemmas (the propositions "you ought to do Q" and "you ought to do $\neg Q$" cannot both be true), these conventional logics seem ill-suited for formalizing Merton's theory. However, given recent innovations designed to capture features of common-sense reasoning, formal logic has become a more valuable tool for sociological analysis of decision-making.

Using a simple non-monotonic logic, Montgomery (2004) developed a model well-suited for understanding the choices made by an individual facing contradictions among social norms, her self-concept, and constraints. In the present paper, I show how this model helps to clarify and extend Merton's original intuition about individual adaptations to such contradictions. Unlike previous critics who have responded to the problems with Merton's micro-level analysis merely by revising his typology (Durbin 1959; Harary 1966; Besnard 1990), I offer a new theoretical method that sociologists could use to clarify particular cultural contradictions, identify the particular patterns of

choices that might result from these contradictions, and explain the particular changes in the self-concept that might be induced by these adaptations.

To proceed, I reanalyze Merton's own example focusing on the cultural contradiction faced by Americans who are unable to satisfy the key societal goal (money) through socially approved means (work). After more fully introducing the model of decision-making from Montgomery (2004), I show how three of the adaptations described by Merton – innovation, ritualism, and retreatism – follow naturally from logical analysis of the individual's choices. I then consider minor modifications of the model (different rankings of goals and means) and the example (introduction of multiple goals) to address other claims made by Merton (1957). Finally, my discussion of the rebellion adaptation leads to a dynamic model of self-concept formation and analysis of the stability of different self-concepts. I close by arguing the relevance of Merton's theory and non-monotonic logic for contemporary sociological perspectives on agency, indicating possible directions for future research.

Merton's example and typology

Merton's (1957, Chap. 4) analysis begins with a stylized account of the contradictions in American culture. This culture prescribes *money* as the primary goal, prescribes *work* as the appropriate means of achieving this goal, and proscribes *crime* as an inappropriate means. Adding some relevant background information, we can begin to specify the individual's decision problem in propositional form as shown in the first column of Figure 1. In attempting to decide whether to work or commit crime, the individual thus reflects on her self-concept (P1), social norms (P2, P3) and constraints (P4, P5, P6). Recognizing that these propositions are mutually consistent, the individual might satisfy all norms and constraints by choosing to work and not commit crime; this pattern of action would generate money and thus confirm the individual's identity as an American. However, a cultural contradiction arises when the individual lacks access to a high-paying job and hence cannot obtain money without committing crime. Formally, suppose that proposition P6 is replaced by P6′ given in Figure 1. Given that the full set

of propositions {P1, P2, P3, P4, P5, P6′} is now inconsistent, the individual cannot choose actions in a way that will simultaneously satisfy all of these propositions.

Merton's informal analysis of this example led to his well-known typology of adaptations given in Figure 2. A *conformist* retains both goals and means; an *innovator* retains the goal but abandons the socially approved means; a *ritualist* retains these means but abandons the goal; a *retreatist* rejects both goals and means; a *rebel* substitutes new goals and means for existing ones. Applying this typology to the present example, a ritualist would reject P2; an innovator would reject P3; a retreatist would reject both P2 and P3. Some obvious problems with this typology have been well-explicated by critics (Dubin 1959; Harary 1966; Besnard 1990). Clearly, the rebellion adaptation seems an ad-hoc addition to the four combinations generated by crossing the options {accept goals, reject goals} with the options {accept means, reject means}. This typology also fails to show that access to socially acceptable means is a necessary condition for conformity, and leaves ambiguous whether access to these means is a sufficient condition. Thus, in the present example, conformity requires P6 rather than P6′, but it remains unclear whether individuals with access to high-paying jobs are always conformists. Further, foreshadowing my analysis below, we might wonder why cultural contradictions never lead to the rejection of roles in the self-concept rather than goals or means. In the present example, wouldn't individuals facing P6′ sometimes reject P1 rather than P2 or P3?

Beyond these (and other) problems with his typology, critics have often highlighted ambiguities in Merton's use of the term "anomie" (Levine 1985; Messner 1988; Besnard 1990; Sztompka 1998; Featherstone and Deflem 2003). Some of this ambiguity derives from Merton's tendency to conflate different levels of analysis. Following Messner (1988), we might distinguish Merton's micro-level analysis of individual adaptations to given cultural contradictions ("strain theory") from his macro-level analysis of the process by which social norms become legitimated or de-legitimated ("anomie theory"). Given this distinction, the present paper focuses exclusively on the micro-level strain theory.[1] Thus, in the following analysis of Merton's example, I take the social norms {P2, P3} as given, examining the individual's adaptations to the

[1] Consequently, much like Merton himself in the earliest statement of his theory (Merton 1938), I make no use of the term "anomie."

personal "strain" induced by these norms, without attempting to explain the emergence or stability of this configuration of norms in the larger society. Building on the present foundation, future research might well attempt a more complete reconstruction of Merton's theory, linking micro and macro levels of analysis. But even without this macro-level addition, I will argue that the micro-level framework used below could also be employed to clarify more recent accounts of "cultural contradictions" and some contemporary sociological perspectives on agency.

A logical framework for reasoning about normative obligations

The gaps and ambiguities in Merton's typology suggest the need for a more rigorous analysis of decision-making in the face of cultural contradictions. To model how an individual might reason about normative obligations, researchers from philosophy, computer science, and artificial intelligence have proposed a variety of deontic logics (Hilipen 1971; Åqvist 1987). Unfortunately, standard deontic logic has many well-known problems (see, e.g., Forrester 1996) that would severely limit its usefulness in sociological applications.[2] Perhaps most crucially, because it presumes that the set of norms is logically consistent, standard deontic logic cannot address the moral dilemmas that arise in real-world social systems.[3]

However, recent innovations in formal logic have addressed many of these problems. Classical logic is "monotonic" in the sense that formulas, once derived, are never retracted. In contrast, recently developed non-monotonic logics allow tentative derivations that may be retracted in light of new information (Brewka, Dix, and Konolige 1997). Non-monotonic logics thus offer a useful framework for modeling common-sense reasoning in social settings where information is either incomplete or contradictory.

---

[2] Standard deontic logic is formulated as a type of modal logic (Chellas 1980; Hughes and Cresswell 1996) in which the modal operator $\pm$ is used to represent "ought." To apply standard deontic logic in the present case, we might begin by restating the norms as modal propositions (e.g., the norm P2 might take the form $\pm (a \rightarrow m)$) and then attempt to derive the individual's obligation to perform actions.

[3] Formally, for any formula Q, standard deontic logic does not permit both $\pm$ Q and $\pm \neg$Q to be true. This restriction might be understood as a consequence of the possible-world semantics. In standard deontic logic, each world is associated with a non-empty set of "ideal" worlds. If $\pm$ Q is true at a given world, then Q must be true at every associated ideal world. Because no world (ideal or otherwise) can contain the contradiction Q and $\neg$Q, it is impossible for both $\pm$ Q and $\pm \neg$Q to be true at any world. See Horty (1994)

Sociologists might especially consider Horty's (1993, 1994) attempt to address moral dilemmas by specifying norms as default rules in a default logic.[4] Using this framework to reason about norms and actions, individuals would attempt to follow as many norms as consistently possible, and conflicting norms would generate multiple "morally defensible" courses of action.[5]

Inspired by this application of non-monotonic logic, Montgomery (2004) used a *level default theory* to model the process by which individuals reason about norms and actions.[6] Formally, a level default theory (LDT) is a set of logical formulas T partitioned into subsets (T_1, …, T_n) ranked by priority. The formulas in T_1 are given higher priority than those in T_2, which are given higher priority than those in T_3, etc. Given this ranking, $E = E_1 \cup \ldots \cup E_n$ is an *extension* of the LDT if and only if $E_1 \cup \ldots \cup E_m$ is a maximal consistent subset of $T_1 \cup \ldots \cup T_m$ for all $m \leq n$.[7] More intuitively, to derive an extension of the LDT, we begin by collecting as many formulas from T_1 as possible while maintaining logical consistency, then extend this subset by adding as many formulas from T_2 as possible while still maintaining logical consistency, then extend this subset by adding as many formulas from T_3 as consistently possible, and so on. Logical consistency of the formulas in T implies a unique extension (the union of all levels) while inconsistency of the formulas in T may result in multiple extensions. For readers unfamiliar with formal logic, Appendix 1 provides additional discussion and shows how to derive the extensions of an LDT using a truth table.

Adopting a slightly modified version of the LDT specified in Montgomery (2004), suppose that individuals reason about actions using the level default theory

(1) $T = (C, S \cup N, X)$

---

for further discussion.

[4] See Nute (1997) for other attempts to reformulate deontic logic as a non-monotonic logic.

