“We Have Met the Grey Zone and He is Us: How Grey Zone Warfare Exploits Our Undecidedness About What Matters to Us in Decision-Theoretically Difficult Scenarios and Conditions of Choice”[[1]](#footnote-1)

Duncan MacIntosh,

Dalhousie University

**Introduction**

Grey zone attacks tend to paralyze response for two reasons. First, they present us with choice scenarios of inherently dilemmatic structure, e.g., Prisoners’ Dilemmas and games of chicken, complicated by difficult conditions of choice, such as choice under risk or amid vagueness. Second, they exploit our uncertainty about how much we do or should care about the things under attack­—each attack is small in effect, but their effects accumulate: how should we decide whether to treat a given attack as something meriting a serious response rather than a mere irritation to be ignored? This chapter brings standard decision theoretic tools to bear on the issues, tools such as the maximization of expected utility principle, and the precautionary principle. But it also develops three innovations. It shows that seeing it as possible to make choices not just of actions given our values, but also choices of values themselves, can extricate us from certain kinds of decisional paralysis. It shows how we can rationally regard ourselves as having bright decision lines even when nothing in the circumstance of choice has these lines in its structure. And it shows how certain conflict escalation risks from the former two strategies can be managed by both parties seeing that cooperation can yield a sharable surplus of goods, and that cooperation can be rational upon a rationally defensible change in their values so that each comes to derive utility from the success of both parties. This yields cooperative solutions to what amount to Prisoners’ Dilemmas. The Cold War was prosecuted on the logic of strategic rationality and that part of game theory that is deterrence theory. We need a new logic from decision theorists for the games that constitute grey zone conflict.

In what follows, Section 1 describes the character of grey zone conflicts, taking work by Michael Mazaar as a point of departure. Section 2 discusses problems of choice under conditions of vagueness in grey zone attacks. Section 2.1 distinguishes attacks not requiring response from attacks that may be misperceived as such, and discuses how to guard against and respond to the latter; 2.2 classes some attacks as having responses dictated by clear cost-benefit analyses; 2.3 discusses attacks requiring choice under conditions of vagueness proper; 2.3 a) covers sorites or true vagueness substructures; 2.3 b), structures in which it is appropriate to use the precautionary principle; Section 3 considers attacks exploiting intransitive preferences; Section 4 the problem of escalation resulting from both parties using the previous problem solutions; Section 5, impure coordination problems; Section 6, bargaining problems. Section 7 notices that, just as drawing lines to solve the problem of vagueness in choice can result in the escalation described in Section 4, above, so can using similar solutions in resolving impure coordination problems and bargaining problems, with the result that parties find themselves in escalating games of chicken. Section 8 distinguishes zero-sum from win-win games, and total-conflict from partial-conflict scenarios, and offers general solutions to each. Section 9 summarizes the new roles for decision- and game-theorists in prosecuting the grey war in contrast to their role in the Cold War. Section 10 extracts thirteen pieces of specific advice from the overall analysis for how to maneuver in grey zone conflicts.

**Section 1 The Character of Grey Zone Conflict**

Grey zone conflict is generally defined negatively as hostility between nation-states that does not rise to the level of hot war prosecuted by their militaries.[[2]](#footnote-2) It manages not to rise to that level by virtue of having one or more of the following features: i) It involves aggressive actions insufficiently consequential to merit a military response, even though they have effects on whole nations and may therefore seem to be such that only state level military agencies have the jurisdiction or ability to make response. ii) It uses the methods of anonymity and misrepresentation and it is therefore difficult to know who to blame for the actions in question.[[3]](#footnote-3) iii) It uses non-kinetic methods or employs kinetic methods (bullets and blast) only marginally. iv) It uses techniques that, while they inflict some sort of harm, are not necessarily illegal. v) It is not necessarily conducted by nation states but possibly occurs with their indulgence or protection, and it is therefore conducted by agents who have the shelter of uncooperative states against criminal prosecution even for actions that are illegal, while not rendering the sheltering nations sufficiently guilty of aggression as to be legally liable to a self-defensive state-level hot war response. vi) It is incremental and subtle by individual deed but may accumulate to large effect. vii) It is typically aimed at non-military targets, e.g., at businesses, informational infrastructures, the deliberative processes and mechanisms of a polity, or at individuals, while yet having effects on the fortunes of and relations between nations; and because its immediate targets are sub-national entities, ones who do not have the authority or power to respond, while yet not being extreme enough to justify state-level military response, it tends to go unpunished. viii) It involves actions response to which would in principle be possible from several agencies, but thereby in effect poses problems of divided agency, since the agencies must sort out which of them shall act, this complicated by the costs and payoffs for the agencies involved, these in turn pitting the agencies against themselves and each other. ix) It is crafted precisely to be mild enough as not to evoke a squelching response, but harsh enough to advance the attacker’s cause.

Discussion of grey zone conflicts is refreshingly free of group think, with three exceptions. First, it appears to be thought that grey zone attacks are a kind of enemy, much as terrorism is often thought of as an enemy. In fact, neither is an enemy. Terrorism, for example, is just the method of various enemies, namely, using violence against civilian populations for a political goal. That apart, the pretexts and solutions to “the problem of terrorism” are as various as the specific grievances of the groups and individuals who resort to it around the world. Likewise, grey zone activity is not an enemy. It is just the method of some enemies, enemies again as various as the specific grievances of the groups and individuals who resort to it.

Second, because grey zone warfare is sometimes thought of as an enemy, it is thought that there should be a ministry of grey zone conflict, a whole-of-nation response to it, national preparedness to engage in it, and institutes for its. This conclusion is not wrong—the form of conflict does indeed need large study and large, distributed and coordinated response. But the premise that grey zone assault is an enemy wrongly leads us to forget that there needs to be specific expertise in and response to, the various occasions of grey zone conflicts, which are many, and distinct. How to think of and respond to calculatedly incremental attacks is a worthy discipline. But the causes and cures of grey zone conflict between the U.S. and the oligarchs of Russia are likely to differ significantly from the causes and cures of such conflict with, say, China.

The most troubling group think idea, however, is this: that the careful studies of the facts of various conflicts (studies which I, too recommend) will by themselves dictate solutions. This is false. We find ourselves troubled by grey zone conflicts not just because of facts about the conflicts, but because of facts about us, in particular, facts to the effect that we haven’t decided what to care about or how much to care about it in the conflicts at issue; and facts about the objectively dilemmatic nature of certain decision problems, these constituted of issues we face simply by being rational agents.

Many wise things have been written about grey zone conflict. A particularly good example is Michael Mazaar’s essay, *Mastering the Gray Zone: Understanding a Changing Era of Conflict*.[[4]](#footnote-4) Mazaar sees grey zone conflicts as caused by nations whom he classes as incremental revisionists, nations who want to see changes in the world order without destroying that order, and who want to be partners in peace but with more influence in and more profit from that order. The essay has five main recommendations, beginning with recognizing and encouraging processes in whose unfolding time is on our side. E.g., in the conflict between Russia and the U.S., the U.S. offers a superior ideology, polity and economy, things which will probably prevail in the long run. Next, to help ensure this, the U.S. must strengthen the institutions and norms which the revisionists would undermine—protect the tools, institutions and processes of deliberative democracy, for example, as well as the current global rules-based order. Third, the U.S. should recognize that not all conflicts need a military response; sometimes accommodation is fine. Next, we need to “build forces, systems, technologies, concepts, and doctrines for a gradualist environment.” And finally, we need to “punish selected revisionist acts and broadcast true red lines”—that is, we need to specify to prospective enemies what sorts of behaviours would elicit retaliatory punishments from us, and what those would be.

