

Implementing the Law by Impartial Agents: An Exercise in Tort Law and International Law

*Eyal Benvenisti** & *Ariel Porat***

Lawmakers regularly delegate authority to agents. Such delegation is accompanied by mechanisms that attempt to ensure that the agents adhere to the will of the lawmakers. But these mechanisms are often ineffective or inefficient. Moreover, at times the very imposition of constraints distorts the agents' incentives and impels them to adopt skewed policies. We suggest that it is possible to reduce such wasteful enforcement costs by delegating authority to certain types of agents who will pursue the lawmaker's policies without constraints imposed by the lawmaker. In this Article we focus on agents who are impartial — but not indifferent! — and skillful enough to identify the proper course of action. The Article encompasses two main arguments. The normative argument is that when skillful and impartial agents can be identified, it makes sense to delegate to them decision-making powers with only limited constraints. Moreover, in such instances it may be more cost-effective to provide agents with incentives (or design agents, like administrative agencies) to act impartially rather than develop

* Professor of Law and Director of Cegla Center, the Buchmann Faculty of Law, Tel Aviv University.

** Dean and Professor of Law, Buchmann Faculty of Law, Tel Aviv University. For helpful comments we wish to thank Alon Klement, Hillel Somer, Steve Sugarman, the participants at the law and economics workshop at Bar Ilan University and Tel Aviv University, the participants at the faculty workshop at the Interdisciplinary Center in Herzliya, and the participants at the joint conference of the University of Pennsylvania Law School and the Cegla Center at the Buchmann Faculty of Law, Tel Aviv University, on "The Role and Limits of Legal Regulation of Conflicts of Interests," held at Pennsylvania University Law School in October 2003. Special thanks go to Leo Katz, our commentator at the conference, and to Dana Rothman-Meshulam, for superb editorial and substantive suggestions.

enforcement mechanisms to impose impartiality. The positive argument is that the law — sometimes explicitly and sometimes implicitly — is compatible with our normative argument in various fields. Although our argument is general in scope and applies to many areas of law, in this Article we begin by focusing on tort law and on international law. We distinguish five categories of cases where the law relies on agents and detect in some of them impartial and skillful agents. This analysis demonstrates that, in some of these categories, the law in fact relies on agents with only minimal constraints and that, in other situations, the constraints imposed by the law are counterproductive. Our normative argument thus serves an explanatory role in understanding many legal doctrines and principles, but, at the same time, offers a critical view on other doctrines and principles, which are not compatible with that argument.

INTRODUCTION

When lawmakers seek specific outcomes, they can impose well-defined rules of behavior that the subjects of the law must respect. But often lawmakers do not have clear outcomes in mind when making laws or are uncertain as to which rules of behavior are required to realize certain outcomes. In other cases, lawmakers encounter tremendous difficulties in enforcing the law. In all such instances, they delegate authority to individuals, public authorities, regulators, and states as their agents, to make decisions that impact people's rights and duties. This delegation of authority raises what the law and economics literature terms the "agency problem": agents who seek to promote their own interests at the expense of the public interest. To reduce agency costs, the delegation of authority to such agents is accompanied by mechanisms that attempt to ensure that, rather than pursuing their own, biased goals, the agents make decisions that reflect what the lawmakers' choice would have been had the lawmakers been involved in the particular decision. These mechanisms include rules setting the agents' composition and decision-making procedures (procedural rules) and rules prescribing the outcomes the agents should reach or how they should conduct themselves prior to making their decisions (substantive rules). Lawmakers can either articulate these rules or delegate the authority to do so to yet another agent, such as a court or administrative agency.

Using these various mechanisms is costly. Lawmakers may err and assign decision-making power to agents whose tendency to betray the lawmakers' trust either cannot be effectively controlled or can be controlled only at a very high price. Conversely, lawmakers can place unnecessary (and costly)

constraints on agents who have no incentive to make biased decisions. Moreover, at times the very imposition of constraints distorts an agent's incentives and leads her to make biased decisions.

It is necessary to uncover the incentives of relevant agents in order to identify those agents who do not need to be constrained, those who would perform well under either procedural constraints or substantive ones, and those whose bias must be constrained both procedurally and substantively. Our thesis is that by distinguishing among the different types of agents, it is possible to eliminate or significantly reduce the costs of delegating authority to some of the agents.

Three basic considerations can be identified for differentiating among types of agents for the purpose of delegating decision-making power and, at the same time, designing a mechanism to constrain the use of the delegated power: their level of impartiality ("the impartiality consideration"); the costs of supervising and sanctioning them if necessary in order to neutralize their tendency to prefer their own interests over those of others ("the costs of enforcement consideration"); and their skillfulness in dealing with the subject matter ("the skillfulness consideration").

Much of the law focuses on the supervision and sanctioning of agents. Focusing on enforcement, however, can be both wasteful and counterproductive, as agents react by adopting defensive behavior. We therefore set out to identify conditions where enforcement is not required. We argue that enforcement is not necessary where agents are impartial and skillful. The Article raises two main arguments, one normative, the other descriptive. *The normative argument* is that the delegation of decision-making powers to skillful and impartial — but not indifferent — agents does not create an agency problem. It therefore makes sense to delegate these powers to such agents, for no enforcement mechanism is required and, more importantly, at times such a mechanism is likely to even distort an impartial agent's considerations and render her biased. Moreover, in general it may be more cost-effective to provide agents with incentives (or design agents, like administrative agencies) to act impartially rather than develop enforcement mechanisms to impose impartiality.

Following the presentation of the normative argument, we identify five categories of cases in which agents are required to balance different interests before making their decision and acting upon it. What distinguishes each category is the type of balance of interests the agent is required to conduct. In the first category, the agent balances her own interests against those of another party; in the second category, the agent balances among the different interests of another party alone; in the third category, the interests of two (or more) different parties are balanced against each other; in the fourth

category, the interests of another party are balanced against social interests; and in the fifth category, the agent balances among her and another party's converging interests. We show that this classification among the different types of balances of interests informs the determination as to whether constraints should be placed on agents operating in different spheres of the law.¹

The positive argument presented is that the law — sometimes explicitly and sometimes implicitly — corresponds with our normative argument in various legal contexts. Our normative argument thus serves an explanatory role in understanding many legal doctrines and principles in those contexts. The same argument, at the same time, offers a critical perspective of other doctrines and principles, ones that are not compatible with that argument.

Part I of the Article elaborates on the impartiality consideration. Part II applies this consideration in the framework of tort law and demonstrates its role in shaping negligence law principles. Part III moves to the sphere of international law and examines instances where the law defers, or could defer, to the unilateral action of one state when that state is deemed, or can be deemed, to be in the position of impartial agent. In Part IV, we summarize and draw some general conclusions.

I. THE IMPARTIALITY CONSIDERATION

A. Outcome Interest, Accuracy Interest, and Reciprocity Interest

In the following section we identify three interests that motivate decision-makers. "Outcome interest" is the interest to reach an outcome that benefits the decision-maker herself. "Accuracy interest" is the interest to pursue the social goals — whatever they may be — that the lawmaker would have pursued were he or she in the agent's shoes. "Reciprocity interest" is the interest to avoid possible retaliation by other agents if the agent externalizes costs to them. An agent may be impartial in the absence of a personal interest in the decision. But even if an outcome interest exists, it may be outweighed by the presence of either an accuracy interest or a reciprocity interest. Of course, the ideal impartial agent will be the agent who has

¹ For further discussion of these same five categories of cases in the context of negligence law, see Ariel Porat, *The Many Faces of Negligence*, 4 *Theoretical Inquiries L.* 105 (2003).

no outcome interest but has strong accuracy and reciprocity interests. But impartiality can be obtained even if not all interests are so aligned.

When an agent has no personal interest in the outcome of her decision (no outcome interest), she is an impartial agent. Impartiality does not require, however, that the agent be indifferent to and have no interest in reaching the "right" decision. On the contrary, the impartiality of the agent who has no outcome interest is further enhanced by her strong accuracy interest. Moreover, the impartial agent must be skillful in the sense that she must understand what an accurate decision is and be capable of identifying it (although she may err and reach an inaccurate decision in specific cases).