[5] Of course, standard logic may remain useful at the metatheoretical level as researchers attempt to verify the logical consistency of sociological theories (Péli, et al 1994, Hannan 1998, Kamps and Pólos 1999). See also the metatheoretical use of non-monotonic logic by Pólos and Hannan (2002) to reconcile some apparent theoretical inconsistencies in organizational ecology.

[6] Level default theories are a simpler alternative to the default logic (Reiter 1980) used by Horty (1993, 1994). See Brewka et al (1997, p. 54 ff) for a discussion of level default theories and their relation to other consistency-based non-monotonic logics.

where C is the set of constraints (assumed to be logically consistent), S is the set of roles comprising the individual's self-concept, N is the set of norms (that may be inconsistent), and X is the set of all possible actions and their negations.[8] Deriving the extension(s) of this LDT, we obtain the set(s) of actions that could be taken by the individual. Intuitively, given the constraints that she faces, the individual attempts to uphold as elements of her self-concept and as many social norms as possible. But in the presence of cultural contradictions, it will be impossible for any extension of the LDT to incorporate the entire set $S \cup N$, and the various extensions of the LDT will reflect alternative adaptations that the individual might make to these contradictions.

To analyze Merton's example, we must first translate the verbal list of propositions given in the first column of Figure 1 into the logical formulas given in the second column.[9] Recognizing that access to labor-market opportunities may vary across individuals, the set of constraints for an individual with access to opportunities is thus

(2) $\quad C_1 = \{m \rightarrow (w \vee c), c \rightarrow m, w \rightarrow m\}$

while individuals who lack these opportunities face the set of constraints

(3) $\quad C_2 = \{m \rightarrow (w \vee c), c \rightarrow m, (w \wedge \neg c) \rightarrow \neg m\}$

For all individuals, the set of roles in the self-concept is

---

[7] Brewka, et al (1997) refer to the set E as a *preferred subtheory* of T. Following Montgomery (2004), I use the term *extension* to highlight the parallel between preferred subtheories of LDTs and extensions of default logics.

[8] Following Montgomery (2004), the set of norms N contains all formulas taking the form *role* $\rightarrow$ Q where Q is a (simple or complex) formula involving roles or actions. Thus, in the present example, both the goal $a \rightarrow m$ and means $a \rightarrow (w \wedge \neg c)$ are types of norms. The final level X of the LDT is included for technical reasons. I assume that the individual faces binary (either/or) choices. In the present example, she must choose either *work* or not *work*, and either *crime* or not *crime*. Thus, the individual's pattern of action must be a maximal consistent subset of $X = \{w, \neg w, c, \neg c\}$. Inclusion of X in the LDT ensures that the individual's choices are complete even if they are underdetermined by the preceding levels of the LDT.

[9] I am adopting standard notation for the logical connectives *and* ($\wedge$), *or* ($\vee$), *not* ($\neg$) and *implies* ($\rightarrow$). Further, to make the notation compact, I am using only the first letter of the atomic propositions *American* (*a*), *money* (*m*), *work* (*w*), and *crime* (*c*).

(4) $S_1 = \{a\}$,

the set of norms is

(5) $N_1 = \{a \rightarrow m, a \rightarrow (w \wedge \neg c)\}$,

and the set of possible actions and negations is

(6) $X_1 = \{w, \neg w, c, \neg c\}$.

Note that, in contrast to standard deontic logic, I have specified each of the norms without using a modal operator to denote the "ought" explicit in their verbal specification. Thus, constraints and norms share the same logical form (material implication). Nevertheless, constraints and norms are distinguished by their relative positions in the LDT. Crucially, norms are placed after constraints. Thus, while every extension will contain the entire set of constraints, logical inconsistencies among norms (or between norms and the self-concept) may generate multiple extensions, each containing only a subset of $S \cup N$.

Conformity

Consider an individual with access to labor-market opportunities, and thus the level default theory

(7) $T = (C_1, S_1 \cup N_1, X_1)$.

Because all of the formulas in the first two levels of the LDT are mutually consistent, this LDT has the unique extension

(8) $E = C_1 \cup S_1 \cup N_1 \cup \{w, \neg c\}$

which implies the unique pattern of action

(9) $A = E \cap X_1 = \{w, \neg c\}$.

In this way, we obtain the first adaptation described by Merton: *conformity*. Given that constraints, self-concept, and norms are mutually consistent, the individual accepts both the goals ($a \rightarrow m$) and means ($a \rightarrow (w \wedge \neg c)$) approved by the society. Formally, the entire set N is contained in the unique extension of the individual's LDT.

Innovation, ritualism, and retreatism

The other adaptations described by Merton arise when a contradiction is introduced into the social system. Replacing the constraint set C_1 with the set C_2, the individual's LDT becomes

(10) $T = (C_2, S_1 \cup N_1, X_1)$.

From the constraint set C_2 comprising the first level of the LDT, it is now possible to derive the proposition that money implies crime ($m \rightarrow c$). Consequently, the union $C_2 \cup S_1 \cup N_1$ is now inconsistent, given that the individual is an American (a) who is supposed to make money ($a \rightarrow m$) but not commit crime ($a \rightarrow (w \wedge \neg c)$). Given this contradiction, the LDT generates multiple extensions. Letting $\mathfrak{E}$ denote the set of extensions of T,

(11) $\mathfrak{E} = \{C_2 \cup S_1 \cup \{a \rightarrow m\} \cup \{\neg w, c\},$
$C_2 \cup S_1 \cup \{a \rightarrow m\} \cup \{w, c\},$
$C_2 \cup S_1 \cup \{a \rightarrow (w \wedge \neg c)\} \cup \{w, \neg c\},$
$C_2 \cup N_1 \cup \{\neg w, \neg c\},$
$C_2 \cup N_1 \cup \{w, \neg c\},$
$C_2 \cup N_1 \cup \{\neg w, c\},$
$C_2 \cup N_1 \cup \{w, c\}\},$

we find that $\mathfrak{E}$ contains seven different extensions.[10] The first and second extensions listed in $\mathfrak{E}$ correspond to the *innovation* adaptation by which an individual retains societal goals ($a \rightarrow m$) but abandons the means ($a \rightarrow (w \wedge \neg c)$). While both extensions imply that the individual will commit crime, the individual may or may not decide to work, corresponding perhaps to cases of "career" and "white-collar" criminals (compare the examples of innovation in Merton 1957, pp 141-149).[11] In contrast, the third extension reflects the *ritualism* adaptation by which an individual abandons societal goals but retains the means.

The four remaining extensions correspond to the *retreatism* adaptation. However, my present conception of retreatism differs in two important ways from Merton's original specification. First, while Merton equated retreatism with the rejection of both goals and means, the present analysis equates retreatism with rejection of a role in the self-concept. Thus, norms are not abandoned – note that final four extensions contain the entire set of norms N_1 – but simply become moot because the antecedent of these norms (a) is not contained in these extensions. Arguably, this conception of retreatism helps reconcile aspects of Merton's own argument.[12] It also makes explicit the importance of the individual's self-concept in Merton's theory.

To expand on this first point, we might attempt to approximate Merton's original argument more closely through a different specification of the individual's LDT in which the *American* role is merely implicit. That is, we might assume

(12) $\quad T = (C_3, \{m, w \wedge \neg c\}, X_1)$

where the constraint set

[10] See Appendix 1 for derivation of these extensions using a truth table.
[11] Alternatively, we might have assumed that time constraints prevented the individual from choosing both work and crime. Adding the constraint $w \rightarrow \neg c$ to the constraint set C_2 would eliminate the second and seventh extensions listed in equation (11).
[12] Merton claims that retreatism is "most likely to occur when *both* the culture goals and the institutional practices have been thoroughly assimilated by the individual and imbued with affect and high value" (Merton 1957, p 153, his emphasis). If we accept this initial claim, the individual's rejection of the role *American* might seem more plausible than Merton's subsequent argument that "the conflict is resolved by abandoning *both* precipitating elements, the goals and the means" (Merton 1957, pp 153-4, his emphasis).

(13) $C_3 = \{m \rightarrow (w \vee c), c \rightarrow m, (w \wedge \neg c) \rightarrow \neg m, w \rightarrow \neg c\}$

has been slightly expanded merely to simplify the analysis.[13] Given Merton's typology of adaptations, we might expect the inconsistency in T to produce the three extensions

(14) $\mathfrak{E} = \{C_3 \cup \{m\} \cup \{\neg w, c\},$
$C_3 \cup \{w \wedge \neg c\} \cup \{w, \neg c\},$
$C_3 \cup \{\neg w, \neg c\}\},$

corresponding to innovation (rejection of means $w \wedge \neg c$), ritualism (rejection of goal m), and retreatism (rejection of both goals and means). However, while the first two sets in $\mathfrak{E}$ are valid extensions of the LDT, the third set is not a valid extension. Intuitively, given the requirement that every extension must be a *maximal* consistent subset, the individual never rejects more formulas than necessary to eliminate inconsistency. If the individual rejects the goal (m), she is not forced also to reject the means ($w \wedge \neg c$). Conversely, if she rejects the means, she can retain the goal. Thus, the LDT has only two extensions. If we retain the plausible assumption that extensions must be maximal consistent subsets – that individuals abandon norms only when forced to do so – then innovation and ritualism are the only possible adaptations. Given our initial specification of the LDT, retreatism emerges as a third possible adaptation because the individual recognizes that both norms are conditioned upon a role (a) that might itself be rejected.

