

**Downs and Rocke, “Tacit Bargaining, Arms Races, and Arms Control”**  
**Tacit Bargaining, Arms Races, and Arms Control. Ann Arbor: University of Michigan. Pages 92-100.**

**Review of Assumption 3:**

**It is possible to regard the decision-making apparatus of a state as a unitary actor?**

Several pieces in this week’s reading have examined the question of whether or not one can assume that the states act with one mind, in one manner. Graham Allison’s analog is that of a single chess player moving pieces with one mind and one hand firmly and predictably attached to it. As an assumption held constant to describe other phenomenon at play in international decision-making, such a model is useful. As a descriptor of how things really work, it is vastly more limited, but not entirely irrelevant. This article outlines some justifications for the cases of traditional unitary actor assumptions, then expands the concept of the unitary actor to include a broader range of actors influencing a Decision Maker, who collectively form a “hypothetical aggregate actor” behaving as a unitary actor.

My reading of this weeks readings suggests that this is a platform of the (still keeping all these categories straight...slippery little suckers that seem to wiggle over time) rationalist realist school.

Summary of the article

The article begins by reviewing (briefly) the members of the realist school – citing Thucydides, Waltz, and de Mesquita – who have assumed state nations behave with self-interest in a manner similar to individuals. It then recognizes the contributions of people like Allison, Cohen and March, Cyert and March, and Halperin and Kanter, who argue that bureaucratic politics play a role in foreign policy-making...a nation is not, in fact an individual, and that governments comprised of individuals behave differently than a single actor would. They use utility function language to describe the phenomenon, but I’m not sure exactly what they mean, since they don’t describe the function very fully.

They then move to a logic of empirics to test validity of different models, but quickly say such an idea is unworkable due to insufficient case studies for testing purposes. I find that odd, given the vast number of decisions made in international relations, but maybe they can’t find the data in the form they need. In any case, they conclude by asking two interesting questions: when can the fp-making apparatus of a country be aptly represented by a solitary actor model, and how often does that occur.

Answers:

1. Strong executive leadership in FP with institutional power backing it up. See de Mesquita’s *The War Trap* for one example of a fleshed out argument. Essentially it says that when there is deadlock, a strong voice can be tie-breaking and shift the

policy to or from war. Moving from modeling to case study, de Mesquita shows that erstwhile democracies tolerate authoritarian tendencies when war needs so require, and gives several examples from the US and Britain. Here are arguments for individual leadership in times of crisis. One is reminded of L'esprit des lois when Montesquieu writes something like "in times of great change men shape institutions, all other times institutions shape men." So in sum, the decision to go to war (which is only one of many FP decisions, we may note, but the one focused on here) tends to give more authority to the individual executive (as opposed to the executive branch writ large) for major decision-making. More mundane stuff sticks with the bureaucratic institutions, which have already had time to become established and to develop "standard operating procedures"...and with bureaucratic politics, again, which have had time to evolve and to develop norms, alliances, interests and constituencies.

1a. Executive Branch dominance is still a big part of the story, though, for implementation of the chief executive's decision, and also helpful because it diverts attention from personality of the chief executive (even if he/she is making the decisions alone) and redirects attention to structural strengths.

Note: Arrow's Paradox, or Arrow's Impossibility Theorem is cited prominently here. I googled it, and for those of you, like me, who don't know what it is, the inestimable source Encyclopedia4U defines the following:

In [voting systems](#), **Arrow's impossibility theorem**, or **Arrow's paradox** demonstrates the impossibility of designing rules for social decision making that obey a number of 'reasonable' criteria. The theorem's content, somewhat simplified, is as follows. A society needs to agree on a [preference](#) order among several different options. Each individual in the society has his or her own personal preference order. The problem is to find a general mechanism, called a *social choice function*, which transforms the set of preference orders, one for each individual, into a global societal preference order. This social choice function should have several desirable properties: **unrestricted domain** or **universality**: the social choice function should create a complete societal preference order from every possible set of individual preference orders. (The vote must have a result.) **non-imposition** or **citizen sovereignty**: every possible societal preference order should be achievable by some set of individual preference orders. (Every result must be achievable somehow.) **non-dictatorship**: the social choice function should not simply follow the preference order of a single individual while ignoring all others. **positive association of social and individual values** or **monotonicity**: if an individual modifies his or her preference order by promoting a certain option, then the societal preference order should change only by (possibly) promoting that same option. (An individual should not be able to hurt a candidate by ranking it *higher*.) **independence of irrelevant alternatives**: if we restrict attention to a subset of options, and apply the social choice function only to those, then the result should be compatible with the outcome for the whole set of options. (Removing some candidates should not have an effect on the relative ranking of the remaining candidates.) Arrow's theorem says that such a social choice function does not exist if the number of options is at least 3 and the society has at least 2 members.

Another version of Arrow's theorem can be obtained by replacing the monotonicity criterion with that of: **unanimity** or **Pareto efficiency**: if every individual prefers a certain option to another, then so must the resulting societal preference order. This statement is stronger, because assuming both monotonicity and independence of irrelevant alternatives implies Pareto efficiency. With a narrower definition of "irrelevant alternatives" which excludes those candidates in the [Smith set](#), some Condorcet methods meet all the criteria.