It is also possible for an agent with an outcome interest but no accuracy interest to be an impartial agent. Such could be the case when an agent is vulnerable to possible retaliation by other agents if she externalizes costs to them or other agents. The threat of reciprocal action — which generates the agent's reciprocity interest — forces the agent to take into account not only her outcome interest but also the costs she would impose on others if she were to pursue her outcome interest, thus neutralizing her partiality to a great extent. Occasionally, a strong reciprocity interest could even turn an agent with an outcome interest into a completely impartial agent. Both the impartial agent motivated by an accuracy interest and the impartial agent motivated by a reciprocity interest might act in accordance with the social goals of the lawmaker — the former seeking accuracy and the latter fearing retaliation. Recall that both types of agents must be skillful in order for impartiality to obtain. But the skills required of each one differ: the accuracy-interested agent must have the necessary skills to identify the lawmaker's goals and act upon them; the reciprocity-interested agent must have the necessary skills to anticipate retaliation and identify the strategy to avoid it.

Note that there is a potential third type of impartial agent. The conditions for this type to emerge are when, for some reason, it is certain that the agent's outcome interest completely converges with the collective social goals the lawmaker would pursue and thus expects the agent to pursue. But since this convergence of interests is seldom verifiable, impartiality — stemming either from an accuracy interest or reciprocity interest — appears to be a significant and central characteristic of the impartial agent the lawmaker can trust.

A typical example of an outcome-impartial but not accuracy-impartial agent is a judge. Generally a judge has no interest in the outcome of her decision, but does have a strong interest in the accuracy of that decision. The judicial system encompasses mechanisms (publicly accessible proceedings, a duty to provide and publish reasoned decisions, judicial review, and a

judicial promotion process) that enhance the accuracy interest of judges. So much so, that we usually assume that the accuracy interest of judges outweighs their outcome interest even when the latter exists. As a result we have no objection to a judge deciding, say, in favor of a bank in which she holds an account (although critical legal scholarship and, frequently, popular sentiment would challenge this claim, since they dispute the assumption that judges have a stronger accuracy interest than outcome interest).

An agent may be indifferent to both the outcome and accuracy of her decision, and as noted, indifference is not tantamount to impartiality in the context of our discussion in this Article. To transform the indifferent agent into an impartial one, it is necessary to create either an accuracy interest or a reciprocity interest for her, for example, by making the decision-making process transparent, by requiring written reasoning for the decision (if the agent is, for example, a government employee), by threatening sanctions or offering subsidies for the agent's functioning, or by providing market-based incentives for the agent to perform accurately based on her reputation.

An agent may have an outcome interest and a conflicting accuracy interest and/or reciprocity interest. The existence of such a conflict, however, does not necessarily indicate that the agent is either partial or impartial. Rather, partiality is a function of which of the potentially conflicting interests is of greater relative strength. Only if the agent's outcome interest is greater than her accuracy interest (or reciprocity interest) does a risk of partiality arise, and the latter must be increased (or the former reduced or changed) in order to secure impartiality in a clash between the two.

It is costly for the legal system to invest in the various mechanisms that can ensure that an agent's outcome interest is weaker than her potentially conflicting accuracy interest or reciprocity interest. It is also costly to identify impartial agents — agents with no outcome interest but having an accuracy (or reciprocity) interest or agents whose accuracy (or reciprocity) interest is either identical in strength to or stronger than their outcome interest. But this does not mean that such an exercise should not be pursued. In fact, sometimes it is less costly to invest in mechanisms for identifying impartial agents than assume that all agents are partial and invest in the enforcement mechanisms of supervision and sanctioning.

Moreover, and more importantly, sometimes the enforcement costs are counterproductive, and they distort the balance between an agent's outcome interest and accuracy interest. For example, an agent who has no outcome interest and whose accuracy interest is relatively weak will function as an impartial agent without the interference of enforcement mechanisms. Yet the presence of enforcement mechanisms may modify her outcome interest and place it in a state of conflict with her accuracy interest. In such cases,

obviously the enforcement mechanisms are counterproductive. Throughout the Article we give several examples of such situations.

Thus, it follows that if the costs of identifying impartial agents are lower than the costs of enforcement, it makes sense to inquire into the identity of "natural" impartial agents, namely, those for whom enforcement mechanisms would be counterproductive.

B. Conflicting Outcome Interests

All agents often have conflicting outcome interests. Indeed, everyone, in his or her capacity as a decision-maker, is quite familiar with the situation of a conflict of outcome interests. Professors have to decide whether to go to a movie or prepare for class; a doctor has to decide between treating all her patients, including the most gravely ill, in order to uphold her Oath, and treating only the most promising patients, in order to ensure a high success rate of treatment. These conflicting outcome interests pull the decision-maker in opposite directions.

A similar conflict may arise between an agent's different impersonal (ideological, political, religious) outcome interests. A judge, for example, may have no personal outcome interest in a case she decides before her but still have an ideological outcome interest, and that outcome is incongruent with the prescribed legal outcome.

Thus, when we seek to identify which agents are impartial and for that purpose focus on the outcome-impartiality of certain agents, we have to take into consideration that their outcome-impartiality can, in itself, be the outcome of an internal conflict. For this reason, when we examine an agent's impartiality, we should look at the balance of her conflicting outcome interests: In the final analysis, which of these interests bears greater weight in the agent's decision-making? The mere existence of conflicting outcome interests often raises the concern that we cannot assess the relative strength of each interest. Perhaps the surgeon in fact has no interest in curing all her patients, and we have no interest in coming prepared to class. The lack of information prevents us from treating the agent as outcome-impartial and requires reliance on strengthening her accuracy or reciprocity interests.

C. Personal and Institutional Impartiality

When an agent is a body composed of a group of agents (for example, a panel of nine judges or a nine-member administrative board), the overall impartiality can result from the individual impartiality of each and every agent within the group or it can result from the conflicting outcome

interests of the individual agents, which cancel each other out to produce outcome impartiality. The first type of impartiality may be termed personal impartiality, and the second — institutional impartiality.

It is not always necessary to ensure the personal impartiality of all the agents within the group for an impartial decision to be secured. Institutional impartiality (when, say, 4 judges on the bench have an interest in outcome *A*, 4 others have an interest in outcome *B*, and the remaining judge is indifferent to either outcome) may be sufficient to ensure impartiality. At times, such institutional impartiality may be preferable to personal impartiality, for example, when the inclusion of individual agents with interests in particular options can inform the group of those options. While there may be legal disputes in which there will be no apparent benefit to including partial agents on the decision-making body, it would be beneficial to have partial agents sitting on, for example, specialized courts (such as representatives of employers and representatives of employees on a labor court or state representatives in international arbitrations) to inform the outcome-impartial agents about the various relevant outcomes. When agents are not required to rely on their expertise to reach their decisions — usually in the context of making political decisions about wealth redistribution — only institutional impartiality is important. Accordingly, legislative bodies are impartial because they include a proportion of partial agents that represents (roughly) the different choices of the population.

II. TORT LAW

When one person (an injurer) exposes another person (the victim) to the risk of injury, the law has two different solutions at its disposal to restrict the injurer's behavior and prevent her from inappropriately preferring her own interests over the interests of the victim. One solution is to use regulations backed by sanctions, such as fines, imprisonment, or tort liability. The other solution is by means of internalization, that is, the injurer is allowed to choose the level of risk that she creates for others, but assumes full liability for the harm she inflicts.² Thus, under the latter solution, the law turns the injurer into its agent who is expected to balance her interests against those of others. Under the former solution, the regulator conducts this balance.

2 Alternatively, the law could impose liability for risks, at the amount of the expected damages of the injurer's wrongful behavior. See Ariel Porat & Alex Stein, *Tort Liability under Uncertainty* 101-29 (2001).

