Activism and Ideological Self-Sabotage

Will C Denton[a]

[a]*From Parts Unknown*

Abstract

A model is developed in which a problem is addressed by activists. There are two types of activist: *pragmatists* care only about reducing the problem; whereas *ideologues* also dislike the problem, but find value in having a problem to solve. The model predicts that ideologues allow a problem to persist longer than necessary because they are driven to counter-productive and even extreme behavior as the problem is mitigated.

Keywords: activism, progress, ideologues, discourse, radicalization

1. Motivation of Activists

The behavior of activists can be perplexing. When PETA engages in a divisive campaign that generates little more than blowback, it becomes evident that their stated intentions—improving the welfare of animals—and their actual intentions—whatever they may be—are not totally aligned. When rioters turn a peaceful protest into a violent one, galvanizing opponents and dividing their own base, the same misalignment becomes evident. When a prominent figure is called out for an infraction that they did not commit, future callouts are met with increased skepticism and more real infractions avoid consequences. At the most extreme, it is impossible to ascribe any sort of productive intent to a terrorist because their actions are reviled even by the overwhelming majority of those with ideological sympathy.

In other words, the behavior of some activists is very obviously counter-productive to their stated causes.

Some activists do not merely want the problem gone; they want the purpose that comes with addressing the problem. Being an activist is an identity. It gives status and signals virtue. But when purpose is found by fixing a problem, the fix negates that purpose and status and identity. The end result is a state of co-dependency in which some activists need the problem to persist so they can continue to fix it.

I call these types of activists *ideologues*. They are activists who have something to lose if the problem is eliminated. Prime candidates include academics and authors whose careers become obsolete if the problem is eliminated; and charities and non-profit organizations whose donations dry up if the problem is eliminated. To quote the (pragmatic) activist Booker T. Washington,

> "There is a certain class of . . . problem-solvers who don't want the patient to get well, because as long as the disease holds out they have not only an easy means of making a living, but also an easy medium through which to make themselves prominent before the public" [1].

Other prime candidates include those who base their identity, sense of self, or sense of purpose on having a machine to rage against. Many college students, lacking real world experience and struggling with self-actualization, gravitate to activism in an almost born-again way. Indeed, the zeal of

Email address: willcdenton@protonmail.com (Will C Denton)

student activism is notorious for its sanctimony and incivility [2]. For many students, this type of behavior is undoubtably just the arrogance and naiveté of youth (e.g. a 20-year-old giving an unsolicited lecture to a 40-year-old); but for others, the existential boon of being an Activist™ becomes central to their being.

In this short paper, I create a model of such conflict of interest. There will be some problem on which two types of activists focus, and there are two types of activists. *Pragmatists* are activists motivated only by the desire to solve the problem. *Ideologues* are activists motivated by the desire to solve the problem, but also find purpose by having the problem to solve. Totally fixing the problem totally destroys that purpose, so ideologues are conflicted. Ultimately, therefore, ideologues want the problem at a small but strictly positive level.

2. Results

When there are no ideologues and all activists are pragmatists, the problem strictly falls over time to zero. On the other hand, if ideologues are present and their counter-productive behavior is unrestrained, then the problem will converge to a non-zero level because ideologues need non-zero problem to maintain a sense of purpose. If ideologues can be removed from discourse for being too extreme, then myopic ideologues will be driven to extreme counter-productive behavior in a desperate attempt to maintain purpose, after which the problem rapidly falls to zero; whereas forward-looking ideologues will restrain their counter-productive behavior to a level that sustains the problem for longer, albeit finitely.

3. The Agents

Let x_t denote the level of the problem at time t. There is measure N of pragmatists. Each pragmatist can maximally fix proportion $\alpha > 0$ of the problem per period; and since fixing is their only motivation, they exert all of their energy toward fixing αx_t of the problem in each period t. Because there are N pragmatists, they jointly reduce the problem by $N\alpha x_t$ in each period t. It will be assumed that $\alpha N < 1$ so that the problem cannot be immediately fixed, in which case the whole enterprise becomes boring. (This also allows some negativity concerns to be avoided.)

If the problem is above the *purpose threshold*, τ, then ideologues decide that fixing the problem is top priority. If the problem is below threshold τ, then ideologues feel that their purpose is too diminished and therefore act in a seemingly counter-productive way that worsens the problem: ideological self-sabotage.