The present analysis diverges further from Merton's original argument by emphasizing that retreatism is not a particular pattern of action, but rather the individual's (lack of) motivation for taking action. Merton seemed to associate retreatism with the set of actions $\{\neg w, \neg c\}$. In contrast, equating retreatism with rejection of a role in the self-concept, retreatism is consistent with every feasible pattern of behavior. Having rejected the *American* role, the individual is not obligated to accept the goals or means of American society, but neither is she prohibited from engaging in the behaviors adopted

[13] The additional constraint ($w \rightarrow \neg c$) eliminates the "white-collar crime" extension of the LDT.

by innovators or ritualists. In the present example, because the individual's self-concept contains no other roles (which would be the antecedents of other norms), her behavior is completely underdetermined. Thus, from the perspective of the analyst (who cannot directly observe the individual's reasoning process), retreatism may be observationally equivalent to innovation or ritualism.

Indeed, given that multiple motivations may lead to the same pattern of action, it is unclear whether the individual herself would necessarily possess a unique motivation. Past actions may be rationalized in multiple ways: "Did I commit crime because I rejected the norm that Americans should not commit crime, or because I rejected the role of American?" On one hand, we might adopt a "rational" specification of the reasoning process which presumes that the individual always possesses a unique motivation at the moment of decision. Formally, we would assume that the individual first selects a particular extension (E) of her LDT and then takes the corresponding action ($A = E \cap X$). In this way, the individual acts *because* she is a retreatist (or innovator or ritualist). On the other hand, adopting a "rationalizing" specification of decision-making, we might assume that the individual recognizes that an action has multiple motivations, and need not (or perhaps cannot) select among them. Formally, given that a particular action (A) is chosen, we might assume that the individual retains every extension that might have motivated that action (every extension E such that $A = E \cap X$). In this way, the individual who adopts a particular pattern of action – say $\{\neg work, crime\}$ – recognizes that her action may be consistent with either innovation or retreatism.

Alternative rankings of formulas in the level default theory

Before addressing the rebellion adaptation, several issues warrant further discussion. First, we might consider alternative rankings of constraints, the self-concept, and norms within the LDT. The analyst's placement of these formulas within the LDT is not arbitrary, but constitutes an important substantive assumption about the individual's level of "certainty" or "commitment" to various beliefs about herself and the social system. Thus, if the individual is more "committed" to some types of formulas (e.g., her self-concept) than others (e.g., social norms), this should be reflected in the relative

placement of S and N in the LDT. Moreover, given our placement of S and N on the same (second) level of the LDT, we lose the potentially important distinction between the belief "I am an American" and the norm "I ought to be an American." For these reasons, Montgomery (2004) assumed that the individual's LDT was given by

(15) $T = (C, S, N, X)$

so that the self-concept is placed before norms. Substituting the sets C_2, S_1, N_1, and X_1 into this LDT, we obtain

(16) $$\mathfrak{E} = \{C_2 \cup S_1 \cup \{a \rightarrow m\} \cup \{\neg w, c\}, \; C_2 \cup S_1 \cup \{a \rightarrow m\} \cup \{w, c\}, \; C_2 \cup S_1 \cup \{a \rightarrow (w \wedge \neg c\} \cup \{w, \neg c\}\}$$

so that $\mathfrak{E}$ contains only the first three extensions from equation (11). Because the individual is now more "committed" to her self-concept than to social norms, retreatism is no longer a possible adaptation to the cultural contradiction facing the individual. Perhaps this reasoning implicitly underlies Merton's own claim that retreatism is "probably the least common" adaptation (Merton 1957, p. 153).

Of course, consistent with Merton's claims that goals and means are differentially assimilated by individuals from different social classes (Merton 1957, p. 151), we might further rank the norms themselves. Placing goals before means, the LDT becomes

(17) $T = (C_2, S_1, \{a \rightarrow m\}, \{a \rightarrow (w \wedge \neg c)\}, X_1)$

so that innovation is the unique adaptation (generating the first and second extensions listed in equation 11). Alternatively, placing means before goals, the LDT becomes

(18) $T = (C_2, S_1, \{a \rightarrow (w \wedge \neg c)\}, \{a \rightarrow m\}, X_1)$

so that that ritualism is the unique adaptation (generating the third extension listed in

equation 11).[14]

Having considered possible rankings among the formulas ($S \cup N$) on the second level of the LDT, we might also consider reinterpreting some of the constraints on the first level as mere beliefs that should be "demoted" to the second level of the LDT. For instance, suppose that the constraint $(w \wedge \neg c) \rightarrow \neg m$ is not really an objective description of the social mobility process, but merely the countercultural belief that "you can't get ahead honestly." Moving this "constraint" to the second level of the LDT (along with S and N), we obtain

(19) $T = (\{m \rightarrow (w \vee c), c \rightarrow m\}, \{(w \wedge \neg c) \rightarrow \neg m\} \cup S_1 \cup N_1, X_1)$

which generates eight extensions. The first seven extensions are identical to those given in equation (11), with the countercultural belief accepted by innovators (first two extensions), ritualists (third extension), and retreatists (final four extensions). Indeed, acceptance of this belief *creates* the contradiction that necessitates these adaptations. But given our new interpretation of the constraint $((w \wedge \neg c) \rightarrow \neg m)$ as an "optional" belief rather than an objective constraint, we also obtain an eighth extension,

(20) $E = \{m \rightarrow (w \vee c), c \rightarrow m\} \cup S_1 \cup N_1 \cup \{w, \neg c\}$.

Given that this extension contains both norms in N_1, it reflects the possibility of *conformity* given rejection of the countercultural belief.

Multiple goals

We might also consider the possibility of multiple cultural goals. Merton recognized that money is not the only goal in American society, and asserted that alternative goals (such as "intellectual or artistic accomplishment") might reduce the

[14] In the present example, rankings of the elements in S and N can thus be handled simply by placing the propositions in these sets on different levels of the LDT. However, future research (addressing more complex examples) might consider generalizing the LDT framework, replacing the levels of the LDT with a partial order on the set of propositions in T (see Brewka, et al 1997, p 57).

strain toward deviant behavior (Merton 1957, p. 157). Of course, achievement of these alternative goals might also be blocked. But to confirm Merton's intuition, consider a simple case where Americans have two goals – money or family – where the latter is interpreted as an action can be achieved by everyone. If Americans are not obligated to achieve both goals simultaneously, we can respecify the first norm as

(21) $a \rightarrow (m \vee f)$

so that the individual's LDT becomes

(22) $T = (C_2, S_1 \cup N_2, X_2)$

where

(23) $N_2 = \{a \rightarrow (m \vee f), a \rightarrow (w \wedge \neg c)\}$,
$X_2 = \{w, \neg w, c, \neg c, f, \neg f\}$.

This LDT has the unique extension

(24) $E = C_2 \cup S_1 \cup N_2 \cup \{w, \neg c, f\}$

generating the unique pattern of action

(25) $A = E \cap X_2 = \{w, \neg c, f\}$.

Thus, following Merton's intuition, alternative (readily obtainable) goals promote conformity. Even if it is impossible to obtain money through work, the individual can uphold both the goals and means of the society by having a family.[15]

[15] If the individual faced better labor-market opportunities (so that the constraint set C_2 was replaced by C_1), the LDT would generate multiple extensions. The union $C_1 \cup S_1 \cup N_2$ remains consistent but no longer determines a unique pattern of action; the individual might choose either $\{w, \neg c, f\}$ or $\{w, \neg c, \neg f\}$.

Given that conformity can be so "easily" obtained through expansion of goals, one might wonder why deviance ever arises in real-world social systems. Perhaps some of the answer may be found in my incomplete specification or misspecification of constraints and norms. For instance, some individuals may be unable to have a family (so that the constraint $\neg f$ is added to the set C_2), or there may be tradeoffs between work and family (so that the constraint $w \rightarrow \neg f$ is added to C_2). Alternatively, in lieu of my disjunctive specification of norms ($a \rightarrow m \vee f$), one might suppose that norms are either conjunctive ($a \rightarrow m \wedge f$) or separate ($a \rightarrow m$, $a \rightarrow f$) so that having a family does not release Americans from the obligation to have money.[16] In any case, questions about the "ease" with which societies may alter norms to promote conformity might be regarded as macro-level concerns to be addressed through a specification of a macro-level process by which norms evolve. Given my current focus on the micro level, I assume that individuals take norms as given, analyzing how their response to particular configurations of norms without questioning the macro-level stability of those configurations.

