A Theory of Shared Values

(Preliminary and incomplete draft)

Elizabeth Maggie Penn*

March 14, 2011

Abstract

In this paper I model groups as collections of individuals who attach value to similar modes of behavior, or ways of life. These values serve as a group heuristic for evaluating the merit of both the familiar and potentially strange or unfamiliar behavior of others. The model is used to examine why some group-specific behaviors persist unchanged in diverse societies, while some behaviors may face pressure to converge to, or diverge from, the behaviors of other groups. The model also offers insight into the relative strengths and weaknesses of different group values from two perspectives: how the group evaluates itself relative to others, and how others evaluate the group with respect to their own heuristics.

*Associate Professor of Political Science, Washington University in St. Louis. Email: penn@wustl.edu. Thanks to Randy Calvert, John Patty and Moses Shayo for helpful comments.

> "[W]e are cultural beings, endowed with the capacity and the will to take a deliberate attitude toward the world and to lend it significance. Whatever this significance may be, it will lead us to judge certain phenomena of human existence in its light and to respond to them as being (positively or negatively) meaningful."
>
> —Max Weber, *Objectivity and Understanding in Economics.*

1 Introduction

It is not difficult to consider a situation in which two different individuals watching the same show, listening to the same debate, or engaging in a conversation with each other have markedly different perceptions of these shared experiences. People take an active role in interpreting their own experiences, and in assigning meaning to those experiences. While these differences can be positive, for example by broadening all individuals' conception of what constitutes a correct way of life, such differences may also pose particular challenges to democratic governance. A group with values that differ from those of society at large may fail to develop trust in legal, educational, and civic institutions. In general, a government seeking to promote the general welfare of its citizens may face obstacles if people differ dramatically in their conception of what the good life is.

In this paper I present a model of groups as collections of individuals with shared values. I use the term *values* to represent a shared belief that a certain way of life, or mode of behavior, merits value.[1] Thus, while members of a group may behave differently from each other, they evaluate the world similarly, and their evaluation of themselves and of others may be different than others'

[1] Bruner, 1990, p. 22.

evaluation of them. From this standpoint, a group's values serve as a heuristic, or what Jerome Bruner terms a "communal tool kit," that is used by group members to both inform their own patterns of behavior, and to make sense of the potentially strange or unfamiliar behavior of others. In particular, I will argue that the values group members share constitute a shared internal narrative or dialogue. By jointly assigning more or less value to different modes of behavior, group members construct a story that explains the relative worth of the different ways that people choose to live their lives.

The model is used to examine why some group-specific values persist unchanged in diverse societies, while other values may face pressure to converge to, or diverge from, the values of other groups. By offering predictions about potential changes to a group's core values, the model provides some insight into the likelihood of intra-group tension, and potentially, conflict. The model can also be used to compare competing group values both from the standpoint of how a group evaluates itself relative to others, and how others evaluate the group with respect to their own heuristics. In this regard, the model can also offer insight into the likelihood of *inter*-group conflict.

Last, the model offers thoughts about a formalization of the concept of value pluralism, or the notion that there may exist multiple true ways of assigning value to different modes of life. While different ways of assigning value may themselves be incommensurable in terms of their correctness, the model shows that different value systems can nonetheless be analyzed positively. From this standpoint, theory speaks to questions of how political and social leaders might, constitutionally or otherwise, seek to embed certain values into a nation's core institutions. With respect to this last point, the model suggests that there are certain special and unique consequences of institutionalized norms of value pluralism — consequences that affect both how citizens of a plural nation will view themselves, and how others will view them. While the results in this paper are framed in terms of a situation in which multiple groups constitute a single society, they can also be applied to the more stylized concept of a single homogenous nation co-existing with other nations.

My ultimate aim with this project is to present a model of politics consistent with the observation that individuals often evaluate policies in terms of their perceived support for, or opposition to, different ways of life. Moreover, political decisions can themselves alter the values that people hold dear. In describing the relationship between individual values and the political self, Aaron Wildavsky writes, "Preferences in regard to political objects are not external to political life; on the contrary, they constitute the very internal essence, the quintessence of politics: the construction and reconstruction of our lives together."[2] Viewed in this way, my aim is to take one step toward formalizing the process whereby values shape preferences, and political institutions may potentially shape values.

2 A Formal Theory of Groups, Activities and Values

I begin by assuming that there are two groups, $g \in \{1, 2\}$, that partition the set of all people, and a collection of *activities*, or things that people can do, denoted $B = \{b_1, ..., b_K\}$. A particular activity, $b \in B$, is a type of observable action that an individual might engage in that could,

[2] Wildavsky, 1987, p.5

potentially, be deemed valuable. Examples of such activities could include spending time with one's family, going to the library, going to the gym, going to church, relaxing, giving money to charity, telling jokes, or different investment and consumption behaviors. The activities are assumed to be non-exclusive, so that an individual member of a group could potentially engage in many (or all) such activities to a greater or lesser extent. The activities are also assumed to be independent, so that engaging in one activity has no direct effect on whether an individual engages in another. Thus, knowing a person chooses to spend a subset of his time with family does not necessarily yield information about whether that person attends church. It will, however, imply that he cannot spend *all* of his time at church.

Individuals are assumed to have a finite reserve of discretionary resources (time, money, physical and mental energy) that they can elect to expend in engaging in any activity. Using Sen's language, I will refer to these discretionary resources as an individual's *capability*, which Sen defines as a measure of what an individual is actually able to do.[3] I will denote the *capability* of a representative member of group g as w^g, with $w^g > 0$. I will also note here that, as in Sen's capability approach, an individual's capability in this model characterizes a set of feasible bundles of activities that the individual could engage in. Unlike Sen, however, the activities I am concerned with are not necessarily similarly or unambiguously valuable. Sen focuses primarily on activities that he terms *functionings*, which are, he argues, activities that fulfill basic human needs, such as good health, nourishment, education, and religious freedom. On the other hand, I am concerned with activities that may or may not directly fulfill such needs, but which people might engage in

[3]Sen 1980, 1982, 1985, 1992, 1999.

nonetheless. Such activities could include the acts of becoming educated or nourished, but might also include acts of self-beautification, pride, risk-taking or modesty. In considering how bundles of activities are valued by different groups, I am interested in the *mixture* of activities individuals engage in. In this regard, even if an individual only engages in activities that could be considered to be "functionings" in Sen's sense, the particular mixture of functionings he engages in is itself something that can be valued and judged.