There is measure M of ideologues. In each period, ideologues jointly affect the problem by proportion $M\beta_t$, where $\beta_t \leq \alpha$ captures how strongly ideologues choose to act. When the problem is above τ, ideologues choose to fix the problem with positive effect $\beta_t > 0$. On the other hand when the problem is below τ, ideologues choose to exacerbate the problem with negative effect $\beta_t < 0$. Leaving β_t unbounded from below captures the possibility of extreme counter-productive or *radical* behavior.

4. The Dynamics

Pragmatists always behave in the same way, so ideologues can be thought of as being second-movers in each period. Suppose $x_t = \tau/(1-N\alpha)$. Then ideologues will do nothing because next period will have $x_{t+1} = \tau$ entirely due to the actions of the pragmatists. When $x_t > \tau/(1-N\alpha)$, then ideologues will choose $\beta_t > 0$ because pragmatists alone cannot reduce the problem to τ. And when $x_t < \tau/(1-N\alpha)$, the problem will fall below τ when pragmatists act, so ideologues become counter-productive and choose $\beta_t < 0$.

To bring the problem to τ as rapidly as possible, ideologues could choose to have effect

$$\tilde{\beta}_t^* = \max\left\{\frac{(1-\alpha N)x_t - \tau}{M x_t}, \alpha\right\}. \tag{1}$$

However, ideologues are *conflicted*. They want the problem to go down because they dislike the problem; but they do not want to problem to go down too much or too rapidly because they would lose

too much of their purpose too rapidly.[1] Compare this to a pragmatist who, facing no such conflict, wants to get the problem as low as possible as quickly as possible.

Ideologues therefore choose to act with effort

$$\beta_t^* = \max\left\{\lambda\left[\frac{(1-\alpha N)x_t - \tau}{M x_t}\right], \alpha\right\}, \tag{2}$$

where $\lambda \in (0,1)$ captures how (un)conflicted ideologues are. For analytical clarity, I will assume that

$$x_1 \leq \frac{\lambda\tau}{\lambda(1-\alpha N)-\alpha M}, \tag{3}$$

which guarantees that the max operator is never binding.

Let $\theta \equiv (1-\lambda)(1-\alpha N)$. The dynamics of the problem in agent optimality can be expressed as

$$x_{t+1} = \theta x_t + \lambda\tau. \tag{4}$$

Because it is assumed that $\alpha N < 1$ and $\lambda < 1$, it follows that $\theta < 1$, and therefore the problem is at its stable steady state when

$$x = \frac{\lambda\tau}{1-\theta}. \tag{5}$$

This can be observed more explicitly by solving x_{t+1} in terms of x_1 through iterated substitution, which yields

$$x_{t+1} = x_1\theta^{t-1} + \lambda\tau\left(\frac{1-\theta^{t-1}}{1-\theta}\right). \tag{6}$$

The result is then obvious because θ^{t-1} limits to zero.

Notice that when $\lambda = 1$, the steady state is at $x = \tau$ because ideologues act aggressively when there are deviations from τ. But when $\lambda < 1$ ideologues do not act as aggressively and so $x < \tau$. It is also interesting to note that the steady state does not depend on M. This is because ideologues act more aggressively when there are fewer of them, so $\beta_t M$ is unaffected by changes in M.

[1] You can think of this as a sort of "purpose smoothing," if you are so inclined. It is easier psychologically to deal with small changes.

Note from equation (5) that the steady state is increasing in τ, the idea being that ideologues act in counter-productive ways sooner and more aggressively for any level of x_t. Likewise, the steady state is increasing in λ because ideologues are acting more aggressively. On the other hand, the steady state is decreasing in α and N because more of the problem is being fixed per period. The takeaway is that the problem will end up at a relatively high level when ideologues act aggressively or derive much of their purpose from their identity as an activist; and the problem will end up at a relatively low level when pragmatists are effective and in large number.

The dynamics of x_t are illustrated in Figure 1. Ideologues anticipate that the problem will fall below their purpose threshold early on and therefore begin acting in counter-productive ways, shown by $\beta_t < 0$. There are enough pragmatists to keep the steady state below τ, but the counter-productive behavior of the ideologues yields a steady state above zero.