Rebellion

Merton's presumption that individuals either accept or reject cultural goals and either accept or reject institutionalized means would seem to lead naturally to a 2×2 typology of individual adaptations, with the four cases labeled *conformity*, *innovation*, *ritualism*, and *retreatism*. The fifth adaptation of *rebellion* thus seems an awkward appendage, prompting critics to try to clarify and elaborate Merton's typology (Dubin 1959; Harary 1966; Besnard 1990). Merton himself conceded that the rebellion adaptation is "on a plane clearly different from the others" because it "represents a transitional response seeking to *institutionalize* new goals and new procedures to be

[16] In standard logic, the formula $a \rightarrow (m \wedge f)$ is equivalent to the conjunction $(a \rightarrow m) \wedge (a \rightarrow f)$. However, in a level default theory, it does matter whether norms are given an independent or conjunctive specification. To see this, compare the extensions of $T = (C_3, S_1, \{a \rightarrow (m \wedge f), a \rightarrow (w \wedge \neg c)\}, X_2)$ with those generated by $T = (C_3, S_1, \{a \rightarrow m, a \rightarrow f, a \rightarrow (w \wedge \neg c)\}, X_2)$. Although both LDTs generate extensions yielding the actions $\{\neg w, c, f\}$ and $\{w, \neg c, f\}$, the former also generates a third extension yielding the action $\{w, \neg c, \neg f\}$. Intuitively, given the latter LDT, rejection of the goal m does not require rejection of the goal f. In the former LDT, both goals are rejected together, causing the individual's choice between f and $\neg f$ to become underdetermined.

shared by other members of the society" and thus "refers to efforts to *change* the existing cultural and social structure rather than to accommodate efforts *within* this structure" (Merton 1957, p. 140, footnote 13, his emphases).

In his discussion of rebellion, Merton thus begins to connect micro-level adaptations to macro-level changes in the normative system. But retaining our narrow focus on micro-level adaptations, it is apparent that rebellion is "on a different plane" from the other adaptations because it presumes the introduction of new roles (along with their associated norms) into the social system. For concreteness, we might suppose that a new role *rebel* (denoted by r) is introduced into the social system, along with the norm that a *rebel* should neither work nor commit crime ($r \rightarrow (\neg w \wedge \neg c)$).[17] Extending Merton's example in this way, I will first examine the individual's behavior when her self-concept contains the *American* or *rebel* roles. I then consider whether any of these self-concepts are likely to be dynamically stable – whether any of these self-concepts is an "absorbing self."[18] While I do not attempt to specify the process by which new roles and norms are created *ex nihilo*, the following analysis nevertheless constitutes an important first step toward a more complete analysis of rebellion.

The individual's choices given the American or rebel self-concept

Extending Merton's example to include the *rebel* role and associated norm, the individual's LDT becomes

(26) $T = (C_2, S \cup N_3, X_1)$

where

[17] While many other conceptions of the *rebel* role seem possible, my present specification might be viewed as a stylized representation of one type of countercultural rejection of conventional American norms. We might further posit the obligation $r \rightarrow \neg m$, reflecting the rejection of conventional goals as well as means. However, this complication would have little effect on the results below because $\neg m$ is a logical consequence of $\neg w \wedge \neg c$ and the maintained constraint $m \rightarrow (w \vee c)$.

(27) $N_3 = \{a \rightarrow m,\ a \rightarrow (w \wedge \neg c),\ r \rightarrow (\neg w \wedge \neg c)\}$

and S is the individual's self-concept. Given that this self-concept may or may not contain the roles *American* and *rebel*, S is a member of the power set (i.e., the set of all subsets) of $\{a, r\}$. That is,

(28) $S \in \{\{a\}, \{r\}, \{a, r\}, \varnothing\}$.

If we continue to assume that $S = \{a\}$, we again obtain seven extensions:

(28) $\mathfrak{E} = \{C_2 \cup \{a\} \cup \{a \rightarrow m, r \rightarrow (\neg w \wedge \neg c)\} \cup \{\neg w, c\},$
$C_2 \cup \{a\} \cup \{a \rightarrow m, r \rightarrow (\neg w \wedge \neg c)\} \cup \{w, c\},$
$C_2 \cup \{a\} \cup \{a \rightarrow (w \wedge \neg c), r \rightarrow (\neg w \wedge \neg c)\} \cup \{w, \neg c\},$
$C_2 \cup N_3 \cup \{\neg w, \neg c\},$
$C_2 \cup N_3 \cup \{w, \neg c\},$
$C_2 \cup N_3 \cup \{\neg w, c\},$
$C_2 \cup N_3 \cup \{w, c\}\}.$

Essentially, this is the same result derived earlier in equation (11): the first two extensions correspond to innovation, the third to ritualism, and the final four to retreatism.[19] Intuitively, if the individual is not herself a *rebel*, the mere addition of this role to the social system does not affect her reasoning about her own obligations.

Alternatively, if the individual's self-concept is $S = \{r\}$, her LDT generates the unique extension

(29) $E = C_2 \cup \{r\} \cup N_3 \cup \{\neg w, \neg c\}$.

---

[18] This terminology is from Montgomery (2004), where change in the individual's self-concept is specified as a transition between states of a Markov chain. In a Markov chain, states are "absorbing" when they can never be exited following entry. Hence, stable long-run self-concepts are "absorbing selves."

[19] In contrast to my earlier results, each of these extensions now contains the norm $r \rightarrow (\neg w \wedge \neg c)$, but this formula is superfluous given that the LDT does not contain the formula *r*.

Thus, the individual who is solely a *rebel* is much like the *American* conformist who faces no contradiction in deriving her (unique) pattern of action. Indeed, we might well say that the individual is "conforming" to the *rebel* role.

If the individual's self-concept contains both roles so that S = {*a*, *r*}, the LDT generates the eight extensions:

(30) $\mathfrak{E} = \{C_2 \cup \{a\} \cup \{a \rightarrow m, r \rightarrow (\neg w \wedge \neg c)\} \cup \{\neg w, c\},$

$C_2 \cup \{a\} \cup \{a \rightarrow m, r \rightarrow (\neg w \wedge \neg c)\}\ \} \cup \{w, c\},$

$C_2 \cup \{a\} \cup \{a \rightarrow (w \wedge \neg c), r \rightarrow (\neg w \wedge \neg c)\} \cup \{w, \neg c\},$

$C_2 \cup \{a, r\} \cup \{a \rightarrow m\} \cup \{\neg w, c\},$

$C_2 \cup \{a, r\} \cup \{a \rightarrow m\} \cup \{w, c\},$

$C_2 \cup \{a, r\} \cup \{a \rightarrow (w \wedge \neg c)\} \cup \{w, \neg c\},$

$C_2 \cup \{a, r\} \cup \{r \rightarrow (\neg w \wedge \neg c)\} \cup \{\neg w, \neg c\},$

$C_2 \cup \{r\} \cup N_3 \cup \{\neg w, \neg c\}\}.$

Given two roles in the individual's self-concept, we might apply Merton's basic typology (composed of his first four adaptations) to each role separately. With respect to the *American* role, extensions 1, 2, 4, and 5 correspond to innovation (rejection of the means $a \rightarrow (w \wedge \neg c)$), extensions 3 and 6 corresponds to ritualism (rejection of the goal $a \rightarrow m$), extension 7 corresponds to Merton's original conception of retreatism (where the *American* role is retained but both the goal and means are rejected), and extension 8 corresponds to my alternative conception of retreatism (where the *American* role is rejected but both the goal and means are retained).[20] With respect to the *rebel* role, extensions 1, 2, and 3 correspond to retreatism (rejection of the *rebel* role), while extensions 4, 5, and 6 correspond to innovation (rejection of the means $r \rightarrow (\neg w \wedge \neg c)$).[21] Extensions 7 and 8 reflect conformity to the *rebel* role.

Finally, suppose that the individual's self-concept contains neither role, so that

[20] Note that, in the present example, both forms of retreatism are observationally equivalent, inducing the same pattern of action $\{\neg w, \neg c\}$.

[21] Given that I did not a specify a goal for the *rebel* role, extensions 4 through 6 might alternatively be understood as a form of retreatism (where the *rebel* role is maintained but means are rejected).

$S = \varnothing$. In this case, the LDT generates four extensions:

(31) $\mathfrak{E} = \{C_2 \cup N_3 \cup \{w, c\},$
$C_2 \cup N_3 \cup \{w, \neg c\},$
$C_2 \cup N_3 \cup \{\neg w, c\},$
$C_2 \cup N_3 \cup \{\neg w, \neg c\}\}.$

Intuitively, if there are no roles in the self-concept, the individual is not forced to reject any norms. On the other hand, because the individual has no reason to choose any particular action, every feasible pattern of action becomes possible.