Underlying the model is an assumption that members of a given group have some shared cultural or historical experience that has led those individuals to attach similar value to a similar collection of activities. The value assigned by a group member to a particular activity is conceptualized as the perceived worth of that activity relative to other things that the individual might do. Thus, the value vector that a particular group g assigns to the various activities is denoted $v^g = (v_1^g, ..., v_K^g)$, with $\sum_{k=1}^{K} v_k^g = 1$, and $v_k^g \geq 0$ for all k. The term v_k^g denotes the relative value assigned to activity b_k by a member of group g. The vector v^g constitutes the *values* of group g.

I do not explicitly model the process whereby a group's values, v^g, are realized. I simply argue that many reasonable scenarios exist that could explain why members of a particular group share similar values. For example, members of a group, or their ancestors, may have faced a similar utility maximization problem, in that these individuals faced similar rewards for engaging in certain activities due to environmental factors. A particularly stylized model could assign Cobb-Douglas utility functions to these individuals, $U = \prod_{k=1}^{K} b_k^{\alpha_k^g}$, where $\sum_{k=1}^{K} \alpha_k^g = 1$, and α_k^g denotes the elasticity of utility for a member of group g with respect to activity b_k. Thus, α_k^g captures the responsiveness of a member of g's utility with respect to changes in that particular behavior, and a utility-maximizing allocation of activities would assign $v^g = (\alpha_1^g, ..., \alpha_K^g)$. We might consider the α^g terms to have been determined by conditions dominating the daily lives of group members, such as social status, the physical environment, or particular conditions of scarcity or abundance.

A different conceptualization of the process whereby a group's values are realized could involve a group's mixture of activities representing a symmetric equilibrium to a game played by the group members. Alternatively, a group's values could be shaped by a constitutional designer or social planner, one who institutionalizes certain behaviors through law in the hope of changing the attitudes and preferences of the governed. The simple point is that members of a group are similar in one respect: they share the same values.

2.1 A spatial model of how values shape perceptions

In this model, a group's values represent a shared belief that a certain way of life, or collection of activities, merits worth. These values serve as a group heuristic for evaluating all individuals on the basis of how they choose to live their lives. In this sense, values are tools that individuals can use to make order out of the highly multidimensional actions of others, by deeming certain behavior combinations to be more or less meaningful than others.

First, note that individuals in group g have a vector of values $v^g = (v_1^g, ..., v_K^g)$, that characterizes a ray in the positive orthant of $K-$dimensional space, originating at the origin. This ray represents the group's *value heuristic*. Second, the representative member of group g has capability w^g, which he expends engaging in the distribution of activities prescribed by his group. Thus, we can consider the representative member of g to be spatially located at $x^g = w^g v^g$. The point x^g

serves as a baseline for a member of g's evaluation of other individuals, and will be referred to as the representative member's *behavior*.[4]

Suppose that a member of group g seeks to evaluate the representative member of group h who is spatially located at (or engages in behavior) $x^h = w^h v^h$. This evaluation takes the form:

$$S(x^g, x^h) = ||x^g|| - ||\frac{x^h \cdot x^g}{x^g \cdot x^g} x^g||.$$

Note that $||v^i||$ denotes the Euclidean norm of vector v^i, and $v^i \cdot v^j$ is the dot product of vectors v^i and v^j. $S(x^g, x^h)$ then represents the difference between the norm of x^g and the norm of the orthogonal projection of x^h onto group g's value heuristic. In this sense, behaviors off the ray defined by v^g are evaluated by members of g on the basis of the point on the ray that most closely fits the deviating behavior. Thus, people observe the behavior of others, and form a possibly biased estimate of the capability of those individuals by considering the point consistent with their *own* values that most closely matches the behavior of the person being evaluated (the projection of x^h onto v^g). The evaluation then consists of comparing their own capability level (measured by $||x^g||$) with the perceived capability of the individual being evaluated (measured by the norm of the projection of x^h onto v^g).[5]

The process of evaluation is akin to a person in group g observing h's behavior and thinking

[4]Not to be confused with specific activities $b \in B$; a member's behavior x^g characterizes the mixture of activities he engages in.

[5]The term $||x^g||$ does not in general equal g's capability level w^g. However, the model simply requires that the norm of any point projected onto on the ray defined by v^g is weakly increasing in the capability term of the behavior characterized by that point.

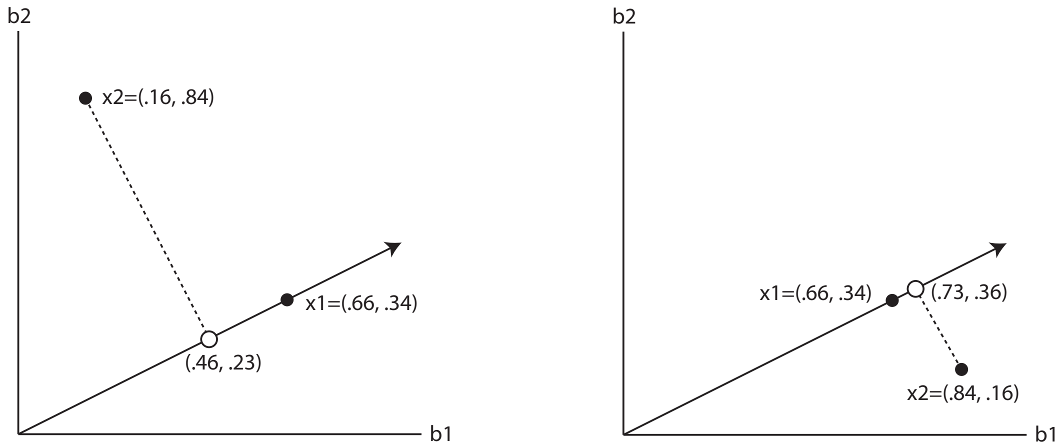


Figure 1: Group 1's evaluation of two different behaviors

"I don't entirely understand your motivations, but you are similar to someone I *do* understand, and that person has capability w, which is unambiguously valuable." The value heuristic therefore provides group members with a method of assigning a transitive ordering to the behavior of all individuals in society, with points closer to the origin being evaluated as less valuable than points farther from it, as they correspond to lower capability levels.