Internalization can be realized in two ways: either by ascribing liability to the injurer for *any* harm she caused (strict or no-fault liability) or by limiting liability to blameworthy behavior only (negligence or fault liability). In both of these alternative ways, the law authorizes the injurer, as its agent, to create risks, but, at the same time, neutralizes her natural tendency to advance her own interests at the expense of others. The threat of liability forces the injurer to take the interests of other parties into consideration, to treat them as though they were her own rather than ignoring them whenever her own interests are at stake. In short, tort law is designed mainly to mitigate conflicts of interest. Its goal is to weaken the injurer's outcome interest in the balance of interests she is required to conduct and to make her an impartial agent.³

A closer look at tort cases reveals that the conflict of interest between the injurer and her victim is, in many cases, either minimal or non-existent. In those cases, when the injurer has either no outcome interest or his outcome interest is neutralized by his reciprocity interest, tort liability does not fulfill its goal, which is to mitigate conflicts of interest, and therefore is often redundant. Moreover, imposing liability on injurers in such cases occasionally generates an outcome interest for injurers and places them in a conflict of interest with their victims that could have been avoided had liability not been imposed. Thus, an injurer who is a relatively impartial agent could become a partial agent merely because she is exposed to tort liability.

In the following sections, we will examine five categories of cases, which represent five different instances of the balance of interests that the injurer is required to conduct before taking action. We will show that the necessity of imposing liability varies from one instance to another, in accordance with the severity of the conflict of interest that the injurer is confronted with before she decides on her action. We will proceed by arguing that imposing tort liability on an injurer is, at times, not only unnecessary but also detrimental to the goal of making her an impartial agent. At other times, reduced liability could be an optimal compromise between full liability and no-liability. Finally, we argue that prevailing negligence law is often consistent with our normative arguments since courts are less willing to impose tort liability on impartial injurers.

3 Tort law also offers compensation for victims, and corrective justice mandates such compensation if the injury is a result of wrongful behavior. *See* Ernest Weinrib, *The Idea of Private Law* (1995). However, in the present article, our focus is on incentives, not compensation or corrective justice.

Category 1: The Agent Decides in a Matter Concerning Herself and Another Party, in which Their Interests Conflict

Example 1: The Driver. John drives his car at a speed of ninety kilometers per hour. He hits Tony and causes damage to him. Had John driven at a lower speed, the damage would have been prevented.

Example 1 represents the paradigmatic case that tort law is based upon. Under a negligence regime, John would be found liable if he had given inappropriate weight to the safety interests of other people using the road when weighing them against his own interests in arriving at his destination in less time. If the Hand Formula is applied to Example 1, John's liability depends on whether his costs of precaution, namely, the time wasted by driving slowly enough to prevent the accident, are lower than the decrease in the expected damages that would have resulted from taking the precaution of driving slowly.

By imposing liability on John if he failed to balance properly the various conflicting interests, the law neutralizes John's tendency to prefer his own interest in arriving at his destination in less time (his outcome interest) over Tony's bodily and propriety interests. If the possibility of liability were not hanging over John's head, he might have an incentive to create too-high risks to other parties and thus become a partial agent. Therefore, the threat of liability is necessary to make him an impartial agent.

Notice that a strict liability rule would have the same effect *in principle*. In fact, it could work better than a negligence rule when the injurer's behavior is not verifiable but the outcome of her behavior is. Moreover, a strict liability rule has a more efficient effect on the injurer's activity level and is probably less costly to administer. The differences between a strict liability rule and a negligence rule are, however, beyond the scope of this Article. For the particular purposes of the Article, it is sufficient to conclude that both negligence and strict liability regimes are directed at mitigating or even abolishing the injurer's initial partiality in balancing interests before taking action.⁴

4 In some cases, the injurer would take the victim's interests into account even if no liability were imposed, because he would fear retaliation by the victim if he were to ignore the latter's interests. Thus, if the injurer can build his house in a non-secure way that may endanger his neighbor and the neighbor can do the same, the injurer is likely to make sure his house is safe so as to avoid the risk of his neighbor's retaliation. The victim's interests are taken into account in this case due to the

Category 2: The Agent Decides in a Matter Concerning Only Another Party

Example 2: The Doctor and the Patient. Jim, who is a doctor, must decide how to treat his patient, Stephen. He has a choice between two possible courses of treatment. Jim has no self-interest in terms of which particular course of treatment to choose. Stephen suffers injury from the treatment chosen by Jim. Had Jim chosen the other course of treatment, the damage would have been prevented.

Under a negligence regime, the court would examine the chances and risks presented by each one of the alternative courses of treatment that might have been considered by a reasonable doctor in Jim's position at the point at which he decided on the course of treatment. The balancing of Stephen's different interests is, therefore, what is under discussion here. If Jim is found negligent, it will not be because he preferred his own interests to Stephen's interests. Indeed, the fact that Jim strove to act in Stephen's best interest in no way contradicts a possible determination that Jim was negligent.

Under the Hand Formula, the absence of any self-interest on Jim's part in the choice between the two courses of treatment means that the costs of precaution Jim could have taken were zero.⁵ Consequently, finding that Jim was negligent in his choice of treatment would mean that he unnecessarily exposed Stephen to risks that could have been avoided at zero cost to Jim.

But if precautions were costless for Jim, from the outset liability was not necessary to encourage him to take precautions. Jim happened to be a completely impartial agent who had no self-interest whatsoever in the outcome of his decision. Non-legal incentives would have been more than sufficient to convince him to act in his patient's best interest, alongside his conscience and likely sense of compassion. If Jim does not face a risk of liability, he will have no outcome interest, on the one hand, and a strong accuracy interest, on the other, which should assure that he will apply impartial discretion in balancing Stephen's interests.

Our discussion thus far raises two questions critical to our argument for no-liability in cases represented by Example 2: Firstly, if precautions were

injurer's reciprocity interest, which forces him to consider the victim's interests as though they were his own interests.

5 Alternatively, one can define the costs of precaution to avoid the risks entailed by one course of treatment as the net risk (risk minus benefit) entailed by the other course of treatment.

costless, why did Jim ultimately fail to take them? Perhaps liability should, nonetheless, be imposed in such cases and something is missing in our analysis. Secondly, even if liability is not necessary in order to provide the injurer with the incentives to take precautions, it still is important because it offers compensation for victims. Moreover, since it is administratively costly to evaluate and verify the costs of precaution, full liability is still the best practical alternative for enhancing the injurer's incentives. Given these two advantages of tort liability, why should Jim be exempt from liability and why should Stephen bear the damage? In what follows, we discuss each of these two questions in turn.

If Jim was an impartial agent, why did he fail to take costless precautions? One possible explanation is that the precautions were not really costless. Jim could have expended more time and effort in order to make the right decision, for example, by consulting other doctors and checking the professional literature more carefully before taking action. Tort liability in this case may, therefore, not be redundant. It turns out that Jim was not really an impartial agent. His self-interest was in saving his time and effort (and he therefore had an outcome interest), and only the threat of liability could neutralize this partiality. Moreover, finding Jim and doctors in similar situations liable would improve the incentives of hospitals and other suppliers of medical care who are vicariously liable for the negligence of their medical staff to hire doctors with better professional qualifications and to make greater efforts to reduce the number of cases where, in treating their patients, doctors fail to take the necessary precautionary measures to prevent potential damage.

However, this explanation cannot justify imposing full liability on Jim. Even if Jim has a self-interest in saving time and effort, the conflict of interest he faces is much less severe than that faced by John in Example 1. Jim's accuracy interest, which derives from non-legal incentives, could suffice to restore his impartiality. But if not, then reduced liability, which is more limited in scope than full liability, would be quite sufficient for this purpose.

In principle, even in Example 1, and in all similar accident cases, liability could have been imposed for the amount of the cost of the precautions that would have prevented the harm instead of for the extent of the harm itself. Tort law imposes liability for the extent of harm and not for the amount of the required precautions (or any other amount) for a number of reasons, but the scope of this Article does not allow for a detailed discussion thereof. Briefly, this is probably so because of the administrative costs necessary to verify the costs of the required precautions (and, perhaps, also in order to pursue corrective justice and to facilitate compensation for victims). But

when, as in Example 2, those costs are close to zero — which can easily be verified — reduced liability can effectively serve the goal of mitigating the conflict of interest between the injurer and her victim.