I generally endorse Mazaar’s recommendations. But I anticipate and seek to resolve some conceptual issues in their implementation.

Mazaar’s first and second recommendations—that we should encourage and be patient in waiting for the effects of historical inevitabilities when they are on our side, and that we should strengthen the norms and institutions on which we rely, and which the revisionists would undermine—have the difficulties that there will likely be grey in determining what counts as a significant threat to these norms and institutions, and what would be appropriate responses, which reprises the main problem. But never mind; Mazaar is clearly right that the objectives he enumerates are important. However, the last three of Mazaar’s recommendations are rife for trouble, each for the same reasons. Take the third, that sometimes we should meet challenge with accommodation rather than military resistance. And contrast it with the fifth, that we should punish selected revisionist acts and broadcast true red lines. These proposals specify not objectives, but strategies; and we can’t implement either strategy until we have decided what we care about and how much. And until then, we can’t decide how to respond to incremental challenges, which means we wouldn’t know how to go about implementing the fourth recommendation, that we should build the resources needed to help us contend with graduated challenges from revisionists.

Evidently Mazaar’s recommendations face us with problems of choice under conditions of vagueness. I shall discuss such choice in detail. We will eventually see that attempts to solve these problems lead us into other types of problem familiar to decision theorists and game theorists, namely, impure co-ordination problems, bargaining problems, games of chicken, and Prisoners’ Dilemmas.

**Section 2 Problems of Choice Under Conditions of Vagueness**[[5]](#footnote-5)

An enemy steals a penny from me (or makes a minor land grab, or hacks my computer to get some business secret or some military secret, or trolls my deliberative democratic venues with inflammatory posts, or is vaguely implicated in the death of a journalist, or finances the leadership of a discontented minority part of my population, attempting thereby to encourage a rebellion). It’s not worth the hassle to punish him. So he steals another penny. If I don’t stop him, he’ll eventually steal all my money. But each theft would be too much hassle to respond to. When should I respond? And how severely?[[6]](#footnote-6) Maybe the problem isn’t urgent after one penny. Maybe it is very urgent after 50% of my pennies are stolen, if not before. What could decide to which penny theft I should respond seriously? And what could decide how many of my remaining pennies I should spend to protect those pennies? It may seem inherently vague where to draw the line; that is, it may not be obvious what further fact I could discover that would decide the matter. (It’s difficult to say how many hairs you have to have on your head to count as not bald, and difficult not because we’re missing some piece of factual information—this exemplifies so-called sorities problems.)

Mazaar makes a proposal which might seem to side-step this problem, drawing on Thomas Schelling.[[7]](#footnote-7) (I don’t know whether Mazaar had this problem in mind when he made the proposal.) He expands on Schelling’s suggestion that to stop a man who would kick a dog, you should provide escalating penalties for each step the man takes closer to the dog. I object that since each step is itself harmless and would cost you something to penalize it, you’ll have difficulty motivating penalizing it. You won’t be moved to do anything either until the last step, when it may be very difficult to stop the man, or after the kick, when it’s too late. This illustrates a problem with Mazaar’s suggestion that we should draw bright red lines to deter incremental aggression: what could motivate us to draw the line at one (sufficiently early) place rather than another?

Many people upon hearing about such issues favor mirroring—do whatever the other party does.[[8]](#footnote-8) (If the man takes a step towards the dog, you take a threatening step towards the man.) This advice is a variant on tit-for-tat strategy, which says always offer peace first, but if you are then met with aggression, aggress back, and meet every aggression with one of your own, but meet olive branches with olive branches. The hope is to teach your enemy to be nice by penalizing nastiness and rewarding niceness. (If the man takes a step back from the dog, take a step back from the man.)

This strategy will induce niceness under some conditions, but not all (e.g., not where there are more inveterately nasty people than inveterately nice). And it has a disadvantage, one which is part of the explanation for why the strategy won’t always work: you’re going to take a hit at the outset if your enemy is bellicose enough. (The man gets a step on you towards the dog.) This may recommend you striking pre-emptively with a hit of your own if you suspect you are about to face a grey zone attack (you take a threatening step towards the man before he does anything), forcing the enemy to accept the first loss (he has lost ground on the matter of the dog).

But a further difficulty with mirroring is that it costs something to engage in a mirroring act, often more than the cost of just absorbing the hit you propose to counter. It is likewise costly to pre-empt, so it will be tempting to delay, hoping there will be no aggression after all.

The problem of how to deal with conflict structured by vagueness is really two problems. The first is how to see it as rational to decide to treat a point on a vague continuum as an inflection point, one arriving at which will trigger greater action. The second problem is how to see it as rational for the parties solving the first problem not to escalate into levels of conflict neither really wants.

Let us consider in more detail the first problem, that of choice in vagueness proper. There are several possible structures that grey zone attacks may exemplify, and we can see the problem by considering the structure that features difficult vagueness alongside some others that don’t. This differential diagnosis will be useful in telling us how to respond to the various sorts of attack.

**Section 2.1 Structure 1: Attacks Really Not Response-Requiring vs. Attacks Illusorily Not Response-Requiring**

I define structure 1 attack types as ones in which it may seem that each attack is trivial, that any number of such attacks would be tolerable, and that punishing any of them would be more costly than accepting them. In that event, arguably there is no problem; you needn’t do anything, and indeed, you rationally shouldn’t do anything. In fact it may then be problematic to see these things as attacks; maybe any vestigial sense that one should respond needs philosophical therapy.[[9]](#footnote-9) Perhaps some so-called grey-zone *conflicts* should really just be seen as acceptable forms of bargaining, or as some other acceptable sort of action.[[10]](#footnote-10) It might be a mistake to represent moves in them as attacks, unjust harms, and so on. They are just pre-contractual maneuvers, actions in a state of nature lacking the established norms needed for some action to count as wrong, bad, inappropriate. (Think of the conflict between Capitalist and Communist countries over ownership of intellectual property; or between countries with differing conceptions of fair competition and the role of government in subsidizing industries.) And our objecting to them merely shows that we should see the issues which source these maneuvers as needing contracting. Until then, the actions are not assaults, free ridings, or any other form of legally recognized harm; for they are not yet illegal, not externalities to an extant deal. Or maybe they represent only borderline infringements on such a deal; or actions that seem hard to classify—they are sort of infringements, sort of not. One country financing separatists in another country, as Russia did in the case of the Donbass region of Ukraine between 2014 and 2021 may be an example.[[11]](#footnote-11)

On the other hand, there are scenarios that may have problematically led to the forgoing willingness to accept the so-called attacks. Suppose I’ve been drawn by steps to which I do not individually much object to a place I would have strongly protested at the outset: has anything bad happened? Maybe I’ve just been induced, by relatively unobjectionable means, to evolve in what I find acceptable. This would be fine if my basic values have remained unchanged and I’ve just discovered that I need to do different things to advance them, or discover that unexpected things count as satisfying them, or if I discover that they were premised on a factually false assumption. But in fact it may be that I’ve in effect acquired so-called adaptive preferences, preferences for a lesser state of affairs that are acquired in stoical resignation upon recognition of the impossibility of attaining a better state.[[12]](#footnote-12) (If I can’t get Scotch, but beer is still available, I may come to prefer beer to Scotch.) And the rationality of forming adaptive preferences is contested. For it seems that unless I retain my current preference I won’t be motivated by it to advance it; so me undergoing adaptive preference change is the same as failing to advance my extant preferences and is therefore irrational. And even if I’ve come to desire a new situation and so am happy by the measure of desire satisfaction to be in it, arguably the acceptability of the situation should have been evaluated by my pre-change desires, which means my undergoing desire-revision was self-defeating. Telling me you’re going to cut my arm off but not to worry, I’ll get used to it, does not make me coming to desire to be one-armed rational.