Dynamics of the self-concept

Having just considered how an individual would act given an exogenous self-concept, we might now attempt to specify the process governing change in the individual's self-concept. Here, I develop a simple iterative approach.[22] Suppose that the individual begins with an initial self-concept S_t, and then selects one of the extensions generated by the LDT

(32) $T = (C_2, S_t \cup N_3, X_1).$

The roles contained in the selected extension determine the individual's revised self-concept S_{t+1}.[23] In the next period, selecting one of the extensions from the LDT

(33) $T = (C_2, S_{t+1} \cup N_3, X_1),$

the individual's self-concept S_{t+2} is determined. Continuing in this way, we obtain a

[22] Montgomery (2004) develops a more elaborate "looking-glass self" model of self-concept formation that generates somewhat different results. See Appendix 2 for further discussion and an alternative analysis of the present example.

[23] Formally, letting E_t denote the selected extension in period t, the individual's self-concept in period t+1 is $S_t = E_t \cap \{a, r\}$.

sequence of self-concepts (S_t, S_{t+1}, S_{t+2}, …) characterizing the individual's trajectory through "role space" (Montgomery 2000).

Given this simple iterative model of self-concept dynamics, the results of the previous subsection imply that the self-concept $\{r\}$ and the empty self-concept are stable long-run outcomes ("absorbing selves"). To see this, recall that the self-concept $\{r\}$ generates a unique extension (given in equation 29) that contains only the *rebel* role. Consequently, once the individual becomes a *rebel*, she would never abandon this role nor adopt the *American* role. Similarly, while the empty self-concept generates four extensions (given in equation 31), note that none of these extensions contains any roles. Consequently, once the individual has an empty self-concept, her self-concept would remain empty. In contrast, the self-concepts $\{a\}$ and $\{a, r\}$ are not stable long-run outcomes. While *Americans* who are innovators or ritualists (adopting one of the first three extensions in equation 28) would remain an *American* in the subsequent period, those who are retreatists (adopting any of the last four extensions) would abandon their *American* identity (transitioning into the absorbing empty self-concept). Thus, given any (arbitrarily small) chance of the retreatism adaptation, $\{a\}$ is not an absorbing self-concept. Similarly, an individual initially holding both identities might temporarily retain the self-concept $\{a, r\}$ if she selected extension 4, 5, 6, or 7 from equation (30). But given any chance of selecting one of the other extensions, she will eventually transition to either the $\{a\}$ or $\{r\}$ self-concept.

Possible transitions between self-concepts are thus characterized by the arrows in Figure 3. A self-concept is absorbing if there is no arrow leading away from (and hence no possible transition out of) that self-concept. Thus, Figure 3 shows that the {*rebel*} self-concept is absorbing while the {*American*} self-concept is transitory. More generally, a particular self-concept is absorbing only if it is possible to conform to every role contained in that self-concept. Thus, for an individual who lacks access to labor-market opportunities, the {*American*, *rebel*} and {*American*} self-concepts are not absorbing. While innovation and ritualism are short-term adaptations that temporarily allow the individual to retain her *American* identity, retreatism ultimately leads to role

abandonment and a new self-concept.[24]

The preceding analysis assumed that the individual lacked labor-market opportunities. Repeating the analysis for an individual with labor-market opportunities who faces the constraint set C_1 rather than C_2, the self-concept $\{a\}$ now becomes a stable long-run outcome. This is because the LDT

(34) $T = (C_1, \{a\} \cup N_3, X_1)$

generates the unique extension

(35) $E = C_1 \cup \{a\} \cup N_3 \cup \{w, \neg c\}$

which contains only the *American* role. Thus, given labor-market opportunities, there are three absorbing selves: $\{a\}$, $\{r\}$, and the empty self-concept. The $\{a, r\}$ self-concept remains unstable because it remains impossible to conform to both roles simultaneously.[25]

Summary and discussion

Non-monotonic logic permits an elegant formalization of Merton's micro-level analysis. In the absence of cultural contradictions, the individual's constraints and self-concept and the norms of the social system (comprising the union $C \cup S \cup N$) are logically consistent. Thus, the individual can retain her current self-concept while adhering to socially approved goals and means. Formally, every extension of the individual's level default theory contains every formula in S and N. But given cultural contradictions (which arise when $C \cup S \cup N$ is inconsistent), the individual can adapt though innovation (rejection of means in N), ritualism (rejection of goals in N), or

[24] Because it permits role adoption as well as role loss, the model of self-concept dynamics in Montgomery (2004) leads to a somewhat different conclusion. See Appendix 2.

[25] Having substituted constraint set C_1 for the set C_2, the only change to Figure 3 would be the removal of the arrow from {*American*} to $\varnothing$. The loss of this potential transition implies not only that the {*American*} self-concept is now stable, but also that an individual would have an empty self-concept only if her initial self-concept was empty.

retreatism (rejection of roles in S). Formally, these adaptations correspond to different extensions of the individual's level default theory – sets of formulas that are as "consistent as possible" given the contradiction facing the individual.

Beyond merely restating Merton's ideas using non-monotonic logic, my formal analysis also helps to clarify and extend his theory. It suggests that retreatism is best understood not as a particular pattern of behavior, but as the loss of a role from the individual's self-concept. In this way, my analysis highlights the importance of the self-concept in Merton's theory. Relatedly, I have suggested that rebellion – often regarded by critics as an awkward addition to Merton's typology – can be partly understood as change in the self-concept arising from adaptation to cultural contradictions. The simple iterative model in the text (and the more elaborate model in Appendix 2) demonstrates how my logical specification of the decision-making process can serve as the foundation for a dynamic model of self-concept formation, permitting an micro-level analysis of the long-run stability (or instability) of different self-concepts.

Given the widely recognized ambiguities in Merton's theory (Besnard 1990) and the influence that this theory still possesses in subfields such as crime and deviance (Adler and Laufer 1995; Passas and Agnew 1997), my present formalization effort should not be dismissed as a sterile exercise in the history of sociological thought. Nevertheless, many contemporary sociologists might still view Merton's theory (and hence my logical formalization) as an outdated perspective on agency. In contrast, my own interest in both Merton's theory and non-monotonic logic stems from my conviction that this theory and these formalisms will provide a useful foundation for clarifying and elaborating more recent perspectives. While further formalization efforts are needed to build a compelling case, my remaining comments suggest some promising directions for future research.

Both older and more recent studies of the urban poor have often explicitly invoked Merton's typology of adaptations (see, e.g., Hannerz 1969, p. 215; Duneier 1999, p. 364). Especially given Duneier's (1999) emphasis on the attempt by sidewalk vendors to lead moral lives, the model developed in the present paper seems well-suited to help clarify and generalize his account of the choices made by these men.[26] In other subfields where Merton's theory is less often acknowledged, researchers nevertheless

[26] See also my formalization of Liebow's (1967) account in Montgomery (2004).

maintain an explicit focus on cultural contradictions and individual adaptations. For instance, consider recent studies addressing the conflict between the roles of mother and worker (Moen 1992; Hays 1996; Macdonald 2003). Using the theoretical method illustrated in the present paper, researchers could clarify the contradictions that exist, explain the choices made by women in the face of these contradictions, and explain changes in a woman's self-concept resulting from these adaptations.[27]

In my discussion of retreatism, I noted that the current model permits both "rational" and "rationalizing" perspectives on the decision-making process. The former assumes that the individual selects a particular extension of her LDT and then chooses the corresponding pattern of action, while the latter assumes the individual selects some defensible pattern of action while retaining every extension consistent with those actions. Perhaps the coexistence of multiple motivations for own behavior is the key to the type of "robust action" described by Padgett and Ansell (1993). Their analysis of Cosimo de Medici emphasizes his "multivocality" which allowed different observers to make different attributions about Cosimo's behavior.[28] But more provocatively, they also hint that Cosimo himself must have maintained multiple, conflicting interpretations of his own behavior.[29] Thus, adopting the "rationalizing" perspective on decision-making, the present model (perhaps in conjunction with the model of the attribution process found in Montgomery 2004) might allow rigorous development of the theory of robust action.

More generally, recognizing that different extensions of a level default theory correspond to different accounts of behavior, future research might begin to clarify perspectives on agency (found especially in the sociology of culture) which emphasize the need for individuals to provide accounts of their actions (Scott and Lyman 1968) and the construction of these accounts from a cultural "tool kit" (Swidler 1986). For example, in his study of the work choices made by Americans, Wuthnow (1996, p. 93)

---

[27] Moreover, Macdonald's (2003) account (which examines interaction between mothers and nannies) suggests an extension of the present model in which two individuals make choices interactively, each creating constraints for the other. This extension would, in essence, represent a role-theoretic alternative to game theory in which choices and self-concepts of multiple actors are jointly determined.

[28] Multivocality is defined as "the fact the single actions can be interpreted coherently from multiple perspectives simultaneously, the fact that single actions can be moves in many games at once, and the fact that public and private motivations cannot be parsed" (Padgett and Ansell 1993, p. 1263).