A second possible explanation for Jim's failure to take costless precautions in order to reduce Stephen's risks is a possible lapse in attention. Tort law considers lapses in attention negligent behavior, even though any reasonable person's attention is known to lapse from time to time. If this explanation is the correct one, the question still remains as to why full liability — as opposed to reduced liability — should be imposed on the injurer.

We have thus far argued that full tort liability is not necessary for encouraging the injurer to take precautions in cases such as Example 2. But even if it is, indeed, unnecessary to enhance the injurer's incentives, why not impose full liability either in order to compensate victims or to avoid administrative costs that other alternatives might entail? This is the second question that must be addressed.

The answer to this question is that imposing full liability in Example 2 and similar cases does not come at zero cost. The very real concern regarding full liability is the negative effect it may have on the behavior of potential injurers. Although this apprehension is not unique to Example 2 or similar cases, it appears that it is particularly compelling in cases where the injurer does not have a strong outcome interest with regard to which course of behavior to choose. If Jim is aware of the risk of legal action, he is likely to opt for a defensive medicine strategy. He is likely to prefer his own interests over the interests of the patient and to choose the course of treatment that reduces the likelihood of his exposure to legal action rather than the best course of treatment for the patient. This would not occur if the two courses of treatment were to entail the same risk of liability for Jim. But it is often the case that one course of treatment could well result in damages to the patient that are more easily attributable to the doctor's decision (such as immediate bodily injury), while another course of treatment could lead to damages that, while no less severe, might well be less ascertainable in a future legal action (such as latent weakening of the patient's immune system). Thus, the asymmetry of the liability threat Jim faces causes him to externalize the costs of one course of treatment to the patient and, at the same time, internalize the costs of the other course of treatment. Why should Jim choose the course of treatment that exposes him to the risk of being sued — even if he believes it to be of greater benefit to the patient — and not the course of treatment that entails a lower risk of legal action, when he derives no personal benefit from either of the treatments? As stated earlier, Jim is sometimes motivated in his choices by his accuracy interest, created by non-legal incentives. However, these, albeit important, incentives

are at times insufficient to dispel the concern described above vis-à-vis the negative effect of imposing full liability on potential injurers. Thus, an injurer's outcome interest often overrides her accuracy interest.

The effect of over-deterrence created by the threat of full liability, as described above, constitutes the main objection to imposing such liability on Jim. If Jim *cannot* be held liable, he will be an impartial (or almost impartial) agent who does his utmost for the good of his patients; if, on the other hand, Jim *can* be held liable, he will be a partial agent who can easily be tempted to betray his patients and to do his utmost to escape legal liability.

The second category of cases does not only comprise malpractice cases. There are many other instances where the injurer is required to balance his victim's interests and where little or no self-interest is involved. In such cases, there is a clear risk of turning an impartial agent into a partial one as a result of the threat of the imposition of liability. Among cases falling into the second category are those instances in which one person acts for the benefit of another person, both when there is a preexisting legal duty requiring her to do so as well as when no such duty exists, and when she has no clear outcome interest with regard to the course of action she has chosen. An example of such a case is when someone rescues another person without any preexisting duty to do so and questions arise with regard to her negligence in the actual execution of the rescue.⁶ Cases in which a parent has negligently injured her child also fall under the second category of cases.⁷ Another example is the case of a teacher who is sued for not properly balancing between her pupil's interest in physical safety and the pupil's interest in freedom of action, which ultimately causes the pupil to suffer bodily injury.

6 In C.A. 380/74, *Ackers v. Kalman*, 30(1) P.D. 611, the plaintiff was crushed by a car that rolled down a hill by fault of the defendant, when the plaintiff was trying to save the defendant from being injured by the car. The comparative negligence of the plaintiff had been determined by the trial court to be at a level of 50%. The Supreme Court attributed the entirety of the liability to the defendant. This ruling was based on the fact that the plaintiff's reaction had been instinctive, natural, and reasonable. For a similar American case, where it was noted that the jurors must determine whether there had been comparative fault on the part of the plaintiff, see *Carney v. Buyea*, 65 N.Y.S.2d 902 (N.Y. 1946).

7 There is good reason to absolve such a parent from liability if she acted in good faith and in the victim's interests, and most legal systems have chosen this solution to the problem. See W. Page Keeton et al., *Prosser and Keeton on the Law of Torts* 904-07 (5th ed. 1984).

Category 3: The Agent Decides in a Matter Involving Only Other Parties with Conflicting Interests

Sometimes an injurer's negligence is manifested by the fact that she did not properly weigh the victim's interests against the interests of a third party. Here, too, the injurer's self interest is not of much (if any) significance among the set of factors that influence her behavior. She can be considered an impartial agent with no outcome interest. Nonetheless, as we will see below, the reasons for imposing liability on such an injurer are stronger in this third category than in the second category, and the risk that the imposition of liability will distort the injurer's incentives and make her a partial agent is much lower in this category than in the second category.

Example 3: The Wounded Person. David lies by the road wounded. Jane is a driver who picks David up in her car and rushes him to the hospital. Due to David's critical condition and the urgent need to get him to the hospital as quickly as possible, Jane drives very quickly. On the way to the hospital, she accidentally hits George, a pedestrian, and injures him.

Under a negligence regime, Jane is considered negligent if she failed to properly balance David's interest in reaching the hospital as quickly as possible against George's interest in his own physical safety. Absent the threat of liability, Jane would balance those interests impartially, since we assume that she has no self-interest in the decision she makes. The threat of liability would cause no change to Jane's incentives. The reason for this is that the threat of liability in Example 3 is symmetrical in the sense that if Jane negligently prefers David's interests over George's and, in the process of so doing, hits George, she will be liable towards George, and if she negligently prefers George's interests over David's and arrives at the hospital too late for David to be saved, she will be liable towards David. Thus, liability makes Jane internalize all the costs and benefits of her decision regarding at what speed to drive her car. She therefore remains an impartial agent. It would appear that the main difference between Example 3 and Example 2 pertains to the symmetry of the liability threat faced by the injurer. In the above discussion of Example 2, we assumed that the threat of liability to Jim was asymmetrical (as in many malpractice cases) — which caused him to externalize the costs of one of the courses of treatment to his patient — whereas in Example 3, we assume a symmetrical threat.

Two main conclusions can be drawn from our discussion thus far: Firstly, liability in cases that belong to Category 3 is not as crucial to the injurer's incentives as in cases belonging to Category 1. Secondly, liability in cases

falling under Category 3 is not as detrimental to the injurer's incentives as in cases falling under Category 2. However, these two conclusions require qualification. Liability in Example 3 is *crucial* if there is a tangible chance that Jane might prefer David's interests over George's because David happens to be her husband, friend, or neighbor. In this case, Jane could no longer be considered an impartial agent, and liability would be necessary to neutralize her initial tendency to prefer David over George.⁸ Liability in Example 3 is *detrimental* if Jane is an impartial agent who does not have any *a priori* preference for either David or George, but the risk of liability she is exposed to if she inflicts harm is asymmetrical. This will be the case if there is a significant chance that one of the potential victims (David or George) does not succeed in recovering damages in court from Jane, while the other is expected to succeed. If Jane is aware of the asymmetrical threat of liability emanating from the difference in the potential victims' chances in future legal action against her, she might be tempted to prefer the interests of the potential victim who has a greater chance of succeeding in court against her. Thus, exposing Jane to a liability threat may turn her into a partial agent with a clear outcome interest. Liability therefore appears to be detrimental to her incentives.

In sum, liability in Category 3 is, in general, less crucial than in Category 1, but more crucial than in Category 2. However, it is important to be aware of the context. Sometimes the third category converges with the first. This occurs in cases in which the injurer equates the interests of one of her potential victims with her own interests. At other times, Category 3 resembles Category 2. Such is the cases when each victim presents a different risk of liability to the injurer. Notice that in Category 3, even when the injurer is mainly impartial, she may occasionally not be completely impartial. As in cases falling under Category 2, in third-category cases, too, balancing the relevant interests can entail a great deal of time and effort. As noted earlier, the existence of non-legal incentives, which create an accuracy interest for the injurer, or, alternatively, the imposition of reduced (as opposed to full) liability could be quite sufficient to neutralize the slight impartiality of the injurer.