But there is a way to think of rational agents on which acquiring adaptive preferences might be perfectly rational: imagine someone who most wants pleasant feelings, and who gets them by having other desires of theirs satisfied. So what they want is as much desire satisfaction as they can get. And there are two ways to get it: alter their circumstance to bring about the ends they desire, or change their desires so that what they desire is what they can get in the circumstances. Either way, they will have the pleasant feeling of desire satisfaction. So if the agent in question is like this, then perhaps their stance is unproblematic. Their new preferences aren’t really ones they have acquired simply because they can’t get their old ones satisfied; rather, they are preferences they have adopted for instrumental reasons—as a way of advancing their desire to have preference satisfaction. They’ve changed their attitudes towards some possible ends as a means of attaining their most highly valued end.

There are, however, doubts about the coherence of such a psychology. For what is it to have a desire, other than to be moved to do what one believes would bring about what it is a desire for? So how could changing one’s desires in these circumstance advance that cause? Meanwhile, if one also wants desire satisfaction in general, indifferently to having it by a) changing the world to attain satisfaction of one’s current desires, or b) changing one’s desires to have them fit what is available in the world, then how can one choose actions, especially where both options are available? Having a desire would seem to entrain the rationality of trying to change the world to fit it, but having a general desire for desire satisfaction would seem to entrain the rationality of not doing this, and instead simply renouncing the desire. And now one is hopelessly conflicted, and should therefore be paralyzed about which course to take.

There is more to say about the intelligibility of the foregoing moral psychology,[[13]](#footnote-13) but let’s move on to another scenario that may have led to me rationally finding both individual aggressions and their sum tolerable, even if irritating: my non-reaction to these steps may be rational on any measure due to facts about the concurrent or evolving contexts in which they occur. E.g., maybe no step is worth punishing, nor any accumulation of steps. Perhaps, for example, in the meanwhile my wealth rises, so that, while each step costs me money, no step individually or cumulatively costs me more than the price of policing further steps. Or perhaps while each step is a harm, it is accompanied by benefits, and the latter would be disrupted by me policing the former. An example would be that, while Russia is harassing our democracy, it is also an ally against terrorism, and its rockets ferry supplies to our space station. In that event, perhaps Russia’s harrying is just part of the price for the useful things. Indeed, perhaps we feel we in some sense owed some of the value to the party whom we would represent as thieving or aggressing, perhaps from some prior action of ours, or as a not unreasonable demand for influence in their sphere. As long as the value they give exceeds what they take, there is no irrationality in our non-response. Or maybe the cost to them of their aggressing is one we’re happy to seem them pay, so our non-reaction is a rope-a-dope play.

But what about this: suppose my enemy attacks me in some minor way, then eases off for a bit, then attacks in some further way, and so on. With this rhythm, perhaps they prevent me from having the experience of an accumulating grievance. Rather, I experience each event only as an isolated episode. Could my non-reactivity be rational? Or am I irrationally allowing to be used against me my psychological tendency not to connect dots interrupted by large gaps?

Relatedly, suppose my enemy regularly does things I find a little outrageous, constantly harassing me. But in so doing, they wear me down, lower my expectations about their good behavior, weary me, and make me think they would stop only if I did something ridiculously escalating. Have I been rational? Have I been manipulated? Both?

Or what about this: my enemy continually harasses in a way that I begin to think is just in their nature, there’s nothing to be done, one just has to accept it. So they have changed what I think my options are. Again, have I been rational? Have I been manipulated? Both?

Or this: my enemy continually harasses me in ways expensive to me, but dangles the possibility of an immense assault should I retaliate. We see this in the conflict between Russia and Ukraine as I write (in March, 2023): Russia’s leader, Vladimir Putin, is nibbling away at Ukraine territory, induction suggests he’ll keep doing this hoping to enlarging Russia to the size of the former Soviet Union; and he threatens NATO with nuclear weapons should it intervene. The odds of him using them are fantastically small given that this would entrain as much harm to him as to his enemies. But the mere chance of their use is experienced as deterring. Here arguably we have been threatened into using the disaster avoidance or precautionary principle as a decision principle, rather than the principle of maximizing expected utility. And this seems irrational on our part. For the effect is to make us care in lesser degree about our ethical values and assets, or at least to make us act as if we cared less about those things.

I should expand a bit on this: the precautionary principle says to choose the option that minimizes the possibility of the occurrence of the worst possible outcome. The expected utility principle says to choose the option that makes as high as possible the sum of the product of the probability and the utility of the possible outcomes of the actions open to you. The latter principle could justify you in doing an action that has a good probability of giving you a good outcome, but also a very small probability of a very bad outcome. Most decision-theorists think the latter principle is the correct one to use whenever you know the desirability to you of each outcome your actions might produce, and the probability of its being produced given each possible action available to you. The principle best tells us how to balance risk and reward. Meanwhile, in such circumstances, the precautionary principle makes us excessively cautious. In the case of Russia and Ukraine, our focusing solely on making sure there is no chance of our triggering a nuclear attack prevents us from acting on other values, e.g., protecting the innocent from aggressors, supporting nascent democracies, and thinking it better to die than to live unfree. Surely these things are important enough that we should act to advance them even if this means a very small chance of a nuclear attack. But the threat of nuclear attack mesmerizes and makes us make choices reflexively against our values. Of course, it might be argued that the fact that it mesmerizes proves that we value not being victim of nuclear attack so much that, in order to avoid it, it is worth it to forgo advancing other things. But I suspect there is something more like a startle response in play here. You accidentally walk into a spider web, you panic, and you start doing spider web kung fu. But in fact you don’t think spiders are a big deal, and you are now acting from reflex, not in a way that rationally expresses your values all things considered. Of course a nuclear attack would be a big deal. But the horror of it, properly discounted by its extremely low probability of occurrence, is a much smaller deal.

Finally, I may be paralyzed by unresolved ambivalence about the morality or justice of my opponent’s action. E.g., Russia is currently trying to conquer Ukraine. Many in The West see this as a morally outrageous and gratuitous act of aggression. But (as we noted earlier) some say this is just Russia rightfully protecting its sphere, justifiably resisting NATO encroachment. And until we sort out which attitude is morally right, we won’t know how to proceed.

In light of the foregoing discussions, we must distinguish between several sorts of cases:

i) My preferences are induced to change adaptively. Arguably this is irrational and I should have resisted it by keeping in mind the preferences I began with, and continuing to focus on what actions would advance them.

ii) My preferences remain the same but my estimate of the odds of any available action being able to advance them has gone down. The only risk of irrationality here is epistemic: I should ensure I haven’t been fooled into thinking the situation intractable.

 iii) I just somehow have my will corroded so that, even though my values and my estimates of the odds of the actions open to me successfully advancing them remain the same, I just don’t act. This is me being weak-willed, something generally thought irrational (although there is more to say case by case: e.g., it isn’t necessarily irrational to fail to act because one is exhausted).

iv) I’m lulled by the intermittency of the attacks into not seeing their cumulating effects. Here, I should have kept a tally and decided what to do in light of a full awareness of my losses as they increase.

v) I’m intimidated into making calculations by the precautionary principle rather than by the expected utility principle. Here I’m being irrational. I should remind myself of my priorities and my estimates of the probabilities of the possible consequences of the possible actions open to me, do an expected utility calculation, and act accordingly, choosing the action with the highest expected utility.

vi) My ambivalence about which ethical lens to use in viewing a problem is being exploited to paralyze me. In this case, I should do moral arithmetic and proceed accordingly.[[14]](#footnote-14)

**Section 2.2 Structure 2: Attacks Retaliation to Which Is Rationally Dictated By Clear Cost-Benefit Analyses**

In structure 2, each attack is trivial, but the attacks have accumulating negative effects; and at some known point, the cost of retaliating would be less than the cost of the accumulating negative effects. In that case you rationally must retaliate to the last acceptable attack—before then would be needlessly costly, after then, excessively costly.