An example is shown in Figure 1, in which the two groups have the same level of capability, and there are two possible activities. In both pictures, Group 1 has a value heuristic of $v_1 = (.66, .34)$. Thus, the members of that group share an optimal mixture of activities that assigns twice as much value to b_1 than it does to b_2. To the members of Group 1, this weighting of activities represents the most reasonable and correct way of life.

When a member of Group 1 evaluates the potentially unfamiliar behavior of a member of Group 2, he does so by projecting 2's location, x^2, onto his own value heuristic. This orthogonal projection generates a unique point on the ray defined by v^1 that is closest to x^2; in other words, 1 evaluates 2 by considering the point consistent with 1's own value system that most closely matches 2's behavior. In Figure 1, $v^1 = x^1 = (.66, .34)$ and $v^2 = x^2 = (.16, .84)$. A member of Group 1 evaluates a member of Group 2 by projecting x^2 onto v^1, which yields the point $(.46, .23)$. Thus, Group 1 views the behavior of Group 2 as inferior to its own, or $S(x^1, x^2) > 0$.

Conversely, Group 1's evaluation of Group 2 when $x^2 = (.84, .16)$ is $(.73, .36)$, a point farther from the origin than 1's own position of $(.66, .34)$. Thus, $S(x^1, x^2) < 0$. In this case, Group 1 holds Group 2 in high esteem, for Group 2 appears to Group 1 as being capable of engaging in quite a valuable mixture of activities. In particular, the representative member of Group 2 engages in a mixture of activities that is most similar to a member of Group 1 whose capability is higher than the representative member of that group.

2.2 Value stability

The simple example described in Figure 1 demonstrated that it is not always the case that groups deem the values of others to be inferior to their own; there can exist situations in which one group's evaluation of another is overwhelmingly positive, to the point of that group evaluating the other's way of life as superior to its own. I describe this scenario as one in which the group may face pressure to change its own values and behavior.

If group g evaluates its own behavior x^g as superior to the other group's behavior, then x^g is termed *stable*. I call a collection of behaviors, $\{x^1, x^2\}$, stable if it is the case that every group's evaluation of its own way of life is deemed superior to its evaluation of every other group's way of life. Conversely, if a group evaluates another group's behavior as superior to its own, then that group may face pressure to change its own behavior. In this case, the group's behavior is termed *unstable*, and is *dominated* by another group's behavior. Formally, the conditions are defined below.

Definition 1 *A group g's behavior is* stable *if it is the case that*

$$S(x^g, x^h) \geq 0.$$

A set of behaviors $\{x^1, x^2\}$ is stable if both x^1 and x^2 are stable. If $S(x^g, x^h) < 0$, then x^g is unstable*, and in particular, is* dominated *by x^h.*

In words, the above conditions simply say that a group's behavior, x^g, is stable if the norm of x^g (or the distance from x^g to the origin) is greater than or equal to the norm of the orthogonal projection of x^h onto v^g, for the other group h. The stability condition given in Definition 1 can be redefined in several intuitive ways, as the following observations demonstrate.

Observation 1 *A group g's behavior is stable iff it is the case that $v^g \cdot x^g \geq v^g \cdot x^h$.*

Observation 1 says that Group g's process of evaluating any group's behavior is similar to g taking a weighted average of the activities undertaken by the group it is evaluating, where the weights are specified by g's value heuristic v^g.[6] Thus, at any given heuristic v^g, g's evaluation of itself is proportional to $v^g \cdot x^g$ and its evaluation of h is proportional to $v^g \cdot x^h$. If g's evaluation of its own behavior x^g is less than its evaluation of x^h, then x^g is unstable.

[6] I use the word "similar" because at any ray v^g, g's status evaluation is scaled by $||v^g||$, a positive constant.

For the same reason as Observation 1, the stability condition in Definition 1 can be defined in terms of the variances and covariances of the vectors v^g, as shown in the following observation.

Observation 2 *Group g's behavior is stable iff it is the case that*

$$w^g[\mathit{Var}(v^g) + \frac{1}{K^2}] \geq w^h[\mathit{Cov}(v^g, v^h) + \frac{1}{K^2}].$$

The above observation follows because both v^g and v^h have the same expected value, $\frac{1}{K}$, where K is the total number of activities possible. Thus, while it is clear that higher levels of capability w^g will always weakly improve others' perceptions of group g, Observation 2 also implies that there is some benefit accorded to groups that value a relatively small collection of activities. In particular, if all groups share the same level of capability, then the stability condition can be stated solely in terms of the variances and covariances of the v^i vectors: A collection of behaviors is stable if and only if $Var(v^g) \geq Cov(v^g, v^h)$ for all groups, g, h.

The next observation shows that, for any collection of behaviors considered, there will always exist a group whose behavior is stable. Later on, this will enable us to consider the behavior of a unique unstable group: when there are two groups, only one can face pressure to change its behavior.

Observation 3 *There exists a group m for which x^m is stable.*

Proof: Let $m \in \arg\max_h ||x^h||$. Then for $g \neq m$, $S(x^m, x^g) = ||x^m|| - \frac{x^m \cdot x^g}{x^m \cdot x^m}||x^m||$. This is weakly positive when $||x^m|| \geq ||x^g|| \cos\theta$, where θ is the angle between x^m and x^g. The angle between any two vectors in the positive orthant is between $0°$ and $90°$, and thus, the inequality holds, as $\cos\theta \in [0, 1]$. It follows that $S(x^m, x^g) \geq 0$ for $g \neq m$. □

Finally, before proceeding to the next section, it will be useful to consider the properties of one particular value heuristic that I will term the "plural point." This heuristic is located at the centroid of Δ^{K-1}, the $K-1$ simplex. It is $v^C = (\frac{1}{K}, \frac{1}{K}, ..., \frac{1}{K})$. Observation 2 showed that if we care about the value stability of an x^g, then there is some value to v^g having a high variance, or assigning an unequal distribution of weights to the various behaviors. Clearly at v^C variance is minimized, and so any behavior characterized by $w^g v^C$ is particularly difficult to maintain as stable. The following paragraphs formalize this and several other properties of plural point v^C.