[29] "Victory, in Florence, in chess, or *go* means locking in others, *but not yourself*, to goal-oriented sequences of strategic play that become predictable thereby" (Padgett and Ansell 1993, p. 1264, my emphasis).

argues that "in order to keep working people must be able to give a legitimate account of themselves."[30] To formalize this idea, we might assume that the individual chooses *work* only if the extensions of her level default theory which lead to this choice would also support any account(s) of behavior that she may need to provide to observers. In this way, the sets of propositions in these extensions do become, in essence, a "tool kit" or "ragbag" from which accounts can be constructed (cf. Wuthnow 1996, p. 95 ff).

Further generalization of the present model might help clarify other suggestive metaphors in Swidler (1986). Given formal representation of ideologies as sets of logical formulas, we could explain how individuals use these "chunks of culture" to derive courses of action in "unsettled" cultural periods. Generalizing the model to allow habit formation, we might explain how culture goes "underground" during "settled" cultural periods.[31] And while Swidler's discussion of "unsettled lives" and "settled lives" seems to conflate micro-level and macro-level processes, micro-level transitions between settled and unsettled selves might be addressed using some version of the dynamic models of self-concept formation offered in the present paper.

[30] Equivalently, he argues that "people work if they can give a legitimate account of what they are doing" and that "having a legitimate account of oneself makes it possible to work." (Wuthnow 1996, p. 93).
[31] To model habit formation, we might assume that, after adhering to the norm *role* $\rightarrow$ *action* for some length of time, the (non-contingent) norm *action* is added to the individual's LDT. In this way, the individual would continue to (try to) choose *action* even if *role* was removed from her self-concept.

Appendix 1. Further discussion of level default theories

This appendix briefly reviews some basic concepts from formal logic, and then shows how to derive the extensions of an LDT using a truth table. Readers might consult an introductory textbook (e.g., Lemmon 1965) for more thorough treatment of propositional logic and truth tables; Brewka et al (1997) offer more discussion of LDTs and other non-monotonic logics.

Logical formulas are constructed from atomic propositions (a, b, c, …) and the logical connectives *and* ($\wedge$), *or* ($\vee$), *not* ($\neg$), and *implies* ($\rightarrow$). Note that a formula might be composed of a single atomic proposition (e.g., the formula a) or involve several atomic propositions and logical connectives (e.g., the formula $a \rightarrow (b \wedge c)$). Each atomic proposition can be either true or false (i.e., each proposition has a "truth value" equal to T or F). Given an assignment of truth values to the atomic propositions, we can use the definitions of the logical connectives to determine the truth value of the entire logical formula. Letting P and Q denote arbitrary propositions (which might be either atomic propositions or more complex formulas), the logical connectives are defined by the following truth tables.

P	$\neg P$
T	F
F	T

P	Q	$P \wedge Q$	$P \vee Q$	$P \rightarrow Q$
T	T	T	T	T
T	F	F	T	F
F	T	F	T	T
F	F	F	F	T

Logical connectives are sometimes called "truth functions" given that they map the truth values of inputs into the truth value of an output. Note that the *not* function maps a single input (the truth value of P) into a single output (the truth value of $\neg P$), while each of the other truth functions (e.g., the *and* function) maps a pair of inputs (the truth values of P and Q) into a single output (the truth value of $P \wedge Q$).

To illustrate how the truth values of the atomic propositions determine the truth values of a logical formula, consider the formula $a \rightarrow (b \wedge c)$. Given that this formula involves three atomic propositions, there are $2^3 = 8$ possible assignments of truth values to the atomic propositions, corresponding to the 8 rows of the following truth table.

a	b	c	$b \wedge c$	$a \rightarrow (b \wedge c)$
T	T	T	T	T
T	T	F	F	F
T	F	T	F	F
T	F	F	F	F
F	T	T	T	T
F	T	F	F	T
F	F	T	F	T
F	F	F	F	T

For each row of the truth table, we may first use the truth values of b and c to determine the truth value of the formula $b \wedge c$. Then, using the truth values of a and $(b \wedge c)$, we may determine the truth value of the formula $a \rightarrow (b \wedge c)$.

Given this background, we may now use a truth table to derive the extensions of a level default theory. Although this truth-table method grows cumbersome as the number of atomic propositions increases, it may be especially helpful for readers less background in logic, and also provides a mechanical procedure that can easily be implemented as a computer algorithm. To illustrate the truth-table method, consider the LDT

(A1.1) $T = (C_2, S_1 \cup N_1, X_1)$

that generates the seven extensions given in equation (11). Writing out each of the formulas in the three levels of this LDT, we obtain

(A1.2) $C_2 = \{m \rightarrow (w \vee c), c \rightarrow m, (w \wedge \neg c) \rightarrow \neg m\}$,

$S_1 \cup N_1 = \{a, a \rightarrow m, a \rightarrow (w \wedge \neg c)\}$,

$X_1 = \{w, \neg w, c, \neg c\}$.

To find the extensions of this LDT, we must first determine whether the formulas in the first level (i.e., the formulas in the set C_2) are logically consistent. By definition, a set of formulas Φ is *logically consistent* if and only if there is some assignment of truth values to the atomic propositions in Φ such that every formula in Φ is true. Given the four atomic propositions (a, m, w, c) appearing in equation (A1.2), there are $2^4 = 16$ possible assignments of truth values, corresponding to the 16 rows in Figure 4. Inspection of this truth table reveals that there are, in fact, assignments of truth values to the atomic propositions such that all formulas in C_2 are true. In particular, all three formulas in C_2 are true given the assignments of truth values in rows 1, 3, 6, 8, 9, 11, 14, and 16. Thus, the formulas in the first level are consistent, and all of these formulas must be contained in every extension of the LDT.

Next, we must determine whether the entire second level ($S_1 \cup N_1$) will be contained every extension. That is, we must determine whether the union $C_2 \cup S_1 \cup N_1$ is logically consistent. Inspection of the truth table now reveals that there is no assignment of truth values (i.e., no row of the truth table) such that all of the formulas in the first two levels are true. Thus, the union $C_2 \cup S_1 \cup N_1$ is inconsistent. Consequently, any extension of the LDT will contain all formulas in C_2 and a maximal consistent subset of formulas in $C_2 \cup S_1 \cup N_1$. Restricting attention to those rows of the table for which all first-level formulas are true, rows 1, 3, 6, and 8 reveal that $C_2 \cup \{a \rightarrow m, a \rightarrow (w \wedge \neg c)\}$ is consistent; row 11 reveals that $C_2 \cup \{a, a \rightarrow (w \wedge \neg c)\}$ is consistent; rows 14 and 16 reveal that $C_2 \cup \{a, a \rightarrow m\}$ is consistent. In this way, we obtain the three partial extensions

(A1.3) $C_2 \cup \{a \rightarrow m, a \rightarrow (w \wedge \neg c)\}$,

$C_2 \cup \{a, a \rightarrow (w \wedge \neg c)\}$,

$C_2 \cup \{a, a \rightarrow m\}$.

While row 9 reveals that $C_2 \cup \{a\}$ is consistent, note that $C_2 \cup \{a\}$ is not a *maximal* consistent subset of $C_2 \cup S_1 \cup N_1$ because it is contained in two of the other subsets, $C_2 \cup \{a, a \rightarrow m\}$ and $C_2 \cup \{a, a \rightarrow (w \wedge \neg c)\}$, already identified.

To complete our derivation, we must add as many third-level formulas as consistently possible to each of the partial extensions in equation (A1.3). Inspection reveals that each of the seven relevant rows of the truth table (rows 1, 3, 6, 8, 11, 14, 16) corresponds to a maximal consistent subset of $C_2 \cup S_1 \cup N_1 \cup X_1$. Thus, we can add either $\{\neg w, \neg c\}$, $\{w, \neg c\}$, $\{\neg w, c\}$, or $\{w, c\}$ to the first partial extension; we can add $\{w, \neg c\}$ to the second partial extension; we can add either $\{\neg w, c\}$ or $\{w, c\}$ to the third partial extension. This yields the seven extensions given in equation (11).

Appendix 2. An alternative model of self-concept dynamics

In the text, I posited a simple iterative model of self-concept dynamics. Montgomery (2004) offers a more elaborate model in which an individual's self-concept gradually conforms to the view of the self held by others. In that model, the individual chooses actions based on her current self-concept. Observers then make attributions (i.e., inferences about the roles contained in the individual's self-concept) which are gradually internalized by the individual. Thus, a feedback loop runs from the self-concept to actions to attributions back to the self-concept. A self-concept S is an "absorbing self" if it is a fixed point of this feedback loop, generating a pattern of action A(S) that causes observers to make the unique attribution S′(A(S)) which is equal to S. Thus, unlike the simple iterative model in the text, the model in Montgomery (2004) incorporates a feedback mechanism that may produce a "looking-glass self" (Cooley 1983 [1902]).