**Section 2.3 Structure 3: Attacks Requiring Choices Under Conditions of Vagueness**

A third possibility: each attack represents an unpleasant but acceptable damage, but also contributes to an accumulating damage that is unacceptable. In such situations there are two further possible sub-structures:

**Section 2.3 a) Sorites or True Vagueness Sub-Structure**

There is a zone in which the attacks would clearly not have accumulated to something worrisome, another zone in which they clearly would, and in between there is a zone where it’s not clear what to say; moreover, the zones have vague borders—the transition between one and the next is not marked by a definite increment. Here, you need to pick a point within the clearly non-problematic zone to make serious responses. Until recently the theory of rationality could not see doing this as rational, since for any point you pick there would be just as good reason for picking another, or, worse, there would be advantage to delaying until a later point. (You are given a bottle of wine and told that for each day you delay opening it, it gets better. No matter which day you pick to open it, there is always the argument that you should wait another day, because it will then be better. But if you follow that argument, you’ll never open the wine, and so never benefit from the bottle.) Scenarios with these characteristics are in the family of so-called Sorites Problems. But as I have elsewhere argued, since regarding the fact that a delayed choice point will always be better as a reason to delay can be predicted to lead you to disaster (you never get to drink the wine), it is irrational to regard that as a reason to delay. So regarding it is not maximizing of one’s expected utility. Instead, it maximizes to take the attitude that some arbitrary point within the good zone must be chosen, chosen by a symmetry breaking technique, like flipping a coin, or throwing a dart at a representation of the continuum, then treating the point so-chosen as the one at which to get serious.[[15]](#footnote-15)

 Chrisoula Andreou argues that such scenarios are counter-examples to the theory that, in each action, a rational agent would always make the choice that maximizes her expected utility. Instead, a rational agent would sometimes make the choice that is part of a course of action the taking of all parts of which would be maximizing in sum. This could result in an agent taking a non-maximizing action provided it was part of a maximizing total course. Not procrastinating retaliation indefinitely is a maximizing course, procrastinating indefinitely, non-maximizing. But to cease procrastinating, one must do a non-maximizing action, vis., take action now, even though it would be maximizing in that moment to delay yet again.[[16]](#footnote-16)

I think the actions Andreou recommends are right, but not because of the rationale she offers. Rather, an agent can see that it is non-maximizing to procrastinate indefinitely, and can see that unless, using a symmetry breaking technique like flipping a coin, she arbitrarily chooses a point to stop procrastinating, she’ll never stop. So she uses the technique to pick a point because using it is maximizing, then she follows the technique because she committed to using it, a commitment I would portray as her coming to prefer to take the action that the technique dictated, which means that she’s maximizing on her preferences even in following through on the outcome of the use of the technique—it has become more important to her to take the recommended action than to pursue further benefit as that was formerly defined by her previous preferences. Thus at no point is she failing to maximize action by action.

At any rate, in political situations where we face choice problems with this true vagueness structure, if we as a nation do not take the foregoing course of action, we are guilty of weakness of will.[[17]](#footnote-17) After all, if we don’t make an arbitrary choice on the continuum, we will never get to consume the good in question, which would be irrational. So now we must seek for the causes of our failing to do this. And so we might ask, drawing on the literature on weakness of will, how to cure ourselves of this tendency. The cures that have been proposed are diagnostic of the problem they are designed to cure: e.g., creating scaffolds to our will, scaffolds in the form of plans (because we have none, and it is hard to implement a resolution without a specific plan), new laws (because the old don’t direct us to respond in some specific way to the problem), the creation of institutions to deal with the problem (because all the current institutions think it is someone else’s problem), clarifying lines of authority, subsuming multiple and competing authorities under one authority—divided authority is something many have complained about in America’s response to grey zone attacks; and apparently this was a factor in the failure to prevent 9-11. No one knew who was supposed to prevent or respond to the catastrophe.

**Section 2.3 b) Precautionary Principle Sub-Structures**

The other possibility is that it is known that sooner or later a zone of disaster would be entered, but it is not known exactly when. In that case the rule to follow is clearly the so-called precautionary principle, which says to minimize the maximum possible disaster and act immediately. For here, we lack the information required to operate the expected utility principle.

**Section 3 Structure 4: Attacks Exploiting Intransitive Preferences**

A fourth possible attack scenario is one in which each attack is unpleasant, responding to it would be even more unpleasant, and yet the attack contributes to accumulating effects that are increasingly unpleasant. Here, many decision theorists would think agents would be irrational if they did not act to prevent the very unpleasant outcomes.[[18]](#footnote-18) But, as I have elsewhere argued[[19]](#footnote-19), since allowing each new outcome is preferred to acting to prevent it, the agent who does not prevent the attacks is not irrational. They just have value structures odd for being intransitive—the agent would rather not have had all these attacks, but would also rather not have done the things needed to not have them. That is, the agent prefers one attack to none (where they can only have none attacks if they deter the first attack, which they prefer not to do), they prefer two attacks to one (because, again, they don’t want to go to the trouble of deterring the second attack), but they also prefer no attacks to two (or, for any number of attacks, n, greater than 0 attacks, they prefer 0 to n).[[20]](#footnote-20) Perhaps this is rational behaviour; and perhaps intransitive preferences can afford a perfectly good rationale for not having to deter a sequence of grey zone attacks.[[21]](#footnote-21)

On the other hand, some would argue that you are being “money-pumped”, and that you should object to this. The following is an illustrative case, and affords a definition of the concept: suppose that, given a choice between a brooch and watch, you’d pick the watch because it can tell time. Given a choice between a watch and a pen, you’d pick the pen because you can write with a pen. And given a choice between a pen and a brooch, you’d pick the brooch, because it’s prettier than the pen. Further, you so prefer each option to the other that it would be worth a dollar to you to trade one for the other. Then you can be money-pumped: you’ll keep offering another dollar to move from having one item to having another, in an endless circle of costly trades. How can you rationally escape being money-pumped? It might be suggested that you would need to recognize that your preference cycle is costing you, and that you would have to have transitive preferences about the cost. You’d then pick an arbitrary point to exit the cycle; for you’d realize that you care about having more money than less compared to having any of these objects. You have intransitive preferences ranging over these objects, but transitive preferences over money, always preferring more to less.

Unfortunately, all this does is present you as conflicted; it doesn’t successfully justify its proposal about how you should resolve the conflict. You’d rather have more money than less, but also rather trade a dollar and the item you currently have for the next item offered to you. What should you do? It seems you have to give up part of your preference psychology. And yet, since your psychology is conflicted, it seems to give you no basis for choosing which part to give up.