Observation 4 *A group with value heuristic v^C evaluates all behavior as equivalent to the capability term characterizing that behavior.*

Observation 4 says that any ranking of groups produced by v^C is neutral, in that it preserves the ranking that would be achieved by simply ranking groups according to true capability. Such values do not differentiate between different behaviors: in assessing the worth of any given behavior, all possible dimensions of action are considered, and are interchangeable. This is because $v^C \cdot x^g = w^g(v^C \cdot v^g) = w^g$.

It is also important to note that any group with value heuristic v^C must be in a position of relative strength: it must have (weakly) higher capability than every other group. Such a group will always evaluate itself as equal to any other group with the same capability level. At the same time, any group with the same capability level as a group located at v^C will evaluate the group located at v^C as weakly inferior. However, while the group at v^C will be viewed as inferior by every other group at the same capability level, it is also the unique location that maximizes the minimum evaluation that any other group could have of it. In this sense, a group at v^C will seem less capable than it actually is to groups that are similarly capable, but it will never appear *too* incapable.

3 Values as narratives

Throughout, I have assumed that a group member's evaluation of others occurs with respect to his own value heuristic, and this evaluation can be positive, negative or neutral. A person with values that only assign positive weight to a single activity, for example, will evaluate others only with respect to that particular activity. Conversely, someone who assigns positive value to a wide range of activities will evaluate others with respect to all of these activities, and will be capable of trading off the worth of one activity for another in his evaluation (captured most perfectly at the plural value heuristic, v^C). One of the motivating ideas of this project is the conception of a value heuristic as representing an internal narrative or dialogue; by assigning value to one or many activities, the person constructs a story that explains the relative worth of the different ways that people choose to live their lives. Two ideas inform this account: first, that values shape the internal reasons, or narratives, that people construct in order to explain the world around them, and second, that this construction renders a multidimensional world unidimensional. These ideas are based on work by Jerome Bruner, who argues that people use narratives in order to construct their own sense of reality, and that these narratives are culturally transmitted—for example, through institutions such as schools and through legal codes.[7]

Jerome Bruner's hypothesis is that while humans have an innate capacity for narrative organization and discourse, our culture equips us with specific and important tools for assigning meaning to our experiences. He argues that what makes a cultural community is "the existence of interpretive procedures for adjudicating the different construals of reality that are inevitable in any diverse society."[8] Taking the existence of conflict as a given in any human society, these narratives serve a peacekeeping role by providing an explanation for conflict. In other words, a shared understanding of the *reason* for conflict serves to mitigate the conflict itself. Therefore, shared narratives neutralize conflict, and a cultural breakdown, Bruner argues, can generally be traced to a problem with this narrative.

Such a problem may be due to a fundamental disagreement about what constitutes an ordinary life versus what constitutes the divergent, as with a cultural revolution. It may be due to the "overspecialization of narrative," when stories become so ideologically self-serving that individuals no longer trust a culturally accepted interpretation of events, as might exist, Bruner argues, under a totalitarian regime. Or it may be due to an "impoverishment of narrative resources," in which a troubled group has been so focused on survival that a "worst scenario" story has come to dominate daily life, and alternate interpretations of events are no longer possible.[9] I will return to these examples later.

In modeling group values as a tool for evaluating the behavior of others, I am drawing from this

[7]Bruner, "The Narrative Construction of Reality," in *Critical Inquiry*,1991, and *Acts of Meaning*, 1996.

[8]*Acts of Meaning*, p. 95.

[9]Ibid, p. 97.

approach by assuming that group members share an internal narrative that provides an explanation for the positive or negative evaluation of others, and that this narrative yields a linear ordering over individuals. A value narrative could be very simple, as in “the person I am evaluating attends church less regularly than me, attending church is all that matters, and therefore this person is living a way of life that is less valuable than my own.” Or the narrative could be more complex, as in “the person I am evaluating attends church less regularly than me, but he studies more than me; I think that both church and studying are valuable, and so I am evaluating him in a neutral light.”

If we buy the argument that societal conflict may be traced to a problem with narrative (for example, disagreement between groups or disagreement between citizens and state) then understanding these narratives – what they imply about both individuals’ evaluations of others and how individuals perceive *they* are evaluated by others, and how these narratives may change over time – is important. If groups fundamentally disagree about which behaviors are and are not valuable, it is likely that those groups will evaluate each other in a negative light, and this may generate conflict.

The concept of value stability described in Section 2.2 relies on groups having a weakly negative evaluation of each other; if one group evaluates another as leading a more successful lifestyle than its own, it faces pressure to change its own mixture of activities. There is nothing inherently negative about a group believing that its own values are best. The problem arises when a group can see no worth at all to another group’s way of life, or when a group knows that another group views its own way of life as close to worthless. In this model this would occur, for example, if $x^1 = (w^1, 0)$ and $x^2 = (0, w^2)$. In these instances, there may emerge a failure of trust either between group and group or between citizens and the state.

The value "narratives" of the different groups that comprise a nation affect the institutional legitimacy of that nation in part because they affect trust in institutions. For example, a parent would likely not want their child taught by a teacher whose way of viewing the world is dramatically different than their own, a teacher who perhaps views the parents' values and lifestyle as worthless. An individual may not trust a legal system if the individual's interpretation of events always differs from the culturally acceptable viewpoint. In the following section I consider how unstable values may change. These results represent my first attempt to model how institutional changes – changes that might for example alter the distribution of capabilities across groups – affect value stability.