The third category of cases is very broad, including all those cases in which the injurer has caused damage to someone and where the prevention of that damage would have placed a third party at risk. Examples are the

⁸ However, if Jane is an ambulance driver, the risk of her knowing David is minimal and the high likelihood that she has a strong accuracy interest would suggest imposing minimal or no liability on her.

case of a judge or an arbitrator deciding on a dispute between a plaintiff and defendant;⁹ a public authority exercising its discretion and preferring the interests of one person over those of another;¹⁰ professionals like lawyers, doctors, or therapists who, in acting on behalf of their clients or patients, cause damage to a third party;¹¹ and manufacturers or suppliers of services who should have considered the safety of their customers as well as the safety of a third party.¹² In many of these cases, the injurer's interest either closely

9 See Keeton et al., *supra* note 7, at 1056-59.

10 See *Hill v. Chief Constable*, [1988] 2 All E.R. 238 (H.L.) (the police were not held liable for omitting to hold a person in custody, who, after he was released, murdered the plaintiffs' relative — the House of Lords ruled that the police owe no duty of care in the circumstances under discussion); *Home Office v. Dorest Yacht Co.*, [1970] 2 All E.R. 294 (H.L.) (the House of Lords ruled that a public authority that operates a rehabilitation camp owed a duty of care to people who were injured by inmates who escaped from custody, since these inmates were under the supervision and control of the authority). For an American case in which liability was imposed on the police, see *De Long v. County of Erie*, 457 N.E.2d 717, 718-23 (N.Y. 1983). (The police were held liable for not arriving in time at the scene of the crime and not preventing a murder. The woman who was murdered had tried to call the police, but the police were delayed in arriving due to an error that was made in inputting the murder victim's address. The victim, it emerged, had made use of an emergency number that the police had publicized, when the reasonable impression was that using this number would ensure the speedy arrival of the police.)

11 *Tarasof v. Regents of Univ. of Cal.*, 551 P.2d 334 (Cal. 1976). (A therapist who did not warn his patient's ex-girlfriend of the patient's intention to murder her was found liable for the death of the ex-girlfriend. The court ruled that the therapist had been obligated to warn her, even if this would have constituted a breach of medical confidentiality.)

12 A decision handed down by an American court is a good example of this kind of case. In the American case, a telephone company subscriber sued the telephone company for the nervous shock he suffered due to a sudden and loud noise from the phone cables that interrupted his phone conversation. It appeared that had the telephone company taken certain precautions to decrease the risk of this occurrence to its subscribers, the risk of electrocution to people on the street would have increased. The court dismissed the suit, emphasizing the importance of protecting the lives of the people on the street even if at the expense of protecting subscribers. *Cooley v. Pub. Serv. Co.*, 10 A.2d 673 (N.H. 1940). The court reasoned as follows:

The defendant's duty cannot, in the circumstances, be to both. If that were so, performance of one duty would mean non-performing to the other. If it be negligent to save the life of the highway traveler at the expense of bodily injury resulting from fright and neurosis of a telephone subscriber, it must be equally negligent to avoid the fright at the risk of another's life. The law could tolerate no such theory of "be liable if you do and liable if you don't."

See also *Lucchese v. San Francisco-Sacramento R. Co.*, 289 P. 188 (Cal. 1930). In *Lucchese*, a train crashed into a truck, and it was claimed that the train engineer had

approximates or is even identical with that of one of the potential victims (and therefore very close to Category 1 cases), while in other cases, the injurer is impartial or almost impartial in balancing among the conflicting interests. In some of these cases, liability is not imposed (e.g., due to the immunity granted to judges and arbitrators¹³), while in other cases there is liability, but only very rarely.

Category 4: The Agent Decides in a Matter Involving Another Party and a Social Interest

Example 4: The Excursion. Dan is a tour guide. He leads a group of people on a hike through the Judean Desert and, at a certain stage, has to choose one of two paths: one path leads through an area in which damage is likely to be caused to the landscape and nature if many people hike through it; the second path goes through a different area and does not pose such a threat to the environment. However, the latter path is more dangerous to the hikers. Dan chooses the second path. Jacob, one of the hikers, falls from a cliff and suffers physical injury while walking along the path.

The determination as to whether Dan was negligent depends on whether he properly balanced the social value of preserving the landscape and nature and the risk that this created for the hikers. The mere fact that the path he chose was riskier than the other path is not, in itself, indication of negligence on his part. However, even if Dan was negligent, holding him liable could be detrimental to his incentives. If there is no risk that Dan might be held liable, he is an impartial agent who has no outcome interest. In such a case, he would likely properly balance the social interest in preserving the landscape and the safety of the hikers. If there is a chance that he might be at risk of being found liable, Dan's fear of liability will hamper his motivation to give proper weight to the preservation of the landscape. The reason for this is that Dan's liability risk for the harm he creates would be asymmetrical: Jacob could potentially sue him, whereas the landscape cannot. As in the defensive medicine example (Example 2), here, too, the asymmetrical threat

been negligent, since he could have braked the train more quickly than he had done. The court found that had the engineer in fact so acted, he would have endangered the train passengers. Accordingly, the court concluded that the engineer had not been negligent, while stressing that his duty to the train passengers superseded his duty to the passengers in the truck.

13 See Keeton et al., *supra* note 7, at 1056-59.

of liability means that one course of action causes Dan to externalize costs to society, while the other course of action does not. Thus, Dan would become a partial agent who cares more about avoiding liability than advancing social interests. No-liability or reduced liability could be a suitable solution for Example 4.

The asymmetrical risk faced by the injurer when he is required to balance the interest of an individual against a social interest is a consideration weighed in many tort cases. However, the application of this consideration changes in cases in which the injurer is a public authority. Arguably, public authorities are typically *a priori* biased in favor of social interests and occasionally fail to give proper weight to individual interests that might collide with social interests that the public authorities strive to preserve. Tort liability for negligence could tilt the scale in favor of individual interests and contribute to the impartiality of public authorities. The risk that liability would make them partial in favor of individual interests can be decreased if the scope of the liability is either reduced or limited to cases of gross negligence only.¹⁴

Category 5: The Agent Balances among Her and Another Party's Converging Interests

Example 5: The Stairs. Lisa slips on the steep stairs in Bill's house and suffers bodily injury. Bill could have decreased the risk of slipping on the stairs had he installed a safety railing, but he did not do so. Had there been a railing, the damage would have been prevented.

An important feature of Example 5 is that had Bill taken precautions (by installing a safety railing), not only would Lisa and his other guests have been the beneficiaries of this action, but Bill himself would have benefited. This means that, in contrast to the first category of cases, where the assumption is that the injurer and the victim have conflicting interests, in this fifth category of cases, their interests partly converge.¹⁵ This characteristic of the fifth category of cases could serve to justify the conclusion that liability is less crucial to the injurer's incentives here than in the first category. This is so because, when deciding whether or not to take precautions, the partiality of the

¹⁴ Porat, *supra* note 1, at 138-40.

¹⁵ Upon closer examination, Example 1 could easily be classified also under this fifth category, since decreasing John's speed would have resulted in a lower risk to both Tony and John.

injurer is mitigated: her self-interest could convince her to take precautions, regardless of the liability rule. Thus, the fact that taking precautions would benefit the injurer herself creates a conflict of outcome interests for the injurer, and these conflicting interests could even cancel each other out and render the injurer an impartial agent.

This does not mean that liability is not necessary in fifth category cases. In fact, liability is necessary for supplying the injurer with the proper incentives to take *adequate* precautions.¹⁶ The reason is that if the injurer is not held liable, she will tend to take fewer precautions than required by efficiency, because she will not care about the risks she imposes on others. Thus, if in Example 5, Bill's self-risk is 2 and the risk to others is 3, if he is not held liable, he will buy safety railing only if it costs no more than 2, even though efficiency would require that he buy the railing if it costs less than 5. We see that even though the partiality of the injurer is mitigated because of her self-interest in taking precautions, it does not disappear altogether.