It is not clear whether there can be creatures with intransitive preferences. But if there can, then if a country’s population has transitive preferences, not intransitive, they should not pick a leader with intransitive preferences, nor one with both transitive and intransitive preferences; for such leaders cannot rationally advance their population’s goals.

**Section 4 The Problem of Escalation**

Let us now return to the scenario of Structure 3 a), above. I said there were two problems with attacks that are grey zone attacks in virtue of implicating vagueness. The first was how to rationally justify picking a point in a vague continuum to decide to act, the problem which we have just discussed. Applied to grey zone conflict, the solution is to create artificial bright lines the crossing of which by your enemy will make you retaliate. You would randomly pick one among the acceptable enforcement points in the spectrum, and one of the acceptable methods/costs. This is done by making a choice of values at the behest of the outcome of the use of a symmetry breaking technique you have resolved to implement, vis., by coming to regard this increment as a difference in quality, not kind. You are choosing to be aggrieved, and are not aggrieved until you so choose. You don’t discover that some harm to you is outrageous. Rather, you undertake to treat it as such. Note that, on this proposal, you are not just picking a moment at which to act, then acting. Rather, you are picking a point, deciding to place a certain value on that point, then acting from that value. The proposal is that you must undergo value change, and undergo it as part and parcel of action selection. You are to do action selection by means of value change selection. For without value change selection, you couldn’t action select, because you’d have the usual reasons to backslide and delay, and so to renounce your choice of action.

An alternative way to construct a point at which you would act would be to pick an arbitrary deontic principle (a rule), which may amount to the same thing as just designating a level of welfare loss as intolerable.

So suppose you acquire deontic attitudes[[22]](#footnote-22): to return to the penny example, you come to regard the theft of some specific number of pennies as an outrage (for violating the rule, “thou shalt not steal *this* outrageous amount”). Once you figure out what you count as an intolerable outrage, and what you’re prepared to pay to correct it, you can maneuver. (Apparently it is a social fact about primates like us that we’ll go to fantastically costly lengths to punish unjust violations of rules, even when the costs are vastly disproportionate to those inflicted on us by the rule violations.[[23]](#footnote-23) The rationality of this is contested.[[24]](#footnote-24))

On the other hand, if I come to regard your refusal to concede pennies as an outrage, and decide I’m prepared to expend resources to defeat your retaliation, our problem returns. (Perhaps I think you were never entitled to the pennies.) It returns multiplied; for now we each have a rationale for escalation. Thus there is a danger that we have embarked on a game of chicken, like two teenagers racing towards each other in cars, each seeking to prove his superior bravery by forcing the other to veer off. The longer the game goes on, the greater the likelihood of death for both parties. And, obviously, converting the problem of choice under vagueness into a game of chicken carries with it new troubles: how are the participants to decide when to keep pushing and when to concede? The more we push, the more we risk disaster, but also the greater the chance that the other party will lose their nerve. So this merely reprises the problem of where to draw the line and decide that we are merely throwing good effort after bad.

The strategy may work, if it tips the enemy into a space they don’t want to go. But it may not. For if they are using grey zone techniques, since these are so cheap, probably there are further prices they are prepared to pay to attain their goal. This may issue in an escalated conflict. Such conflicts are still contests of wills, after all, and are won only when one party loses the will to continue.[[25]](#footnote-25) The result then may be more expensive grey zone conflict.

Another possibility is to make the conflict non-grey by responding kinetically (blowing something up or shooting someone). This could work if the enemy was never prepared to engage in a shooting war. But such response is likely to be perceived by many parties as unjustly disproportionate. As I write, this is what has happened as Russia, feeling threatened by the possibility of Ukraine joining NATO, and so feeling victim of grey zone maneuvering, attacked Ukraine in a hot war. Now most of the world is on Ukraine’s side against perceived over-reaction by Russia.[[26]](#footnote-26) So radical escalation can be self-defeating. The lesson here will not much constrain liberal democracies, who are already resolved to maneuver justly in accord with international law and the laws of war, and so who would rather keep escalations to legal things, like sanctions. But the lesson should be cautionary to nations less committed to the rule-based international order.

We shall return to the possible consequences of the above discussed strategies, but first, on to another sort of problem featuring in grey zone conflict, namely…

**Section 5 Impure Coordination Problems**[[27]](#footnote-27)

You and I would be better off making and keeping to a deal to split the profits from some possible joint enterprise than with no deal and each continuing in individual ventures giving each of us some meagre reward. There are only two possible deals available. One slightly more profits me in some way not shareable (e.g., giving me more dignity, a reputation as a power broker, perception of my style of government or economy as superior), the other, you. Which deal should we agree on? And how should we decide? This logic often characterizes situations in which nation states are engaged in proxy warfare against each other by backing different parties in regional conflicts in parts of the world distant from both sponsoring countries. The conflicts are typically brutal civil wars, they are expensive for both sponsoring states, both would love to end the conflict, but each somewhat favours a somewhat different resolution. The same for the directly warring parties. Examples abound in The Middle East.

If there is such a conflict between you and me, it might be thought that any given impure co-ordination problem goes away to the degree that I decide to care less about the minor advantage to me of one of the deals, and more about the advantage of there being some deal rather than none.

On the other hand, the problem would equally disappear if you went that way. So maybe I should hold out for that. And now we are, again, in a game of chicken, again, with an incentive to escalate, but with a recognition that, unless at least one of us discovers a limit to this strategy, disaster awaits.[[28]](#footnote-28)

On to one last form of problem featuring in grey zone conflict, namely…

**Section 6 Bargaining Problems**

You and I are aggressing against each other. Each of us gets some benefit from each of our aggressions. But each of us would be even better off if we both stopped. This is proposed. But I say I’ll only stop if you let me have two free aggressions. Should you take the deal? Am I reasonable in demanding it? (The aggressions could be insults in a marital spat— “I just want to make a couple more points.” They could be small encroachments across a national border in a land grab, or hacks of the opponent’s computers to steal business or military secrets.) What would a rational bargainer do?[[29]](#footnote-29)

It might be thought that any bargaining problem I face goes away if I realize you won’t agree to a deal that involves more pain to you than another possible deal, and won’t keep to any deal that leaves you desiring revenge for pre-deal actions from me.[[30]](#footnote-30)

On the other hand, the problem would equally well go away if you came to this realization. So maybe I should hold out for that. We are in an impure coordination problem again, the goods at stake that cannot be shared constituted of the requested additional aggressions. And we again find ourselves in a game of chicken, hoping the other concedes first.

**Section 7 Common Structures to Escalation-Prone Grey Zone Problems**

We saw difficulties with the idea that the above problems could be resolved by one or the other party becoming intransigent: if I decide I’ll destroy you for any penny theft, for any solution to an impure coordination problem that doesn’t favour me, for any solution to a bargaining problem that involves me conceding the right to inflict more pain to you, then maybe I can back you off to your second best choice. This is full commitment to see the game of chicken through to the possibly disastrous end. But my enemy has intransigence as a strategy available too. Of course, the U.S. is the most powerful country, so can’t it prevail by intransigence? No. For it is not more powerful than all the other countries combined; brutal intransigence from it would only induce a coalition against it; the U.S. could find such intransigence suicidal in the form of courting nuclear threats from other nations; and in any case, the U.S.’s deployment of great force would have the self-defeating effect of destroying other goods, like international trade, on which the U.S. depends.