4 Value change

The stability condition outlined in Section 2.2 stated that a collection of behaviors $\{x^1, x^2\}$ is *stable* if and only if it is the case that for all g, h:

$$S(x^g, x^h) = ||x^g|| - ||\frac{x^g \cdot x^h}{x^g \cdot x^g} x^g|| \geq 0.$$

If $||x^g|| < ||\frac{x^g \cdot x^h}{x^g \cdot x^g} x^g||$ for $h \neq g$, then x^g is *unstable*, and in particular, is *dominated* by x^h. Using this condition we can characterize not only those situations in which a group faces pressure to change its values, but also the particular activities a group will face pressure to change.

Rewriting the above stability condition explicitly in terms of capabilities w and values v we get that stability requires for g, h:

$$S(x^g, x^h) = w^g (\sum_{i=1}^{K} (v_i^g)^2)^{\frac{1}{2}} - w^h (\sum_{i=1}^{K} v_i^g v_i^h)(\sum_{i=1}^{K} (v_i^g)^2)^{-\frac{1}{2}} \geq 0. \quad (1)$$

If S is negative then group g faces pressure to alter its bundle of activities so as to improve its evaluation of itself relative to h. If S is positive, then the stability condition is met, and g faces no pressure to change in response to its evaluation of group h. For the remainder of the paper I will consider S to be an explicit function of w^g, w^h, v^g, and v^h, as in Equation 1, because these are my main variables of interest.

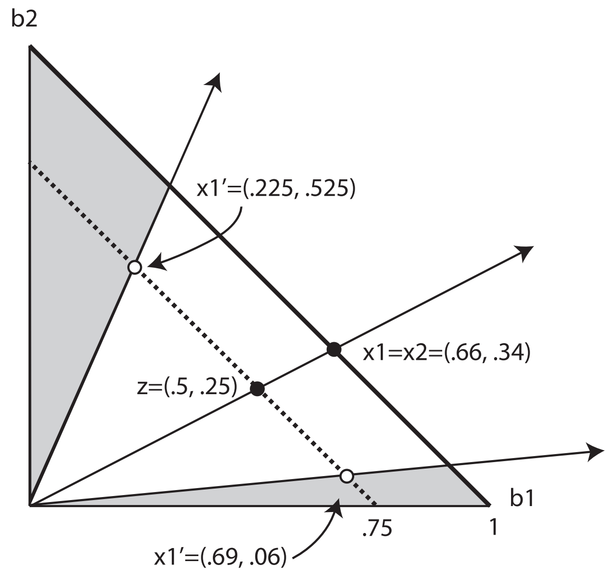
While the maxima of S with respect to v^g are corner solutions, they are easy to compute, and provide insight into how value stability is affected by the relative differences between group capabilities. Letting $\beta^K \subset \Delta^{K-1}$ denote the standard basis for $\mathbb{R}^K$, an examination of Equation 1 reveals that these maxima are:

$$\arg\max_{v^g \in \beta^K} S = \{v^g : v_i^g = 0 \text{ if } v_i^h \neq \min v^h\}. \quad (2)$$

Clearly there are numerous stable collections of behaviors. For example, any situation in which group g places all weight on a single activity ($v_i^g = 1$) and group h places all weight on a different activity ($v_j^h = 1$ for $j \neq i$) will generate a collection of behaviors that is stable for all capability levels $w^g, w^h > 0$. At the same time, as Observation 3 demonstrated, it is always the case that at most one group can face pressure to change its behavior. Therefore, it is useful to consider how such a group will attempt to alter its behavior in order to improve its status, holding the (stable) behavior of the other group constant. Before proceeding to more general results, the following example illustrates a simple dynamic that is mirrored in several of the propositions that follow.

Consider the situation depicted in Figure 2 in which Groups 1 and 2 are initially located at the same point $x^1 = x^2 = (\frac{2}{3}, \frac{1}{3})$, and thus, both groups have the same capability, $w^1 = w^2 = 1$.

Figure 2: Group 1 response to a negative capability shock



Next, suppose that Group 1's capability decreases to $w^{1'} = \frac{3}{4}$. If Group 1's values were to remain unchanged then Group 1 would be located at $z = \frac{3}{4}(\frac{2}{3}, \frac{1}{3}) = (\frac{1}{2}, \frac{1}{4})$. However, the set of behaviors $\{z, x^2\}$ is not stable; at point z Group 1 receives a strictly negative status term. In fact, when $w^{1'} = \frac{3}{4}$ and $x^2 = (\frac{2}{3}, \frac{1}{3})$ the only values for which Group 1 experiences weakly positive status relative to Group 2 are characterized by rays that lie in the shaded regions of the figure, or situations in which either $b_1 \geq .69$ or $b_1 \leq .225$. The points labeled $x^{1'}$ in the figure correspond to a Group 1 status of zero, equal to Group 1's status at its initial location x^1. Thus the figure demonstrates that a change in the relative capability levels of the two groups may at the very least lead a low capability group to alter its bundle of activities as a status-improving measure. At the most, these changes may cause the group to adopt behaviors at the periphery of its capability constraint (in this case, at the edges of the line connecting b_1 and b_2). Moreover, if we consider Group 1 and Group 2 to have initially been the *same* group, then this figure shows that altering the relative capability of a segment of a group may cause the group to split, in terms of its values. In this regard, the model could potentially be used to look at questions of endogenous group formation.

An examination of Equation 2 leads directly to the following proposition, which captures this insight. Throughout, I will use the term r to denote the ratio of group h's capability relative to group g's capability. Thus, $r = \frac{w^h}{w^g}$, and $w^h = r \cdot w^g$.