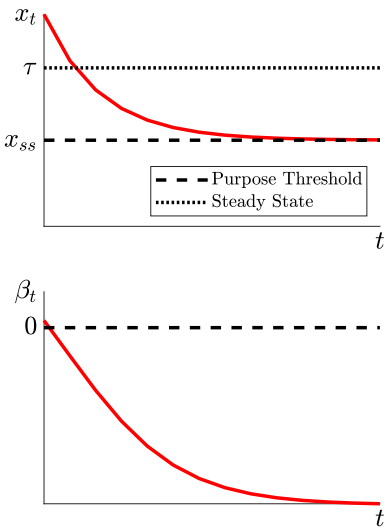

FIGURE 1: When ideologues anticipate or observe that the problem is below their purpose threshold, they act counter-productively as shown by $\beta_t < \tau$.

5. Exit from Discourse

This simple model explains why activists sometimes behave in obviously counter-productive ways, and why extreme and divisive and dubious

rhetoric (e.g. concept creep and witch-hunt mentality) and behavior increases even as the problem becomes subdued. It will be instructive to introduce dynamics of activist populations as well.

The key new assumption is that there is a negative threshold of β_t, below which ideologue behavior becomes destructive enough (e.g. rioting) to imply a non-negative probability of exiting discourse (e.g. being arrested, removed from leadership, or becoming persona non grata). Ergo over time, *radical* ideologues will be removed from discourse.

Let $\epsilon < 0$ denote that threshold, which I call the *extremism threshold*. When $\beta_t < \epsilon$, ideologues will exit discourse according to dynamics

$$M_{t+1} = \max\{(1+\delta\beta_t)M_t, 0\}, \tag{7}$$

which says that the number of ideologues changes depending on how extreme ideologues become (a large negative β_t) and how sensitive exit is to extremism, δ. If ideologues are not extreme so that $\beta_t \geq \epsilon$, then the number of ideologues is steady at $M_{t+1} = M_t$.

Because M_t is now potentially time-variant, dynamics change. Ideologues choose to have effect

$$\beta_t^* = \max\left\{\lambda\left[\frac{(1-\alpha N)x_t - \tau}{M_t x_t}\right], \alpha\right\}. \tag{8}$$

Note again that for lower M_t, ideologue effort β_t increases in magnitude. So when $\beta_t < \epsilon$, it follows that M_{t+1} will be lower; which means $\beta_{t+1} < \beta_t$; so M_{t+2} will be lower; which means $\beta_{t+2} < \beta_{t+1}$; and so on. The implication is that if ideologues are myopic and freely choose $\beta_t < \epsilon$, then it is guaranteed that ideologues will radicalize and eventually will be removed from discourse.

For $\beta_t < \epsilon$, it is required that

$$x_t < \frac{\lambda\tau}{\lambda(1-\alpha N) - \epsilon M_1} \equiv \overline{x}, \tag{9}$$

which I call the *radicalization point*. To focus on interesting phenomena, it is assumed that $x_1 > \tau$ so that x_t falls over time, and therefore approaches its steady-state x monotonically from above. For x_t to reach the radicalization point, it must be the case that the steady state is below the radicalization point, which in fact will be the case when

$$-\epsilon M_1 < \alpha N. \tag{10}$$

In words, x_t will fall far enough to incite unacceptably counter-productive reactions when there are many pragmatists who exert a lot of effort; when there are already few ideologues; and when the behavior of counter-productive ideologues is quickly considered unacceptable.

Myopic ideologues eventually disappear, upon which the dynamics of the problem will reduce to

$$x_{t+1} = (1-\alpha N)x_t, \tag{11}$$

which converges monotonically to zero.

The dynamics the case with myopic ideologues is illustrated in Figure 2. It is important to note that when society becomes less accepting of extreme behavior (i.e. ϵ is higher), then the behavior of ideologues becomes *more* extreme. Compare the dynamics of β in Figure 1 to that of Figure 2. When there is no exit (i.e. when ϵ is low enough to be irrelevant), β_t remains at a low but finite level. On the other hand, when ϵ is high enough to be crossed, β_t becomes extremely negative. In other words, when society disapproves of extremism, society gets a burst of extremism. The upside is clear waters after the radicals fires have been rapidly extinguished.