**Section 8 The General Form of the Solution to Escalation-Prone Grey Zone Conflicts**

In considering possible solutions to escalation problems, there are two important distinctions to be added to Mazaar’s analysis of the competition between incremental revisionists whom he sees engaging in grey zone conflict: the distinctions between zero-sum and win-win games, and between total-conflict and partial-conflict games. In zero-sum games, one party can improve their position only if the other’s position is worsened, while in win-win games, both can improve their situation if they coordinate. Relatedly, in total conflict scenarios, each party desires the non-satisfaction of the desires of the other party, so that, logically, neither can improve their position without the other’s position being worsened; while in partial conflict scenarios, both parties can improve their positions by cooperating in making and keeping to some agreement, but each party is incentivized by the prospect of even greater gain, and by the need for self-protection from the other party’s seeking same, to defect from any such agreement. Zero-sum games aren’t necessarily total conflicts. It’s just that both parties want the same thing and they can’t both have it; but the reason they want it isn’t to frustrate the other party’s wants. Likewise, win-win scenarios aren’t the same as partial conflicts. For in win-win scenarios, neither party can improve its position unless both parties’ positions get improved; while in partial conflicts there is a way for both parties’ positions to be improved, but also the possibility of each improving its position even more at the other’s expense.

 In escalation scenarios in zero-sum games, there is little alternative to seeing the game of chicken to the end, increasing the costs of competition to the other in the form of the risk of disaster in the hopes that they will desist. We can see these as contests of wills in the sense that the parties are discovering who wants what is at stake more, and who has more resources to expend in seeking it. But there may still be resolution short of maximum escalation; for it often happens that one side can attract more allies to its cause, or more powerful allies, and so prevail. Or a third party will have its own stake in the conflict and try to create a solution by force or positive incentive. There can then arise a meta-conflict to the original conflict if another third party takes a conflicting position with the first. And so on.

 Meanwhile, it is always worth investigating whether a situation really is zero-sum. Russia and Ukraine both want Crimea. Well, why? For its seaport? Its resources? Surely these could be shared, or there could be lease or free trade arrangements. Surely these options would be better for both sides than grinding war. On the other hand, if the issue is national identity, or national pride, or anything else that involves relative standing, then hot war may seem inevitable. If two hockey players both want to make a million dollars a year, that could be arranged. If each wants to be uniquely the highest paid player, someone must accept less. But then, why does relative standing matter? Does it matter in itself, or only because it is a means to something else? Maybe Russia wants to rule Ukraine because it sees no other way to be secure. And maybe Ukraine wants to rule Ukraine because *it* sees no other way to be secure. But then maybe the rest of the world could provide security arrangements to each, and civilized relations could resume. (All of these ideas have been floated about the conflict.)

The lesson: people can think they are in an intractable situation requiring escalating violence when they are not. So an important tool in grey zone conflict is investigation into ways to see the situation as a win-win scenario or as merely a partial conflict. E.g., while the parties may be in a zero sum game about one sort of thing, they may be in a win-win or partial conflict about another—perhaps only one of us can put a church on that hill, but cooperating on mitigating climate change is more important. So we are not all things considered in a zero-sum game.[[31]](#footnote-31) Indeed, the kind of game the parties find themselves in can evolve with changes in their fortunes and attitudes. Thus the value of decision-theoretic wisdom, political scholarship, diplomacy, and mediation. And, psychology: for by investigating the motivations behind conflict, there is the hope of arriving at genuine human needs, things it may be easier to satisfy than the imagined urgencies officially premising conflict. (What do leaders bent on conquest *really* need? What does the child in them need? And how can we give it without ruining the world?) And we may discover a win-win or partial conflict conception of the situation, the latter having the potential for a cooperative solution.

On, then, to partial conflicts. One might think that if two parties are engaging in grey zone belligerence, that is evidence that a cooperative solution cannot be had.[[32]](#footnote-32) Quite possibly. Such solutions are available only where, given the parties’ desires and circumstances, there is a surplus of shareable goods to be had from cooperation, if only each party could come to trust the other to do its part. That is, cooperation is possible only where there is not a total conflict, where each party derives welfare directly from the others’ illfare.

Total conflict is rare. More commonly, each party benefits from a circumstance that can cause the illfare of the other, but neither party intrinsically delights in that illfare. Mazaar’s analysis of grey zone conflict implies that it is not intractable in the way of total conflict—the parties merely want a greater share of goods. This may mean they are in a zero-sum game. But it doesn’t follow that they want each other to suffer unless that is a means to a new sharing arrangement.

However we need to know more about the specific grievance in any given grey zone conflict before we can generalize to solutions. [[33]](#footnote-33) Some, for example, might be solvable by resource sharing. Others, only by the extraction of a pound of flesh in vengeance. Yet others might require only an apology. Or a promise of future non-belligerence.

A related concern is that some people just want power over others.[[34]](#footnote-34) Think of Vladimir Putin and his desire to subordinate Ukraine. But even these conflicts may be only partial. Perhaps the party seeking power is prepared to offer the other party something it would accept in exchange for standing in the subordinated relationship, so that both are winners by the measure of their values. There is more to say here, of course. E.g., if you offer to inflict less violence on me should I become subordinate to you, have we really made a trade in which we both benefit? Perhaps. Maybe you value being the boss more than I do. Maybe I value peace more than you do. So we trade. But let us move on now to cleaner cases of Prisoners’ Dilemmas, ones requiring less cynicism to see as able to be solved by cooperation.

We saw that grey zone conflict can require cultivating intransigence as a defense from attacks, and that if both parties do this they find themselves disastrously in escalating games of chicken. How are they to de-escalate? After all, each is such that, if it goes intransigent, and the other does not, then the first party wins by the measure of its current values and visions. Unfortunately, this is true of both of parties, so it cannot be relied upon as a unilateral strategy. Nor can the opposite strategy of being yielding. For if one party takes that strategy and the other party does not, the first loses.

Welcome to the Prisoner’s Dilemma, that game-theoretic situation in which one party can do well only if the other does poorly, but where both can do fairly well by making and keeping agreements not to exploit each other, and instead to share the fruits of cooperative enterprises. (In chicken, maybe the drivers call off the game after each has shown enough bravery, thence to enjoy the cooperative surplus of long lives.)

To solve some grey zone problems, then, requires cooperatively solving the Prisoner’s Dilemma. Can it be so solved?

In abstract Prisoners’ Dilemmas, there is a pre-defined possible surplus of goods from the parties cooperating, and a pre-defined distribution of the goods to each party. The problem then becomes how the parties are to trust each other to do their parts in the solution. But in real world Prisoners’ Dilemmas, because the conflict involves destructive actions, the magnitude of the surplus is greater the sooner the conflict ends. This incentivizes speedy resolutions. However, the share of the surplus that will flow to each party must be bargained out. And each party can demand a greater share by threatening to destroy some of the surplus—that is the effect of a party prolonging the conflict. Fortunately, two factors push against extreme bargaining strategies. First, since the parties don’t wish each others’ destruction, only a greater share of goods, each has a stake in the other doing reasonably well—each will benefit from the other as a trading partner, for example, so each will want the others’ economy to thrive well enough to be able to make purchases. Second, the parties see this as a grey zone conflict, so each wants to avoid hot war.[[35]](#footnote-35)

These factors limit escalation inclinations, and inform what bargain for distributing the surplus is likely to be agreeable. But bargaining theory does not offer a consensus on what the exact distribution ought rationally to be. No matter; parties manage to muddle through.