This appendix shows how the model in Montgomery (2004) would be applied to the example in the present paper. To begin, suppose the self-concept is placed before norms in the individual's LDT so that

(A2.1) $T = (C_2, S, N_3, X_1)$

given the self-concept $S \in \{\{a, r\}, \{a\}, \{r\}, \varnothing\}$. For each of these possible selves, we can derive the extensions of the LDT and hence the pattern of actions that the individual might choose. The self-concept $S = \{a, r\}$ generates four extensions:

(A2.2) $\mathfrak{E} = \{C_2 \cup \{a, r\} \cup \{a \rightarrow m\} \cup \{w, c\}$

$C_2 \cup \{a, r\} \cup \{a \rightarrow m\} \cup \{\neg w, c\}$

$C_2 \cup \{a, r\} \cup \{a \rightarrow (w \wedge \neg c)\} \cup \{w, \neg c\}$

$C_2 \cup \{a, r\} \cup \{r \rightarrow (\neg w \wedge \neg c)\} \cup \{\neg w, \neg c\}\}$.

The self-concept $S = \{a\}$ generates three extensions:

(A2.3) $\mathfrak{E} = \{C_2 \cup \{a\} \cup \{a \rightarrow m, r \rightarrow (\neg w \wedge \neg c)\} \cup \{\neg w, c\}$,

$$C_2 \cup \{a\} \cup \{a \rightarrow m,\ r \rightarrow (\neg w \wedge \neg c)\} \cup \{w, c\},$$
$$C_2 \cup \{a\} \cup \{a \rightarrow (w \wedge \neg c),\ r \rightarrow (\neg w \wedge \neg c)\} \cup \{w, \neg c\}\}.$$

The self-concept S = $\{r\}$ generates the unique extension

(A2.4) $\quad E = C_2 \cup \{r\} \cup N_3 \cup \{\neg w, \neg c\}.$

The empty self-concept S = $\varnothing$ generates the extensions

(A2.5) $\quad \mathfrak{E} = \{C_2 \cup N_3 \cup \{w, c\},$
$$C_2 \cup N_3 \cup \{w, \neg c\},$$
$$C_2 \cup N_3 \cup \{\neg w, c\},$$
$$C_2 \cup N_3 \cup \{\neg w, \neg c\}\}.$$

The relation from self-concepts to possible actions may thus be characterized by the binary matrix

(A2.6) $\quad \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

where the rows of **A** correspond to the self-concepts ($\{a, r\}$, $\{a\}$, $\{r\}$, $\varnothing$) and the columns correspond to the patterns of action ($\{w, c\}$, $\{w, \neg c\}$, $\{\neg w, c\}$, $\{\neg w, \neg c\}$).

Now consider the attributions that might be made by observers – the self-concepts that might be attributed to the individual on the basis of observed outcomes and actions. Suppose that observers make these attributions using the LDT

(A2.7) $\quad T' = (A, N_3', \{a, \neg a, r, \neg r\})$

where A is the set of outcomes and actions observed, and N_3' is the set of equivalences

(A2.8) $N_3' = \{(a \leftrightarrow m, a \leftrightarrow (w \wedge \neg c), r \leftrightarrow (\neg w \wedge \neg c)\}$

corresponding to the set of norms N_3. As discussed in Montgomery (2004), inclusion of the set N_3' in this LDT reflects the tendency for observers to reason abductively, making (logically problematic) inferences of the form *action* and *role* → *action* therefore *role*. In the present example, four possible patterns of outcomes and actions might be observed:

(A2.9) $A \in \{\{m, w, c\}, \{\neg m, w, \neg c\}, \{m, \neg w, c\}, \{\neg m, \neg w, \neg c\}\}$.

Considering each of these cases, the set of observables $A = \{m, w, c\}$ generates two extensions:

(A2.10) $\mathfrak{E} = \{\{m, w, c\} \cup \{a \leftrightarrow (w \wedge \neg c), r \leftrightarrow (\neg w \wedge \neg c)\} \cup \{\neg a, \neg r\},$
$\{m, w, c\} \cup \{a \leftrightarrow m, r \leftrightarrow (\neg w \wedge \neg c)\} \cup \{a, \neg r\}\}$.

The set of observables $A = \{\neg m, w, \neg c\}$ generates two extensions:

(A2.11) $\mathfrak{E} = \{\{\neg m, w, \neg c\} \cup \{a \leftrightarrow (w \wedge \neg c), r \leftrightarrow (\neg w \wedge \neg c)\} \cup \{a, \neg r\},$
$\{\neg m, w, \neg c\} \cup \{a \leftrightarrow m, r \leftrightarrow (\neg w \wedge \neg c)\} \cup \{\neg a, \neg r\}\}$.

The set of observables $A = \{m, \neg w, c\}$ generates two extensions:

(A2.12) $\mathfrak{E} = \{\{m, \neg w, c\} \cup \{a \leftrightarrow (w \wedge \neg c), r \leftrightarrow (\neg w \wedge \neg c)\} \cup \{\neg a, \neg r\},$
$\{m, \neg w, c\} \cup \{a \leftrightarrow m, r \leftrightarrow (\neg w \wedge \neg c)\} \cup \{a, \neg r\}$.

The set of observables $A = \{\neg m, \neg w, \neg c\}$ generates the unique extension:

(A2.13) $\mathfrak{E} = \{\neg m, \neg w, \neg c\} \cup N_3' \cup \{\neg a, r\}$.

The relation from sets of observables to potential attributions may thus be characterized by the binary matrix

$$\text{(A2.14)} \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

where the rows of **B** correspond to the patterns of action ($\{w, c\}$, $\{w, \neg c\}$, $\{\neg w, c\}$, $\{\neg w, \neg c\}$) and the columns correspond to the self-concepts ($\{a, r\}$, $\{a\}$, $\{r\}$, $\varnothing$).

The Boolean product of the matrices **A** and **B** gives the possible transitions between self-concepts:

$$\text{(A2.15)} \quad \mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

where both the rows and columns of this matrix correspond to the self-concepts ($\{a, r\}$, $\{a\}$, $\{r\}$, $\varnothing$). These transitions between self-concepts are also characterized by the arrows in Figure 5. Note that {*rebel*} is an absorbing self given either the simple iterative model of self-concept dynamics in the text (Figure 3) or the more elaborate model applied here (Figure 5). But given the present model, the empty self is no longer absorbing. Intuitively, given that the empty self-concept might lead the individual to choose any feasible pattern of action, observers might attribute either the {*American*} or {*rebel*} self-concept to the individual. In the short run, the individual might repeatedly transition in and out of the *American* identity, moving between the {*American*} and empty self-concepts. But {*rebel*} is now the only self-concept stable in the long run.

References

Adler, Freda and William S. Laufer, eds. (1995) *The Legacy of Anomie Theory*. New Brunswick, NJ: Transaction Publishers.

Åqvist, Lennart (1987) *Introduction to Deontic Logic and the Theory of Normative Systems.* Naples, Italy: Bibliopolis.

Besnard, Philippe (1990) "Merton in Search of Anomie," pp. 243-54 in *Robert K. Merton: Consensus and Controversy*, edited by Jon Clark, Celia Modgil, and Sohan Modgil. New York: Falmer Press.

Brewka, Gerhard, Jurgen Dix, and Kurt Konolige (1997) *Non-Monotonic Reasoning: An Overview*. Stanford, CA: Center for the Study of Language and Information.

Chellas, Brian F. (1980) *Modal Logic: An Introduction.* Cambridge University Press.

Cole, Stephen (1975) "The Growth of Scientific Knowledge: Theories of Deviance as a Case Study," pp. 175-220 in *The Idea of Social Structure: Papers in Honor of Robert K. Merton*, edited by Lewis A. Coser. New York: Harcourt Brace Jovanovich.

Cooley, Charles Horton (1983 [1902]) *Human Nature and the Social Order*. New Brunswick, NJ: Transaction Publishers.

Dubin, Robert (1959) "Deviant Behavior and Social Structure: Continuities in Social Theory," *American Sociological Review* 24:147-164.

Duneier, Mitchell (1999) *Sidewalk.* New York: Farrar, Straus and Giroux.

Featherstone, Richard and Mathieu Deflem (2003) "Anomie and Strain: Context and Consequences of Merton's Two Theories," *Sociological Inquiry* 73:471-89.

Forrester, James Wm. (1996) *Being Good & Being Logical*. Armonk, NY: M.E. Sharp.

Hannan, Michael T. (1998) "Rethinking Age Dependence in Organizational Mortality: Logical Formalizations," *American Journal of Sociology* 104:126-64.

Hannerz, Ulf (1969) *Soulside: Inquiries into Ghetto Culture and Community*. New York: Columbia University Press.

Harary, Frank (1966) "Merton Revisited: A New Classification for Deviant Behavior," *American Sociological Review* 31:693-697.