Proposition 1 *As group h's capability increases relative to group g's, value stability will ultimately require that v^g place all weight on a single activity.*

Proof: Equation 2 shows that g's status is maximized at $v^{g*} \in \underset{v^g \in \beta^K}{\arg\max}\, S = \{v^g : v_i^g = 0 \text{ if } v_i^h \neq \min v^h\}$. Thus, holding x^h fixed, the maximum status g can attain is $w^g(1) - rw^g(\min v^h)$. It follows that for $r = \frac{1}{\min v^h}$, the unique stable behavior vector for g is located at $x^g = w^g v^{g*}$. For $r > \frac{1}{\min v^h}$ there is no stable x^g. $\square$

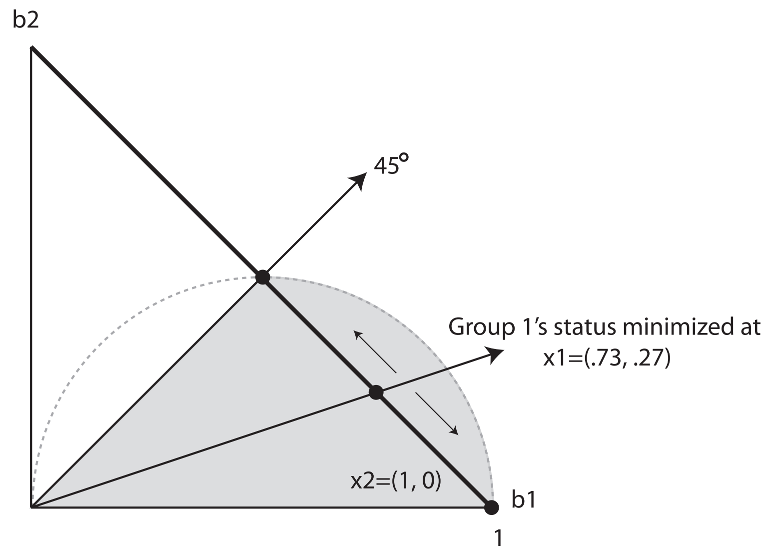
Proposition 1 shows that for a high enough difference in capabilities between groups the only stable values are those in which a low capability group places all weight on a single behavior, and moreover, this behavior must be the activity *least* valued by the high capability group. At the same time, this result says nothing about the initial location of the low capability group. It may be reasonable to assume that group values change slowly, so as to produce small status improvements. If this is the case, the direction of change will depend on the initial location of the low capability group, which will be considered in the next section.

4.1 Directional changes in values

The previous section showed that high capability differences between groups may lead a low capability group to adopt more "extreme" behaviors, in the sense of placing all value on a single activity. However, knowing that S is maximized at a corner does not provide any information about the likely direction of change for a low capability group; it simply says that in the most extreme setting of capability differentiation in which value stability actually exists, then a low capability group must be adopting status-maximizing values. In this section I argue that the dynamics governing how group values change are fundamentally different than those governing how, say, individual behavior adapts to changes in environment. Groups are not sentient beings, and likely do not decide to suddenly adopt a radically different set of norms from those that they had heretofore held. Values evolve over time, and in this section I consider instability of a group's values as instigating a process of small behavioral changes for that group. I simply argue that status-improving changes are more likely than status-decreasing changes.

To begin, consider Figure 3 in which two groups have the same capability, $w^1 = w^2 = 1$ and in which Group 2's behavior is located at the point $x^2 = (1, 0)$. Group 1 could potentially locate anywhere on its capability constraint, which is the line connecting $(1, 0)$ and $(0, 1)$. The dashed half circle represents the projection of x^2 onto any value heuristic (ray) that Group 1 could choose. Thus, when the intersection of Group 1's values and this curve lies to the right of 1's capability constraint then Group 1 receives negative status; when it lies to the left of the constraint then 1's values are stable. The shaded region corresponds to Group 1 values that are unstable.

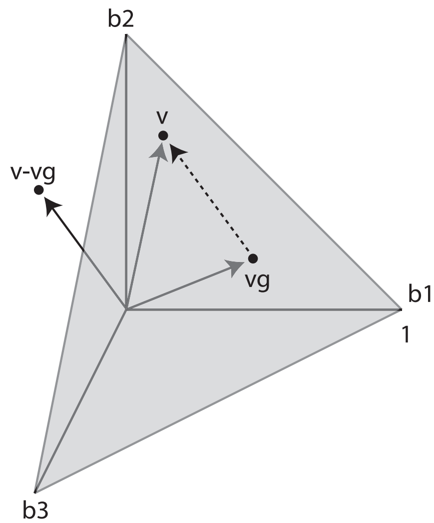
Figure 3: An example of status-minimizing behavior



In this example, 1's status is minimized at the point $(.73, .27)$. Moving away from this point in either direction is status-improving for Group 1. If Group 1 is initially located at a point at which $v_1^1 \in (.73, 1)$, a small movement toward $(1, 0)$ is status improving, and a small movement toward the plural point $v^C = (.5, .5)$ is status lowering. The opposite holds if Group 1 is initially located at a point at which $v_1^1 \in (.5, .73)$. Thus, even though Group 1's status is maximized at $x^1 = (0, 1)$, Group 1 may never face pressure to increase the weight it places on activity b_2, because doing so would result in a small decrease in status. In this example Group 1 faces pressure to engage in behavior that is the reverse of what would lead to his status-maximizing point.

In this section I will assume that value instability results in movement toward locally, as opposed to globally, status-improving behaviors. In particular, I am interested in calculating g's change in status as v^g moves along g's capability constraint (a K-1 dimensional hyperplane) to-

Figure 4: Value changes along the capability frontier



ward some new $v \in \Delta^{K-1}$. This is pictured in Figure 4 for the 2-simplex. The rate of change of S through v^g in the direction of vector $v - v^g$ is equivalent to the projection of the gradient vector of S with respect to v^g onto the K dimensional unit vector that is codirectional with $(v - v^g)$. The gradient of S with respect to the elements of v^g is:

$$\nabla S = \{(w^g v_i^g - w^h v_i^h)||v^g||^{-1} + w^h v_i^g (v^g \cdot v^h)||v^g||^{-3}\}_{i=1}^K,$$

and the rate of change of S in the direction of vector $v \in \Delta^{K-1}$ is $C(v^g, v) = \nabla S \cdot \frac{v - v^g}{||v - v^g||}$, or:

$$C(v^g, v) = \sum_{i=1}^{K} \Big(\frac{w^g v_i^g - w^h v_i^h}{(\sum_j (v_j^g)^2)^{\frac{1}{2}}} + \frac{w^h v_i^g (\sum_j v_j^g v_j^h)}{(\sum_j (v_j^g)^2)^{\frac{3}{2}}}\Big)\Big(\frac{v_i - v_i^g}{(\sum_j (v_j - v_j^g)^2)^{\frac{1}{2}}}\Big). \quad (3)$$

We can now use this formula to characterize situations in which instability of v^g might, for example, lead g to engage in a mixture of activities more or less similar to those engaged in by group h.