After bargaining, how are the parties to trust each other to keep to the bargain. David Gauthier has a plausible proposal:[[36]](#footnote-36) each party is rational in becoming such (i.e., in acquiring a disposition to the effect) that it will cooperate with the other in mutually beneficial peace provided the other is likewise disposed. The disposition’s adoption of by either party would incentivize same by the other—for the promise of a share in the fruits of cooperation. And yet each is protected from vulnerability since action on the disposition is called for only if the disposition is detected in the other.

The problem now becomes an Assurance Game: how can each party assure the other that it will cooperate? Traditionally this is thought possible only in so-called Iterated Prisoners’ Dilemmas, situations in which the parties will interact repeatedly with agents who will know whether they have kept cooperative commitments, and so be incentivized to cooperate to preserve their reputation as cooperators, thence to induce cooperation from others with the aforementioned disposition.

But Gauthier thinks it possible to signal, by taking steps to be transparent, that one has acquired a reliable disposition to be cooperative, and so can be expected to cooperate even without the possibility of enforcement in later iterations.[[37]](#footnote-37) Indeed, I have elsewhere suggested that one must come to value cooperation with other conditional cooperators as valuable for its own sake. Or one must become such that one’s utility is partly conditional on that of the other party—we must each come to feel a stake in the desire-satisfaction of the other.[[38]](#footnote-38) Only in these ways would we have become inclined to refrain from exploiting other parties who have behaved trustingly towards us in non-iterated dilemmas, situations where it would be possible to escape punishment for breaking cooperative bargains.

Mutual intransigence is inherently escalating; mutual love, de-escalating. The resolution of the foregoing problems, then, requires all parties coming to have values and visions with the property that, if all parties adopt them, all will then be moved to act in ways that will improve the positions of everyone by complying with cooperative solutions to our Prisoners’ Dilemmas.

**Section 9 The Cold War and The Grey War**

While the central decision theoretic problem of the Cold War was whether it was rational to form and fulfill intentions action on which would leave everyone vastly worse off, the problem of the “Grey War” is whether it is rational to form and fulfill intentions the mutual fulfilling of which would leave both parties better off than had neither ever formed them.[[39]](#footnote-39) If we act as we commit to acting as part of threatening in the Deterrence Paradox game that defined the Cold War, everyone is dead. But if we act as we are to commit to acting in cooperative solutions to the Prisoner’s Dilemma, everyone gets a share of the cooperative surplus. The Deterrence Paradox is necessarily a one-shot game—there will be no survivors for iterations. Grey War Prisoner’s Dilemma typically iterate, however, so the problem is easier to solve. In common with the solution to the Deterrence Paradox, the decisions that will cooperatively resolve the Prisoner’s Dilemma will involve not just what actions to do, but also what attitudes to have, in particular, concerning what things to care about and how much. Additionally, it will involve decisions about what degree of altruism to adopt to the other parties in the conflict, and about what bargains to come to see as just.

**Section 10 Conclusion**

This chapter’s contribution to those decisions has been to make the following specific recommendations:

1) Eschew adaptive preferences.

2) Guard against enemies who posture as more intransigent than they really are.

3) Do not get worn down by repeated bad opponent action.

4) Do not be intimidated into using the precautionary principle instead of expected utility maximization.

5) Do moral analysis to decide whether one’s opponent is morally right in its actions, this in order to avoid the paralysis that comes of being ambivalent on the matter.

6) Keep track of start-stop-start aggressions so as to see them as accumulating to ever greater harms.

7) Keep a tally of harms.

8) Avoid leaders who have intransitive preferences if you don’t.

9) Where cost-benefit analysis doesn’t afford natural bright lines as decision points for whether to counter grey zone attacks, use symmetry breaking techniques to choose points in continua of harms as points to regard as egregious enough to merit response.

10) Adopt preferences whose adoption consists in coming to regard these points as serious wrongs, the responses, things desired intrinsically for their own sakes, this to avoid backsliding into further indecision.

11) Analyze conflicts to find out whether they are total or only partial, and zero sum or win-win.

12) In partial conflicts, defuse escalation by bargaining to solutions which recognize that the parties don’t wish each others’ destruction, that each prefers to avoid hot war, and that both have a stake in maximizing and sharing the cooperative surpluses from cooperative solutions to the Prisoners’ Dilemmas that define these conflicts.

13) Come to place intrinsic value on compliance with these solutions, or on the welfare of the other party, provided the other party is doing likewise, in this way becoming a party who can be trusted not to avail of opportunities to exploit other parties in the resulting peace.