In case you got bewildered, none of this logic, according to the article, is relevant for our purposes. The massive executive branch foreign policy resources serve, as it were, at the pleasure and for the pleasure of the President, as a sort of well-tuned engine to do his will (or should I write His will?).

2. Fundamental attribution bias. The logic goes as follows: nation A ignores its ordinary pluralist features and behaves as a unitary actor, thereby giving nation B incentive to be more unitary-actorlike in response. The concern is that if one side is being pluralist, it might jump around a bit in some sort of confusing manner, which means the other side is more likely to misinterpret actions, leading to increased danger. Seems a bit of a stretch in this day and age, but I don't doubt that more complexity in decision-making implies a requirement for more sophisticated analysis from afar, and therefore the chances of misinterpretation are raised by sheer law of probability. In its extreme, this bias doesn't make sense on a number of dimensions (leaders can shut down pluralism at will; analysts aren't clever enough to account for non-consistent policy shifts), but as a muted trend, it seems to have some explanatory power.
3. Fear of Exploitation. States don't want to get duped when it counts or...preemptive attribution. States with naturally sketchy relations are constantly pushing the boundaries to see what they can get from the other. If you cede ground gradually, you're likely to end up losing more substantially over time, giving states with hostile relations incentives to preempt confusion about intent, interpretation, and consequence. When Nation A sends a clear signal of intent...it will hold Nation B responsible for potential event X...the chances of event X happening decrease, as responsibility and reaction are established ex-ante. By definition, everyone knows how everything is going to be interpreted. Thus, states "force" unitary actordom on their rivals for pragmatism's sake...out of a fear of exploitation.
4. Formal Models. Back to Arrow (I knew he'd be useful) and the Condorcet paradox (<http://www-rohan.sdsu.edu/~jwingram/condorcet.html> models it, but in essence describes the insensitivity of social preference obtained by pairwise vote). We see that majority preference gives us unstable outcome predictions, or in other words, we can't easily predict in any individual case what the majority preferred action will be, holding constant for bias. This is problematic, but, Black has modeled a utility function that helps us think about things a bit differently. He notes that if the function is single-peaked (a big if, but sometimes holds true), the median voter wins. Others (Caplin and Nalebuff) show that mean voter preferences rule if decision criterion are 64 percent majority, which setting aside some hoary assumptions, basically wants us to believe that most foreign policy decisions have a relatively large degree of consensus. If I read this right, the models are useful in explaining partial factors under highly stylized circumstances, but not yet ready to integrate to answer our "unitary actor" question as a whole. And, they are also somewhat inconsistent...with some of the earlier ones demonstrating "the people are muddled" effects, and the later ones arguing that the preferences of median or mean voters can have predictive capacity.

## 5. Plausible Combinations of Pluralist Preferences (PCPP's)

This is an interesting argument. Mind meld of advisors/actors resulting in a “buck stops here” executive. The authors call this executive “a hypothetical aggregate actor whose preferences are a weighted average of those of all the actors”.

This is where the fun begins. They're assuming a quadratic function (I'm sorry not to know why this is the norm, but they assert it is so, and build equations accordingly) with complete freedom of choice for optimizing ideal points left to each actor relevant to the model. Everyone's objective is to get as far as possible towards his or her ideal outcome (which the authors designate  $v(1)$ ) without over expending their own individual political capital, which presumably is the binding constraint for each individual  $v(i)$ . There is a whole bunch of fancy stuff that basically boils down to horse-trading amongst actors to maximize their own  $v(i)$ 's while limiting their capital expended  $(1/2b(i)a(i)E(o)a(i))$ .

The Decision Maker (chief executive) maximizes his final utility when he/she is able to take his or her own “undisturbed” maximized utility, and discount it by the sum of the pressure that the other actors have applied.

To do that, the DM starts by solving the FOC (first order condition) that maximizes his own choice without outside influence...i.e. what he/she wants to do in a pressure vacuum.

The proof then moves on to model the utility for actor  $i$  at the DM-optimal choice. Using the same quadratic function, the model substitutes the first variable,  $x$ , in the general equation, with the DM optimal point  $x^*$ , then subtracts the  $v(i)$  point (this uses a Euclidean framework, or basic graph theory) in the standard quadratic formula (which I'm still taking on faith). And then it subtracts by the cost that the individual actor incurs for advocating a departure from the DM's ideal point. The “chit cost”.

By differentiating out the equation  $dx^*/da(i)$ , with  $a(i)$  representing the “action vector” chosen by actor  $i$ , you get a value for  $a(i)$  which you can then plug into the FOC for the DM, described above, which incorporates the discounted preferences of the individual actors into the DM's optimization equation.

All of this, then, gives us a simplified model of our “hypothetical aggregate” DM.

But what does it mean?