In particular, a move toward v is locally status improving for group g whenever $C(v^g, v) > 0$, or (dividing out positive terms):

$$\sum_{i=1}^{K} \left(w^g v_i^g - w^h v_i^h + \frac{w^h v_i^g \sum_j v_j^g v_j^h}{\sum_j (v_j^g)^2}\right)(v_i - v_i^g) > 0. \tag{4}$$

The first result I consider concerns the plural point, v^C. Observation 4 showed that only a high capability group can be stably located at v^C. However, the discussion in Section 3 implied that, from the standpoint of reducing intergroup conflict and increasing intergroup trust, there may be some benefit to groups having values near v^C, because such groups are neutral in their evaluation of others. This result then concerns whether it can ever be possible for a low capability group to find it status-improving to move toward v^C if a high capability group is located there. The result is negative.

Proposition 2 *It is never strictly status improving for g to move toward v^h when $v^h = (\frac{1}{K}, \frac{1}{K}, ..., \frac{1}{K})$.*

Proof: Letting $v_i = v_i^h$ in Equation 4 and collecting terms, we can show that it is locally status improving for g to move toward v^h only when:

$$\sum_{i=1}^{K} v_i^g v_i^h - (v_i^g)^2 - r(v_i^h)^2 + r\frac{(\sum_j v_j^g v_j^h)^2}{\sum_j (v_j^g)^2} > 0.$$

If $v^h = (\frac{1}{K}, \frac{1}{K}, ..., \frac{1}{K})$, this reduces further to a condition that:

$$\frac{1}{K} - \sum_j (v_j^g)^2 - \frac{r}{K} + \frac{r}{K^2 \sum_j (v_j^g)^2} > 0.$$

However, satisfaction of this condition would require that $\sum_j (v_j^g)^2 < \frac{1}{K}$, which cannot occur as $v^g \in \Delta^K$. $\square$

Proposition 2 showed that while behaviors represented on the ray passing through the centroid possess several desirable qualities from the standpoint of how such a group evaluates others relative to itself, these behaviors may not be attractive to low capability groups. The next result concerns a situation in which two groups initially have the same values, $v^g = v^h$. If $w^g < w^h$, then v^g is unstable. The following proposition characterizes the kinds of behaviors g will face pressure to change in order to improve its status.

Proposition 3 *Suppose $v^g = v^h$ and $w^g < w^h$. Then g's status is improving in the direction of v if and only if $\sum_{i=1}^{K} v_i v_i^h - (v_i^h)^2 > 0$.*

Proof: Recall that $w^h = r \cdot w^g$, and let $v^g = v^h$. A move toward v is status improving whenever $C(v^g, v) > 0$. Equation 3 implies that:

$$C(v^h, v) = \sum_{i=1}^{K} \Big(\frac{w^g v_i^h - r w^g v_i^h}{(\sum_j (v_j^g)^2)^{\frac{1}{2}}} + \frac{r w^g v_i^h (\sum_j (v_j^h)^2)}{(\sum_j (v_j^h)^2)^{\frac{3}{2}}}\Big)\Big(\frac{v_i - v_i^h}{(\sum_j (v_j - v_j^h)^2)^{\frac{1}{2}}}\Big),$$

which is positive when

$$\sum_{i=1}^{K} (v_i^h)(v_i - v_i^h) > 0. \square$$

Proposition 3 shows that when $v^g = v^h$, g will find it status improving to increase the weight it places on those activities that are highly valued by both the high capability group, h, and itself. Thus, if we consider the case where a portion of a group suddenly becomes relatively disadvantaged, the disadvantaged group members will find it locally status-improving to move in the direction of placing all weight on the activity that the group initially valued most highly (in other words, they will find it status-improving to “beat their former group members at their own game.”)

This result speaks to questions of endogenous group formation, and is invariant to r, the size of the capability shock experienced by g. A corollary of this result is that the disadvantaged segment will not attempt to differentiate itself from its former group members by moving in the direction of placing all weight on the behavior *least* valued by those individuals:

Corollary 1 *When $v^g = v^h$, a move toward g's status-maximizing point is always weakly locally status-decreasing.*

Proof: Follows directly from Equation 2 and the logic above. □

The final result assumes that $v^g \neq v^h$, and in a sense shows that Proposition 3 is knife-edged. In this case, when r, the capability difference between g and h, is sufficiently high g can always improve its status by moving in the direction of placing all weight on the behavior that maximizes the difference between g's *perception* of h's behavior and h's actual behavior. This condition is not particularly intuitive, and so Figure 5 is provided to provide some insight into the dynamic.

Proposition 4 *For high values of r, when $v^g \neq v^h$ it is always weakly locally status-improving for group g to move in the direction of placing all value on the activity that maximizes $v_i^g \frac{v^g \cdot v^h}{v^g \cdot v^g} - v_i^h$.*

Proof: Equation 4 can be decomposed into two terms:

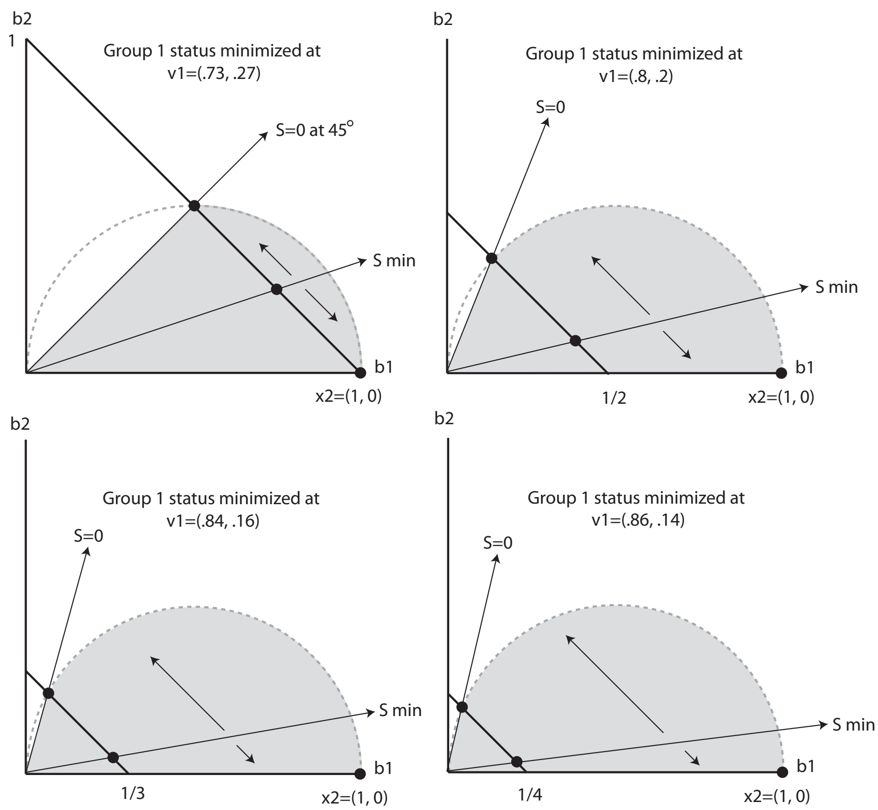
$$\sum_{i=1}^{K} v_i \big(v_i^g - r v_i^h + \frac{r v_i^g \sum_j v_j^g v_j^h}{\sum_j (v_j^g)^2}\big) - \sum_{i=1}^{K} v_i^g \big(v_i^g - r v_i^h + \frac{r v_i^g \sum_j v_j^g v_j^h}{\sum_j (v_j^g)^2}\big).$$

Each term is a convex combination of the same quantity, and because v^g is fixed, Equation 4 is weakly positive whenever v_i places all weight on the maximum of the quantity, which is the maximum of $(v_i^g - rv_i^h + \frac{rv_i^g \sum_j v_j^g v_j^h}{\sum_j (v_j^g)^2})$. For sufficiently large r, this term is maximized by setting $v_i = 1$ for some $i \in \underset{i\in\{1,\ldots,K\}}{\arg\max}\, v_i^g \frac{v^g \cdot v^h}{v^g \cdot v^g} - v_i^h$. □

Corollary 2 *For sufficiently high values of r, it is always weakly locally status-decreasing for group g to move toward placing all value on the behavior dimension along which the difference between g's perception of h's behavior and h's actual behavior is minimized.*

Proof: Immediate from above. □

Figure 5: Changes in relative capability



To see what these statements mean, consider Figure 5. In this figure, the high capability group,

Group 2, is located at $x^2 = (1, 0)$. The four figures depict Group 1's optimal direction of change depending on the initial location of x^1 and the capability differential r, where $w^2 = 1$ and $w^g = \frac{1}{r}$. Clockwise from the upper left, the images show $r = 1, 2, 3$ and 4, with the capability constraint of Group 1 pictured. The status-minimizing ray for Group 1 is pictured, as is the ray that grants Group 1 a status of zero. Group 1's perception of Group 2, or the projection of Group 2 onto every possible v^1, is pictured by the dashed half circle. As in Figure 3, the shaded regions represent the rays along which g receives negative status.

The dimension that maximizes Group 1's perception of Group 2's behavior minus Group 2's actual behavior is b_2: Group 2's actual behavior on b_2 is zero, while Group 1 evaluates him almost everywhere as strictly positive. Similarly, b_1 minimizes this difference, as Group 2's actual behavior on b_1 is one, while Group 1 evaluates him almost everywhere as strictly less than one. For each possible initial location of v^1, the arrows show the locally optimal direction of change for Group 1. Thus, as r increases Group 1 will benefit from moving in the direction of b_2 for an increasingly large set of initial values of v^1.

5 Concluding thoughts

This paper is concerned with two, difficult to attain, goals of any heterogeneous nation. One goal is to promote trust between groups that may have very different ways of viewing the world. The other goal is to induce widespread trust in state institutions by citizens with different values. With respect to this second goal, I consider a scenario in which political leaders may choose to embed values into a nation's institutions – in a sense, to choose a "national" value heuristic. If a group's trust in institutions are partly a function of how the group is evaluated by those institutions, then the model presented here suggests that there are certain special and unique consequences of institutionalized norms of value pluralism, by which I mean institutionalizing values that lie at the centroid $v^C = (\frac{1}{K}, \frac{1}{K}, ..., \frac{1}{K})$. Such values do not differentiate between different behaviors, and from Bruner's standpoint, yield the most "complex" narrative: in assessing the worth of any given behavior, all possible dimensions of action are considered and are interchangeable. Such values are uniquely neutral, in that they evaluate every individual's capability correctly. In practical terms, institutionalizing these values could be accomplished by, for example, prohibiting public institutions from favoring one lifestyle over another.

At the same time, and with respect to the first goal mentioned above, plural values can be stable only for a group that is in a position of relative strength. The model suggests that low capability groups seeking to improve their self-evaluation may frequently find it desirable to adopt more extreme values, in the sense of reducing the collection of behaviors they place value on. Thus, reducing the capability of a group that is a value outlier will produce increased incentives for that group to become more extreme, rather than more mainstream (Proposition 2). The results of this paper suggest that the best way of promoting value pluralism, and consequently improving groups' evaluations of each other, is by equalizing capabilities across groups. The only instances in which it can be stable for all groups to adopt values close to the plural point is when all groups have close to the same capabilities.

This model makes many strong assumptions, one of which is that a group's behavior is located on its value heuristic (i.e. the representative member of a group chooses a mixture of activities equal to the relative weights he places on each activity). At the same time, one could also consider a value heuristic as distinct from a group's behavior. In this case, a group member simply uses the heuristic to evaluate both himself and others. For example, if the representative member of a group places equal weight on activities b_1 and b_2 it may not be natural to say that the member actually engages in those behaviors in equal amounts. In future work I plan to consider a related model in which groups may locate at points off of their own value heuristic.