Cases belonging to the fifth category include car accidents in which the injurer was also exposed to an increased risk because she failed to take precautions; property maintenance accidents where the owner or occupant of the property who was responsible for its maintenance was exposed (together with the victim) to the risks emanating from the property; and cases of criminal attacks where it was the defendant's duty to protect the victim from criminals and where the precautionary measures she could have taken to fulfill this duty would also have reduced the defendant's own risks.¹⁷

III. INTERNATIONAL LAW

International law views states as the primary agents in the implementation and enforcement of global public order. The idea is that if states keep internal order and also maintain good neighborly relations with one another, world order will be maintained.¹⁸ In the international arena, the benefits of

16 When applying the Hand Formula, the decision-maker should consider both risks to others and self-risks that would have been reduced had precautions been taken, and not only risks to others, as courts mistakenly do. See Robert Cooter & Ariel Porat, *Does Risk to Oneself Increase the Care Owed to Others? Law and Economics in Conflict*, 29 J. Legal Stud. 19 (2000).

17 *Id.* at 30.

18 The *Island of Palmas* arbitration emphasizes that sovereignty represents an obligation of the sovereign towards the international community to maintain domestic law and order. The *Island of Palmas* (or *Miangas*), Arbitral Award of Apr. 4, 1928, *reprinted in* 22 Am. J. Int'l L. 867, 876 (1928). "[T]erritorial sovereignty involves the

relying on states as agents for promoting global interests are obvious. The definition of international norms and their enforcement are public goods whose production requires complex collective action among agents with diverse interests and capabilities. Hence, effective international norms are constantly undersupplied. At the same time, there are instances in which specific individual states are uniquely situated in a position to define or enforce international law unilaterally. This is the case when the individual state is outcome-impartial, has a strong accuracy interest, and is sufficiently skillful to achieve the accurate outcome. This is also the case when states have reciprocity interests that prompt them to adopt cooperative strategies in the hope of fostering cooperation with other states. Fortunately, as we will demonstrate below, there are several situations in which the conditions of impartiality and skills of a particular state can be ascertained with relative ease. In such circumstances, international law effectively relegates authority to such impartial states to define and refine the law and defers to their unilateral enforcement of the law against violators. Consequently, when state action — whether domestic involving its own citizens or international affecting foreign states and citizens — is reviewed in formal procedures (litigation before the International Court of Justice, the European Court on Human Rights, the European Court of Justice, or the World Trade Organization, for example) or informally, by other states that react, attention is paid to the acting state's degree of impartiality as decision-maker in the matter. Those states that are found to act as impartial decision-makers are not scrutinized, or are scrutinized less strictly than the states perceived as being in a conflict of interest. Moreover, the impartial states' actions and declarations are given greater attention when judges and scholars sift through state practice in an effort to learn about customary international law norms.¹⁹ No doctrine in fact requires such consideration of a state's level of impartiality. Despite the importance of impartial agency for the implementation of international law, there is no explicit recognition of the partiality factor in positive international

exclusive right to display the activities of a state. This right has a corollary duty: the obligation to protect within the territory the right of other states."

19 The elusive concept of "*opinio juris*" (i.e., the belief of states that a certain practice is required by the law) distinguishes between state practice that has no legal effect and state practice that creates customary international law that is binding on all states. State action that is based on sheer self-interest will not be as good a proof of *opinio juris* as an action by a state whose commitment to that practice imposes significant costs on the state. On the impact of the costs of compliance in the evolution of customary law, see Michael Byers, *Custom, Power and the Power of Rules* 151-56 (1999).

law. However, and quite implicitly, attention to impartiality permeates both the law and the legal scholarship.

What factors drive an individual state's accuracy interests? One strong motivation is reputation. Some states cultivate a reputation of what may be called "engaged neutrality." Their neutrality is not indifference to the outcome of a conflict between rivals, but, rather, reflects an interest in facilitating a resolution of the conflict according to the law. States thus motivated seek to acquire the reputation of an honest broker. Countries like Switzerland and the Netherlands come to mind as agents who are ready to invest significant resources to facilitate peaceful conflict resolution among states. Other states wish to attract foreign investments and, hence, seek to develop rules and dispute-resolution bodies that do not adversely affect foreign interests. Still other states have an accuracy interest based on internal forces that demand adherence to the law. States in which international law is incorporated as part of domestic law will be more likely to conform to international norms as they are interpreted by the domestic legislature and courts. Domestic public opinion, informed by human rights or environmental NGOs, can also be the source of a strong accuracy interest that influences the government in its actions.

Other states will be impartial due to their reciprocity interest. Thus, for example, the acts and declarations of riparian states situated on the upstream of one international river and the downstream of another international river will be accorded more deference than the acts and declarations of riparian states located only on the upstream or downstream of an international river. This is in fact quite a prevalent source of states' impartiality, as will be discussed in more detail in the framework of Category 1 cases, below, and perhaps the catalyst behind the evolution of international law.

Finally, there are those states that act not out of outcome impartiality, but in pursuit of their own goals, which are identical to the international collective interest. This is especially true of relatively very strong and skillful states that declare that their actions are aimed at increasing both individual state and global welfare. The US attack on Afghanistan in 2001 aimed at protecting all of civilization is one recent example. The claim of impartiality derives from the correspondence between the individual and collective goals and the skillfulness at identifying those goals and implementing them. Such claims made by relatively strong regional or global powers will be reviewed by other actors, which may or may not accept them as valid, and deference will result either from acceptance of the claim or from acquiescence to the overwhelming relative power positions that render prohibitive the costs of providing better-skilled agents.

In addition to states, there are other agents operating in international law, to

whom the law assigns legal responsibility as "subjects" of international law. These agents include international organizations whose powers are defined by a treaty between two or more states. Their activities are also governed by international law, or at least their activities may be attributed to the states that are parties to those institutions. These organizations can have personal impartiality in the sense discussed in Part I, but in the international law arena, we more frequently encounter institutional impartiality, when such institutions include representatives from different states. These institutions may prove outcome-impartial as well as possessing a strong accuracy interest. Arbitrators, for example, seek to acquire a reputation for accuracy impartiality, as do international tribunals. The actions of such institutions are, therefore, likely to be aimed at attaining social interests more than the unilateral actions of individual states. In this Part, we focus on the actions of individual states, which is a more difficult case to argue.

As in tort law, often there is no clear-cut distinction between partiality and impartiality, and often states are motivated by a host of considerations, merging their self-interests with the interests of other states. We argue that, in the latter case, the need for external enforcement does not necessarily have to amount to the same type and level of enforcement that may be necessary when the agent pursues only its self-interest.

Category 1: The Agent Decides in a Matter Concerning Itself and Another State, in which Their Interests Conflict

Example 6: State A decides not to invest in preventing pollution. As a result, polluted air reaches crops in State B and ruins them.

As in torts, this is the paradigmatic case that raises concern regarding the law's scope of protection. This is also the paradigmatic case that exposes the fundamental weaknesses of international law, namely, the under-supply of generally accepted norms and commitment to effective dispute resolution mechanisms. Much collective effort has been invested in defining the proper conduct of states and the modalities for dispute resolution. The results are often less than satisfactory.

Moreover, similar to tort law, international law recognizes responsibility based on fault and responsibility based on a strict liability rule. Like tort law, international law recognizes direct and indirect liability of state actors for the actions of individuals acting on state authority or in state territory.²⁰

But now let us consider a further development of Example 6:

20 The corpus of laws dealing with such questions is called the law on state

Example 6.1: State A decides not to invest in preventing pollution. As a result, polluted air reaches crops in State B and ruins them. State B retaliates by imposing duties on imports from State A.

Unlike tort law, international law, due to its fundamental weaknesses, permits or acquiesces to self-help as an acceptable or tolerable action in several contexts. The prospect of reciprocal action by State B, when such a prospect exists and assuming a correlation between State B's action against State A and the harm State B suffers from State A's action, will lead State A to internalize the harm it creates. State A will reduce its emissions in anticipation of the subsequent responses of rival states. The literature on the evolution of cooperation among states is packed with examples and game-theoretical explanations of such motives,²¹ which we call in this Article reciprocity interest. Many of the most basic norms of international law, such as the laws of war and the laws on diplomatic immunity, are preserved through the possibility of unilateral retaliation allowed by the law. "Tit-for-tat" norms increase the impartiality of every state that can be adversely affected by such action.