1. This chapter began as a paper for a conference entitled Challenging the Grey Zone: The Changing Character of Warfare and the Application of International Law, co-sponsored by the Center for Ethics and the Rule of Law at the University of Pennsylvania Law School, The Middle East Center, likewise at U. Penn, and The Judge Advocate General’s Legal Center and School in Charlottesville, Virginia, U.S.A., April 2019. For useful discussion my thanks to Claire Finkelstein, Sheldon Wein, L.W., and students in my classes on the theory of rational decision. Thanks to Mitt Regan and Aurel Sari generally for their editorial work on the project; and to Sari particularly for some extremely helpful and insightful comments on the penultimate draft. [↑](#footnote-ref-1)
2. Richard Harknett, addressing the conference, Democracy in the Crosshairs: Cyber Interference, Dark Money and Foreign Influence, held by the Center for Ethics and the Rule of Law at the University of Pennsylvania Law School, November, 2018, argued that grey zone conflict is the new norm, and that, unlike traditional conflict, which was fought with money, men, materiel and kinetic warfare, it is more like a perpetual wrestling match, undertaken through cyber hacking, social media trolling, anonymous political influencing, trade wars, naming and shaming, law-fare (using nations’ own laws against them by, for example, suing them), and so on. Each party seeks to influence the other by short-of-war techniques, sometimes called “retorsion”. [↑](#footnote-ref-2)
3. Someone’s stealthing their action so that we can’t attribute responsibility for it is sometimes a good thing. For if we could identify the perpetrator, we’d have to go through the trouble of making a response. In that pay-off structure, the secrecy of their action is almost a courtesy. [↑](#footnote-ref-3)
4. Michael Mazaar, *Mastering the Gray Zone: Understanding a Changing Era of Conflict* (Carlisle Barracks: United States Army War College Strategic Studies Institute, 2015), pp. 126-137. <https://press.armywarcollege.edu/monographs/428?utm_source=press.armywarcollege.edu%2Fmonographs%2F428&utm_medium=PDF&utm_campaign=PDFCoverPages> [↑](#footnote-ref-4)
5. This section develops ideas from Duncan MacIntosh, “Intransitive Preferences, Vagueness, and the Structure of Procrastination” in Chrisoula Andreou and Mark D. White, eds., *The Thief of Time: Philosophical Essays on Procrastination* (Oxford: Oxford University Press, 2010), pp. 68-86. [↑](#footnote-ref-5)
6. This is sometimes called Camel’s Nose warfare: your camel wants in your tent. He sticks his nose in. You don’t go trouble to shoo him away—what matters a camel’s nose in a tent? So he sticks his whole head in…then a foot…. [↑](#footnote-ref-6)
7. Thomas Schelling, *The Strategy of Conflict* (New Haven, CT: Yale University Press, 1960), p. 42. [↑](#footnote-ref-7)
8. Thanks to Michael Watkins for discussion on this point. [↑](#footnote-ref-8)
9. Although it’s also possible that, while the actions in question are not significant attacks, it is significant that one would put up with them—symbolism may be in play, and that can matter. Maybe a foreign leader evidences in his smirk that he thinks our leader is a pawn. That’s not an attack, but still, maybe we should do something about it, or, at the very least, take offense. Or do better next time we pick a leader. [↑](#footnote-ref-9)
10. For example, as I write, in November, 2022, the U.S. Presidential administration of Joe Biden seeks to conceive jousting between the U.S. and China as competition, not conflict. [↑](#footnote-ref-10)
11. The Russians might think that they have a right to protect their sphere of influence by meddling in another country, even though no one has intruded into their country. They experience NATO as standing too close. And Ukraine is used to Russians, and contains many Russian nationals and loyalists, so that there is some pretext for separatist impulses within Ukraine and some rationale for their support from Russia. [↑](#footnote-ref-11)
12. For more on the distinction between rational and irrational changes in one’s values, see Duncan MacIntosh, “Preference-Revision and the Paradoxes of Instrumental Rationality”, *Canadian Journal of Philosophy*, Vol. 22, No. 4 (December 1992): pp. 503-530. [↑](#footnote-ref-12)
13. Moral Psychology is the study of what a mind must be like for its action choices to be subjectable to rational criticism. [↑](#footnote-ref-13)
14. The West uses grey zone warfare, too. As I write, President Joe Biden is trying to “boil the frog” in helping Ukraine against Russia. He is gradually ramping up the amount and quality of aid, being careful not to cross Russian red lines, but also stretching them. Interestingly, he is doing this by providing massive quantities of lethal weapons. But this is only a grey zone attack against Russia since it is legal to give weapons to countries who are merely defending themselves; and since the U.S. is not itself engaged in a hot war. Note that the responses to grey zone attacks I am proposing here can be used by our opponents, too. This will figure, below, when we consider the problem of mutual escalation of conflict. [↑](#footnote-ref-14)
15. See MacIntosh, “Intransitive”. [↑](#footnote-ref-15)
16. Chrisoula Andreou, “Temptation and Deliberation”, *Philosophical Studies*, Vol. 131, No. 3 (2016): pp. 583-606. [↑](#footnote-ref-16)
17. See MacIntosh, “Intransitive” for more, especially the introduction in which my analysis of procrastination as imprudent delay could be reframed as an analysis of irrational failures of response to grey zone conflict. [↑](#footnote-ref-17)
18. E.g., Chrisoula Andreou would say this were she to apply to this case the logic from her, “Environmental Damage and the Puzzle of the Self-Torturer”, *Philosophy and Public Affairs* 34 (2006) (1): pp. 95-108. [↑](#footnote-ref-18)
19. MacIntosh, “Intransitive”. [↑](#footnote-ref-19)
20. For more on such scenarios see MacIntosh, “Intransitive”, especially pp. 70-78. That passage covers Warren S. Quinn, “The Puzzle of the Self-Torturer”, *Philosophical Studies* 59 (1990) (1): pp. 79-90. [↑](#footnote-ref-20)
21. Chrisoula Andreou, “Environmental Damage”, claims there are many real-world choice problems with this structure. E.g., you don’t want to stop polluting, but you don’t want the accumulating effects of polluting. You may now have intransitive preferences: for each moment at which you might stop polluting you can argue that one more polluting act won’t make a difference and it would be a pain to stop, so you don’t stop. But at some point, you’ve polluted so much that you’d love to go back to a point at which you haven’t polluted too much, but you can’t. So you prefer not stopping at each point, but prefer stopping at some point to not stopping. That is you have intransitive preferences.

The same might be true of many Europeans on stopping Russia’s invasion of Ukraine: it’s a pain to fight them over Ukraine, so you don’t. But if you don’t fight them over something, you’ll wind up in Russian-Ruled hell. So while there is no point at which you want to decisively fight the Russians, once they rule, you’ll wish you’d picked a point to fight. You have intransitive preferences about Russian encroachment and the other things in your life. And perhaps the Russians are exploiting this fact about you. Are you irrational? (See MacIntosh, “Intransitive” for more on the structure of such scenarios.) [↑](#footnote-ref-21)
22. We used to have lots of these, e.g., “better dead than red”. There was some advantage to having them: they made us resist all threats of communism, and left us free of grey zone confusions—any whiff of communism sent us into hysterics of retaliation. But there was also a disadvantage. E.g., here, any whiff of communism sent us into hysterics of retaliation! The virtue of the principle is that it makes you resolute against Communism, from which, therefore, the principle may save you. The downside is that it makes you expend resources on trivialities (in this case, on suppressing trivial wafts of Communist thought unlikely to cumulate). [↑](#footnote-ref-22)
23. See Joseph Heath, *Following the Rules: Practical Reasoning and Deontic Constraint* (Oxford: Oxford University Press, 2008). [↑](#footnote-ref-23)
24. See, for example, Jordan Howard Sobel, “Must Constrained Maximizers Be Uncharitable?”, *Dialogue: Canadian Philosophical Review* XXXV, No. 2 (1996): pp. 241-254. [↑](#footnote-ref-24)
25. Brigadier General Ian Langford of Australia taught me this view of conflict and war. [↑](#footnote-ref-25)
26. For more on this sort of scenario, see Duncan MacIntosh, “The Sniper and the Psychopath: A Parable in Defense of the Weapons Industry”, in Daniel Schoeni and Tobias Vestner, eds., *Ethical Dilemmas in the Global Defense Industry* (New York: Oxford University Press, 2023), pp. 47-78 [↑](#footnote-ref-26)
27. This section develops ideas from Duncan MacIntosh, "Buridan and the Circumstances of Justice (On the Implications of the Rational Unsolvability of Certain Co-ordination Problems)", *Pacific Philosophical Quarterly*, Vol. 73, No. 2 (1992): pp. 150-173. [↑](#footnote-ref-27)
28. Sometimes the only solution to be hoped for is one imposed by outside dictators with their own agenda. See MacIntosh, “Buridan”. See also Duncan MacIntosh, “Re-drawing the Boundaries of Sovereignty: Permissible and Obligatory Interventions in the Affairs of Sovereign Nations” (ms. Dalhousie University, 2017). [↑](#footnote-ref-28)
29. This section develops ideas from Frederic Schick, *Having Reasons: An Essay On Rationality and Sociality* (Princeton, New Jersey: Princeton University Press, 1984). [↑](#footnote-ref-29)
30. See David Gauthier, *Morals By Agreement* (Oxford: Oxford University Press, 1986). [↑](#footnote-ref-30)
31. Recall Section 2.1 and Structure 1 scenarios, from above, in which there is merely an illusory need to reply to attacks. [↑](#footnote-ref-31)
32. Thanks to Jack Whitmer for this worry. [↑](#footnote-ref-32)
33. Again, my thanks to Whitmer. [↑](#footnote-ref-33)
34. Thanks to Jana Robinson for this point. [↑](#footnote-ref-34)
35. My thanks to Aurel Sari for suggesting that this aversion to hot war finds a place in the analysis of grey zone conflict I’ve been offering. [↑](#footnote-ref-35)
36. Gauthier, *Morals*. [↑](#footnote-ref-36)
37. Gauthier, *Morals*. [↑](#footnote-ref-37)
38. This develops ideas in Duncan MacIntosh, “Assuring, Threatening, a Fully Maximizing Theory of Practical Rationality, and the Practical Duties of Agents”, *Ethics*, Vol 123, No. 4 July (2013): pp. 625-656; that article proposes a variant on ideas originally proposed by Gauthier, *Morals*, above. [↑](#footnote-ref-38)
39. I here apply a distinction made in David Gauthier, “Assure and Threaten”, *Ethics* Vol. 104, No. 4 (July, 1994): pp. 690-721. [↑](#footnote-ref-39)