Since we're differentiating based on  $a(i)$ , by definition we're looking at manipulating the FOC depending on the direction and force (length) of each vector applied to the DM's original  $v^*$ , subject to two conditions: no final decision point can land outside the DM's utility parameters (defined as orthogonal to  $a(i)$ ), and the cost to  $i$  of influence is proportional to  $a(i)$  squared (which isn't explained very well – why not three times, or whatever, but as a place-holder estimate is probably as good as any other).

Or in other words, we're looking at a situation where the unitary actor utility is the sum of the DM's (as a baseline) with all the others factored as appropriate given their divergence from the DM, the force of the sentiment, and the amount of power expended to influence the DM. The example given in the text is something like the President, who takes his views, then has to incorporate the Congressional leadership and the cabinet members. Given the wide latitude that the President (DM in our model) has over many FP decisions, starting with this model makes some sense...unlike, say, in more domestic issues where power is more evenly distributed amongst many DM's.

So...these authors find more credibility in the idea that the unitary actor assumption, in spite of its heavy criticism, than one might assume. Executive instruments such as agenda-setting power (factor 1 in the outline, or in thePCPP model, starting from a DM  $i^*$ , and diverging thenceforth), the hierarchical nature of bureaucracies (condition 1a in our outline), ability to appoint key officials, and access to intelligence information not widely dispersed all combine, along with other factors not mentioned, to give the President more authority to be our DM in foreign than in domestic policy-setting. In addition, the accountability argument (a blend of conditions 2 and 3), along with our model of organizational decision-making (that nasty proof we just slogged through) give us what we need to reconsider the usefulness of the unitary actor.

#### My Comments:

Since I just wrote more pages (just about) than the article itself was, I'll keep it brief. Basically, I'm still a skeptic. The model is interesting, to be sure, but really it could pretty easily be flipped around to reduce the supremacy of the DM as the base starting point...it's all just a matter of how you state the FOC, so the proof is a bit circular. It asks us to assume the central role of the DM, then, recognizing that the DM isn't in a bubble, incorporates other views. But, it never shows us why the DM is the starting place in the first place, which is the point of the unitary actor theory, so it doesn't prove the basic point too convincingly, in my opinion. The other models are hard to evaluate, but mutually contradictory, so I'm not sure what to do with them either.

But my real question about this model is why anyone would think it qualified as unitary actor proof in the first place. I'm not trying to be snippy at all, but really am confused. The whole point of the model is that DM's are a dominant force, but still influenced by the position of others around them, which is pretty much the antithesis of the unitary actor model. I think what's been described is actually much more a variation on some model of bureaucratic politics in the White House, where everyone has an opinion, some strongly held, others less so, and there's a scramble for influencing the President, who comes to the table with his own ideas as a starting point. Isn't this the definition of bureaucratic politics? And can you really think that any sort of model that incorporates consensus, which this one does with the second half of the proof, is unitary over time, even if it might be in a single case (in which case this is a repeat of deMesquita's argument in the first point anyway). Using the language of the proof, for any one decision, you have a set range of variables making the ultimate DM outcome set, but for

round two, the variables will shift on all dimensions, I would argue, likely somewhat randomly. So while the  $i^*$  may exhibit consistency over time, as the DM has a predictable set of preferences and ideological beliefs, this model gives us dozens or hundreds of other variables that effect the outcome...ranging from the beliefs of many actors, to their willingness in any individual case to use capital to influence the DM. If an incentive of the unitary actor model is in some sense to enhance consistency/predictability in foreign policy making, as asserted by factors 2 and 3, this model shows precisely why many actors make that unlikely in multiple round modeling.

Well, an interesting talking point.

Working backwards, we have factors three and two, which are basically flip sides of the same coin...the first one says we require you to act unitary-like so we understand what you're doing, the second says we're going to start acting unitary-like and you'll follow so we don't miscommunicate. But this also misses the point, because they mix up input and output. The point of unitary actor is NOT that the nation behaves predictably, or rationally, or consistently so we can all communicate effectively; it is indifferent to effective policy. This model argues that one person (the chess player, president, DM) makes and executes decisions largely devoid of either bureaucratic processes or politics. The assumption is that the outputs will reflect the inputs, but the theory itself is defined, as I read it, by the inputs. These two factors I believe are best employed in informing trends and likelihoods for a tendency towards unitary actordom, but not in and of themselves proof, or descriptors. And until the link between effective communication and unitary actordom is proven, rather than asserted and assumed, we can't really know if these factors actually matter, or are perceived to matter by those in authority.

That leaves us with the first factor, which I find the most convincing. The one case they really don't go into is the most obvious: an all-powerful dictator. It falls under category 1, but isn't described as an example. Clearly, Stalin (or some other dictator with a strong executive) could pretty much act as a unitary actor a lot of the time. So the idea that there are times in countries' histories when the strong fist connects to a single brain is credible, and I think more demonstratable than the paper suggests. One could in fact argue that the unitary actor is a *characteristic* of strong authoritarian regimes, as well as of ordinarily pluralist regimes in time of great change, which the authors describe aptly.