But there are risks to such a policy. State B's response can be excessive, and the tit-for-tat might only intensify the conflict. The major concerns regarding tit-for-tat tactics are the possibility of reaching an equilibrium of non-cooperation after a series of bilateral breaches and the adverse impact this response can have on relatively weaker parties, who would be intimidated by potential retaliation from stronger states. Thus, international law must choose between prohibiting, restricting, tolerating, and even encouraging self-help. This choice, we submit, is influenced by the extent to which it can be expected that the exercise of self-help will be abused by, for example, the stronger states, who will not fear revenge and whose reciprocity interest will be limited. In short, the ideal self-helper is the state whose self-help is constrained by its accuracy or reciprocity interests. In the context of the use of force, for example, the law allows for reaction in self-defense, but regulates it (by demanding immediacy, necessity, and proportionality) to reduce the likelihood of uncalled-for escalation. A similar

responsibility. See International Law Commission, Draft Articles on Responsibility of States for Internationally Wrongful Acts (2001); James Crawford, *The International Law Commission's Articles on State Responsibility Introduction, Text and Commentaries* (2002).

21 See Robert Axelrod, *The Evolution of Cooperation* (1984); Todd Sandler, *Collective Action* (1992); Eyal Benvenisti, *Collective Action in the Utilization of Shared Freshwater: The Challenges of International Water Resources Law*, 90 Am. J. Int'l L. 384, 390-94 (1996).

decision mechanism should be set with respect to self-help in other instances, say, in a situation of environmental self-help,²² in light of the potential costs and benefits involved.

But often there is much vagueness in the law on these issues, and our observation is that where State B's response reflects its impartiality, the law will tolerate it as a legitimate exercise of self-help action. Impartiality in this context would mean that State B has only an interest in wiping out the consequences of State A's action, without gaining other benefits (which would render it partial). Proportionate response is an important factor for understanding the law's attitude towards self-help. When State B responds to a breach of an obligation by State A, State B is less likely to be considered as acting non-proportionately if it incurs costs as a result of its response. For example, State B, in responding to State A's breach by curtailing trade with State A, is expected to lose from this loss of trade if other states do not join in the boycott. Such a loss can be seen as a credible commitment to the promotion of the international norm breached by State A and, thus, reflecting State B's outcome interest — and sometimes also accuracy interest — in promoting general interests, even if the effort is costly. Indeed, the boycott of Apartheid South Africa was not seen as a non-proportionate countermeasure by the states that participated in the boycott, since it actually harmed the economies of those states. In this type of situation, State B could be classified under Category 5, since the state balances among its own interests and another party's converging interests. A different incentive structure occurs when the retaliating state does not reap any gains (as in Category 2 below), and yet another arises when it can gain from the countermeasures (the paradigmatic Category 1 case). For reasons explained below, it is necessary to develop a sensitivity to the different incentives that motivate states so as to ensure optimal deterrence of retaliating states. International law has an interest in promoting optimal unilateral retaliation, and attention to impartiality is the key to identifying the optimal means of deterrence.

22 Such a self-help claim was made by Hungary, the lower riparian state on the Danube River, when it reacted to the unilateral diversion of the river by Slovakia, the upper riparian state. *Case Concerning the Gabčíkovo-Nagymaros Project (Hung. v. Slov.)*, 1997 I.C.J. 7 (Sept. 25, 1997), *reprinted in* 37 I.L.M. 167 (1998).

Category 2: The Agent Decides in a Matter Concerning Only Another State

Example 7: State A intervenes militarily in State B to prevent the genocide of State B citizens.

Often states that breach international obligations claim exceptions to these obligations, exceptions that are motivated by concern for the welfare of other states. Take the case of "humanitarian intervention," namely, the incursion into the territory of another state to save the citizens of that state from internally inflicted genocide. As a rule, when a neighboring state resorts to this claim, it rings hollow. Thus, such claims made by India (in 1971) and Vietnam (in 1979) to lawful humanitarian intervention in the neighboring states of Bangladesh and Cambodia, respectively, were rejected as unacceptable by the United Nations, despite the fact that their interventions did in fact bring an end to wide-scale genocide. The international community viewed both India and Vietnam as having partial interests as neighboring states seeking partial gains and using the atrocities as a pretext for achieving their own goals. On the other hand, when intervention in a state to prevent atrocities was practiced not by the immediate neighbor, who would be suspected of furthering partial interests, but, rather, by non-neighboring states having no self-interest but, instead, motivated by the desire to end human suffering, these instances were deemed acceptable by the international community as lawful instances of humanitarian intervention. Thus, the interventions of Tanzania in Uganda (1979) and of NATO forces in Kosovo (1999) were deemed morally and politically acceptable, and international lawyers worked to find a formal legal basis to justify or excuse them.²³

Our emphasis on a state's outcome-impartiality as a determining factor in analyzing lawful action is crucial in this context. For this reason, most international lawyers have tried to justify the humanitarian intervention doctrine by focusing on the situation in the invaded state (to what extent individuals there were being massacred or about to be massacred and whether or not the local government was behind it). They have imposed a very high threshold for justified intervention because they are concerned about the possibility of abuse by regional bullies who act out of self-interest. Sharing this concern over possible abuse, we suggest focusing directly on

23 See, e.g., Louis Henkin, *Kosovo and the Law of "Humanitarian Intervention,"* 93 Am. J. Int'l L. 824 (1999); Bruno Simma, *NATO, the UN and the Use of Force: Legal Aspects,* 10 Eur. J. Int'l L. 1 (1999).

the motivation of the intervening state. If it can be ascertained beyond reasonable doubt that the responding state was motivated by either accuracy interests (domestic public opinion has proven to be a significant factor in recent years motivating governments to respond to atrocities in other countries) or outcome interests that conform to the global social interest, then its action can be justified. The muffled responses to the NATO attack on Serbian forces in Kosovo in 1999 can be seen not as a relaxation of the prohibition on the use of force but as (or also as) an affirmation of the delegation of enforcement powers to impartial agents.²⁴

Focusing on the motivation of the intervening state is important if the international community wishes to enable states to intervene when needed. On the one hand, the intervening state has an accuracy interest in preventing a massacre. On the other hand, it has an interest not to be deemed an illegitimate aggressor that violated basic international law obligations and to potentially suffer adverse consequences as a result. A too-high threshold for a justified intervention would provide strong disincentives for impartial states to intervene. If impartiality becomes a relevant factor in determining legality for humanitarian intervention, there is a higher likelihood that impartial states will not be deterred from intervening when necessary.

Category 3: The Agent Decides in a Matter Involving Only Other States with Conflicting Interests

Example 8: A national court in State A decides in a conflict between citizens of State A or of State B concerning claims of violations of international human rights.

State institutions may often decide matters involving the application of international law in disputes involving two or more conflicting interests that do not implicate the interests of the deciding institution. National courts perform this function in two types of cases: when they adjudicate conflicting claims of two non-nationals and when they adjudicate conflicting claims of two of their own nationals. These institutions lack an outcome interest, but they do have a strong accuracy interest, especially when they adjudicate claims of their own citizens.