Hays, Sharon (1996) *The Cultural Contradictions of Motherhood.* New Haven: Yale University Press.

Hilpinen, Risto, ed. (1971) *Deontic Logic: Introductory and Systematic Readings.* Dordrecht, Holland: D. Reidel Publishing Company.

Horty, John F. (1993) "Deontic Logic as Founded on Nonmonotonic Logic," *Annals of Mathematics and Artificial Intelligence* 9:69-91.

------ (1994) "Moral Dilemmas and Nonmonotonic Logic," *Journal of Philosophical Logic* 23:35-65.

Hughes, G. E. and M. J. Cresswell (1996) *A New Introduction to Modal Logic.* New York: Routledge.

Kamps, Jaap and László Pólos (1999) "Reducing Uncertainty: A Formal Theory of Organizations in Action," *American Sociological Review* 104:1774-1810.

Lemmon, E. J. (1965) *Beginning Logic.* London: Chapman and Hall.

Levine, Donald N. (1985) *The Flight From Ambiguity: Essays in Social and Cultural Theory*. Chicago: University of Chicago Press.

Liebow, Elliot (1967) *Tally's Corner*. Boston: Little, Brown.

Macdonald, Cameron L. (2003) *Shadow Mothers: Nannies, Au Pairs, and the Social Construction of Mothering*, unpublished book manuscript.

Messner, Steven F. (1988) "Merton's Social Structure and Anomie: The Road Not Taken," *Deviant Behavior* 9:33-53.

Merton, Robert K. (1938) "Social Structure and Anomie," *American Sociological Review* 3:672-682.

------ (1957) *Social Theory and Social Structure.* New York: Free Press.

Moen, Phyllis (1992) *Women's Two Roles: A Contemporary Dilemma*. New York: Auburn House.

Montgomery, James D. (2000) "The Self as a Fuzzy Set of Roles, Role Theory as a Fuzzy System," *Sociological Methodology* 30:261-314.

------ (2004) "The Logic of Role Theory: Role Conflict and Stability of the Self-Concept," revised and resubmitted to *Journal of Mathematical Sociology*.

Nute, Donald, ed. (1997) *Defeasible Deontic Logic.* Boston: Kluwer.

Orrù, Marco (1990) "Merton's Instrumental Theory of Anomie," pp. 231-40 in *Robert K. Merton: Consensus and Controversy*, edited by Jon Clark, Celia Modgil, and Sohan Modgil. New York: Falmer Press.

Padgett, John F. and Christopher K. Ansell (1993) "Robust Action and the Rise of the Medici, 1400-1434," *American Journal of Sociology* 98:1259-1319.

Passas, Nikos, and Robert Agnew, eds. (1997) *The Future of Anomie Theory*. Boston: Northeastern University Press.

Péli, Gábor, Jeroen Bruggeman, Michael Masuch, and Breanndán P. Nualláin (1994) "A Logical Approach to Formalizing Organizational Ecology," *American Sociological Review* 59:571-593.

Pólos, Lázsló and Michael T. Hannan (2002) "Reasoning with Partial Knowledge," *Sociological Methodology* 32:133-181.

Reiter, R. (1980) "A Logic for Default Reasoning," *Artificial Intelligence* 13:81-132.

Scott, Marvin B. and Stanford M. Lyman (1968) "Accounts," *American Sociological Review* 33:46-62.

Swidler, Ann (1986) "Culture in Action: Symbols and Strategies," *American Sociological Review* 51:273-86.

Sztompka, Piotr (1998) "Robert K. Merton's Four Concepts of Anomie," pp. 163-75 in *Robert K. Merton and Contemporary Sociology*, edited by Carlo Mongardini and Simonetta Tabboni. New Brunswick, NJ: Transaction Publishers.

Wuthnow, Robert (1996) *Poor Richard's Principle: Recovering the American Dream Through the Moral Dimension of Work, Business, and Money.* Princeton, NJ: Princeton University Press.

Figure 1. Propositional inventory

proposition	*logical formula*
P1. *I am an American*.	a
P2. *An American ought to have money*.	$a \rightarrow m$
P3. *An American ought to work and not commit crime*.	$a \rightarrow (w \wedge \neg c)$
P4. *Money can be acquired only through work or crime*.	$m \rightarrow (w \vee c)$
P5. *Crime generates money*.	$c \rightarrow m$
P6. *Work generates money.*	$w \rightarrow m$
P6.′ *Work (without crime) does not generate money*.	$(w \wedge \neg c) \rightarrow \neg m$

Figure 2. Typology of individual adaptations

modes of adaptation	*cultural goals*	*institutionalized means*
I. Conformity	+	+
II. Innovation	+	–
III. Ritualism	–	+
IV. Retreatism	–	–
V. Rebellion	±	±

Source: Merton (1957). Following Merton (1957, p 140), + denotes "acceptance," – denotes "rejection," and ± denotes "rejection of prevailing values and substitution of new values."

Figure 3. Possible transitions between self-concepts

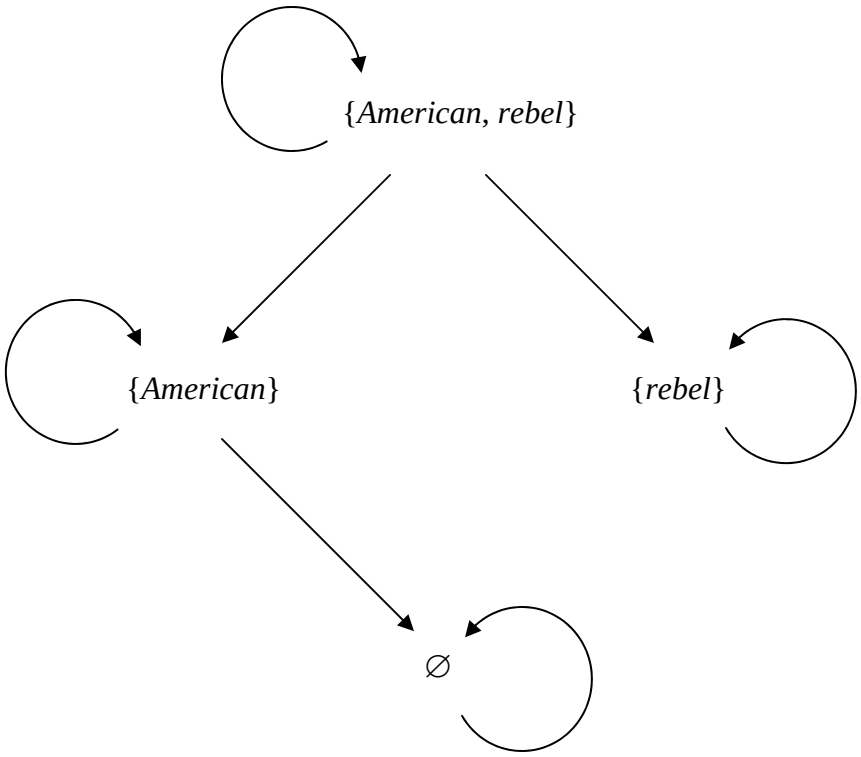

Figure 4. Truth table for the level default theory $T = (C_2, S_1 \cup N_1, X_1)$

	atomic propositions				*formulas in first level*			*formulas in second level*			*formulas in third level*			
	a	m	w	c	$m\rightarrow$ $(w\vee c)$	$c\rightarrow m$	$m\rightarrow$ $(w\wedge\neg c)$	a	$a\rightarrow m$	$a\rightarrow$ $(w\wedge\neg c)$	w	$\neg w$	c	$\neg c$
1.	F	F	F	F	T	T	T	F	T	T	F	T	F	T
2.	F	F	F	T	T	F	T	F	T	T	F	T	T	F
3.	F	F	T	F	T	T	T	F	T	T	T	F	F	T
4.	F	F	T	T	T	F	T	F	T	T	T	F	T	F
5.	F	T	F	F	F	T	T	F	T	T	F	T	F	T
6.	F	T	F	T	T	T	T	F	T	T	F	T	T	F
7.	F	T	T	F	T	T	F	F	T	T	T	F	F	T
8.	F	T	T	T	T	T	T	F	T	T	T	F	T	F
9.	T	F	F	F	T	T	T	T	F	F	F	T	F	T
10.	T	F	F	T	T	F	T	T	F	F	F	T	T	F
11.	T	F	T	F	T	T	T	T	F	T	T	F	F	T
12.	T	F	T	T	T	F	T	T	F	F	T	F	T	F
13.	T	T	F	F	F	T	T	T	T	F	F	T	F	T
14.	T	T	F	T	T	T	T	T	T	F	F	T	T	F
15.	T	T	T	F	T	T	F	T	T	T	T	F	F	T
16.	T	T	T	T	T	T	T	T	T	F	T	F	T	F

Figure 5. Possible transitions between self-concepts in Appendix 2