If ideologues are forward-looking, then instead of choosing $\beta_t < \epsilon$ and triggering radicalization, they choose $\beta_t = \epsilon$ exactly. This guarantees $M_t = M_1$ for all t, so ideologues are able to remain engaged in discourse as long as there is discourse to be had. However, it still guarantees that the problem will be eliminated over time, just at a slower rate. Because forward-looking ideologues deliberately reign in their reaction when facing the extremism threshold, the equation (5) is no longer characterizes a steady state; x_t always decreases. The dynamics with forward-looking ideologues is shown in Figure 3.

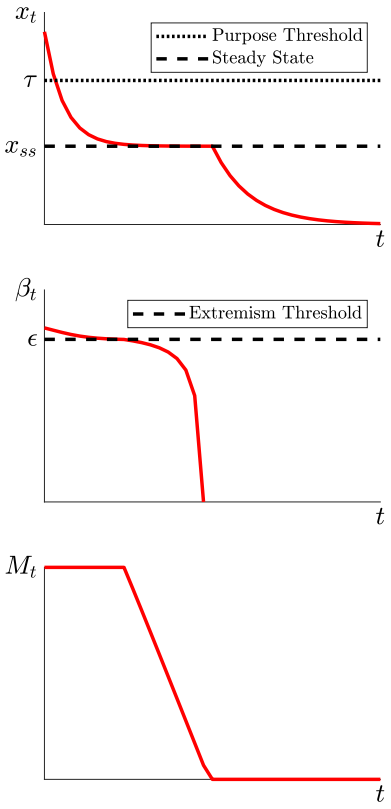

FIGURE 2: β_t eventually drops below ϵ because $x_t < \tau$, thereby inciting radical ideologues. The number of radical ideologues M_t falls over time as they are removed from discourse, which pushes remaining ideologues to become even more extreme. Eventually all ideologues are removed from discourse and the problem drops to zero.

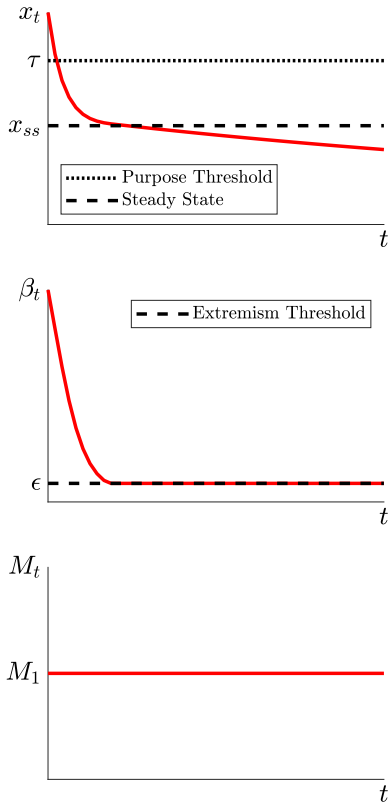

FIGURE 3: When β_t gets close to ϵ, forward-looking ideologues refuse to radicalize and therefore keep $\beta_t = \epsilon$. Accordingly, the number of ideologues M_t remains constant. However, ideologues are unable to be sufficiently counterproducive to maintain a positive level of x_t.

6. Conclusion

While simple, the model explains several observable phenomena. It explains why activists can behave in obviously counter-productive ways. It explains why even though certain problems have improved precipitously over time, and even though there are more activists with stronger tools than ever, rhetoric has become more divisive and vitriolic. It explains why a safer society can drive ideologues to bursts of extremism.

Whether bursts of extremism followed by relative calm are preferred to a more moderate society dealing with a prolonged but subdued problem is unclear. What is clear is that the incentives of activists need be scrutinized. Incentives matter, and it is in error to analyze behavior without accounting for them. When an activist becomes too attached to a cause, the potential for fanaticism emerges, to which no functional, equal, democratic society can acquiesce.

References

[1] H. Brotz, B. W. Austin, African-American Social and Political Thought: 1850-1920, Routledge, 2017.

[2] A. Binder, J. L. Kidder, Student activism is often uncivil. We can change that., The Chronicle of Higher Education (2019). URL: https://www.chronicle.com/article/Student-Activism-Is-Often/247168.