When a court in State A adjudicates claims of or against a citizen of State B, in addition to an accuracy interest, the court may have also a reciprocity

²⁴ On the analysis of the U.S. attacks on Afghanistan and Iraq, see *infra* text accompanying notes 33-34.

interest, as it expects courts in State B to act in a similar way when they, in turn, adjudicate claims of or against a citizen of State A. Either the accuracy interest or the reciprocity interest might overcome here the temptation of national courts to rule in favor of their own citizens. But note that the courts in both states may be in a prisoner's dilemma: if the impartiality of the counterpart court is not guaranteed, the reciprocity interest will be weak. Under such a scenario, the national court will still have an accuracy interest. But this accuracy interest may be offset by a negative reciprocity interest (i.e., an interest in retaliating for the partiality of the other state's court) or by a strong national outcome-interest. Also, especially where the application of international law is concerned, the accuracy interest may not be so particularly strong. As a result, the court may decide not to act impartially and, instead, will choose to favor its own national. This analysis explains the phenomenon of national courts' disregarding their accuracy interest when international law is involved and their tendency to defer to the national interest as perceived by their government when they adjudicate claims involving national and foreign parties.²⁵

One illuminating example of this point is national courts' treatment of the illegal abduction of suspected criminals. The question numerous courts have faced is whether the abduction of a person from the territory of State B by State A officials in violation of international law or of an extradition treaty precludes the prosecution of that person in State A courts. Invariably, State A courts have found the violation of international law irrelevant and have declared that it has jurisdiction over the abducted person.²⁶ On one occasion, however, the UK House of Lords decided that an abduction of a person from South African territory invalidated its jurisdiction in the case.²⁷ The Law Lords explained their groundbreaking decision, *inter alia*, on the basis of the fact that the new, post-Apartheid South African national court reacted in a similar way to a previous abduction from a third country to South Africa. Faced with a clear promise of reciprocation, the House of Lords responded with a similar stance based on reciprocity interest. But

25 Eyal Benvenisti, *Judicial Misgivings concerning the Application of International Law*, 4 Eur. J. Int'l L. 159 (1993).

26 This is the general rule, known in the U.S. as the "*Ker-Frisbie* doctrine," that "the power of a court to try a person for crime is not impaired by the fact that he had been brought within the court's jurisdiction by reason of a 'forcible abduction.'" *Frisbie v. Collins*, 342 U.S. 519, 522 (1952) (citing, *inter alia*, *Ker v. Illinois*, 119 U.S. 436 (1886)).

27 *Reg. v. Horseferry Road Magistrates' Court, Ex parte Bennett* (No. 3), [1995] 1 Crim. App. R. 147.

the chain of reciprocation collapsed when the suspect was subsequently transferred from the British police to the Scottish police and the Scottish court found a way to uphold its jurisdiction over the abductee.²⁸

As a result of a general lack of a reciprocity interest, on the one hand, and a usually strong outcome interest, on the other hand, national courts are generally not a good example of impartial actors in the international arena. They can be referred to as such only in rare instances.

One such rare instance is the litigation under the U.S. Aliens Tort Claims Act of 1789. This Act was interpreted as granting U.S. federal courts jurisdiction to adjudicate tort claims brought by a non-national for the violation of international law by another alien. Beginning in 1980 with the famous *Filartiga* case,²⁹ this tool was instrumental in quite a few disputes, including, recently, the successful litigation against Swiss banks and German companies for their illegal gains during the Holocaust. Although the assumption of universal jurisdiction by the U.S. Congress is an odd phenomenon, no serious claim has been made that this is a violation of international law. We suggest that this is due to the apparent outcome-impartiality of the U.S. federal courts, which is accompanied by at least a minimal accuracy interest that ensures that the courts are not indifferent to the claims they adjudicate.

It is, of course, possible that U.S. interests may be implicated in this type of litigation, if one of the parties has an interest that is similar to the interest of the U.S. government (for example, the claimant brings a suit against a foreign government, and the U.S. government has an interest in the outcome of the claim). In such a case, the court, given the observed tendency of national courts to conform to their governments' interests,³⁰ will likely abandon its presumption of impartiality, unless it rules against the national interest.

Our focus on the impartiality of agents can explain another phenomenon in international law: the tendency of international institutions to pay greater deference to decisions of democratic national institutions that reflect a genuine political give-and-take among different domestic interests.

Example 8.1: The European Court of Human Rights decides in a conflict between a citizen of State A and the government of State A concerning claims of violations of international human rights.

28 *Bennett v. HM Advocate*, 1995 S.L.T. 510.

29 *Filartiga v. Pena-Irala*, 630 F.2d 876 (2d Cir. 1980).

30 Eyal Benvenisti, *Margin of Appreciation, Consensus and Universal Standards*, 31 N.Y.U. J. Int'l L. & Pol. 843 (1999).

The European Court of Human Rights developed in this context the doctrine of "margin of appreciation," according to which the Court will defer to the decisions of national legislatures and courts on matters involving domestic conflicts concerning the application of human rights law. The underlying justification is respect for national sovereignty and the assessment that the national decision-maker is a better agent than the international one in assessing the proper balance to be struck between the conflicting interests. In our terminology, the doctrine recognizes the reciprocity interests that play out in the domestic democratic process and the strong accuracy interest that shapes the national courts in these democratic countries.

It has been suggested that this doctrine of margin of appreciation makes sense only when the decision-making process within the relevant state ensures equal opportunity for all relevant domestic voices to influence the outcome. But if the domestic process tends to disenfranchise a minority of citizens, then the assumption does not obtain and the margin of appreciation doctrine should not apply.³¹ In the terms of this article, when national institutions are not influenced by either accuracy interests or reciprocity interests, they cannot be deemed impartial, and hence they should not be given the deference they would otherwise deserve.

This discussion leads to a more general observation concerning the international legal approach towards the behavior of democratic and non-democratic state agents. In general, we can say that international bodies accord greater deference and respect to the decisions of democratic regimes concerning their own citizens than to non-democratic regimes, although formally, international law regards all states equal regardless of their domestic structure of government. Thus, for example, democratic domestic decisions concerning the ways of addressing the crimes of previous regimes (South Africa, Chile) will be deferred to by judicial bodies exercising universal jurisdiction that are contemplating criminal prosecution of foreign criminals (unless the decision was biased against the minority within the democratic state).

Democracies merit more deference than dictatorships not only because they are expected to pursue better outcomes than dictatorships, but also because they are expected to better resolve internal conflicts of interest, due to the internal reciprocity interests of the various political actors and the accuracy interests of impartial courts. A dictator is less influenced by either reciprocity or accuracy interests. In other words, outsiders may assume that a democratic decision related to, say, setting standards for permitted pollution

31 *Id.*

that takes place within the state's boundaries internalized both the interest of the polluters and those who suffer from the pollution, because both sides participated in the formation of the policy. A democracy that resorts to an armed attack, if that attack passed through a democratic decision-making process that entailed debate with the opposition parties, would need to exert less effort than a dictatorship in a similar position to convince the international community of the justification of its resort to force. There is, of course, no official document to even imply such an attitude in international law. Nevertheless, there is vast literature in international relations discussing why democracies do not go to war against other democracies.³² Our focus on the impartiality factor helps to respond to this riddle: democracies do not go to war as often as dictatorships because their internal processes that balance conflicting domestic interests render them more impartial than dictatorships in their international relations.

As in Category 3 tort cases, international intervention in domestic tribunal decisions may prove counterproductive. On the one hand, the possibility of review by the international tribunal will increase the accuracy interest of the domestic tribunal. But on the other hand, the international tribunal will be less skillful, being less well-informed about the considerations that motivated the domestic court. It may not have an accuracy interest as high as that of the domestic court. There is also the risk that, in order to avoid intervention, the domestic court will decide in a way that decreases the chances of intervention instead of increasing the chances of accuracy (for example, if the intervention damages the reputation of the domestic court). Hence, it makes sense to defer to domestic courts that are strongly motivated by accuracy interests and have no outcome interests, in democratic societies where the democratic process allows the participation of all relevant voices.

Category 4: The Agent Decides in a Matter Involving Another State and a Social Interest

Example 9: State A, in an effort to eradicate terrorist activities or drug production in State B, imposes trade or other sanctions against State B. These sanctions are not significantly costly to State A. Instead of unilateral sanctions, State A could have requested the extradition of

³² See, e.g., Spencer R. Weart, *Never at War: Why Democracies Will Not Fight One Another* (1998); Felix Knupling, *Democracies & War: An Investigation of Theoretical Explanations* (2000).

the terrorists or drug dealers (which would probably have been denied by State B).

State A has an outcome interest that is identical to the general social interest. Its success or failure will impact not only global interests but also, and equally, its own interests. It thus has an incentive to promote the collective goal.³³ But if its reaction could trigger international backlash as being a violation of international law, State A is likely not to promote the global interest and, instead, will opt for the ineffective extradition request or remain idle. And so would other states act. If this is the case, the collective good — promotion of general goals by individual agents — will be under-supplied.

The fact that State A has only an outcome interest and no reciprocity interest or accuracy interest suggests that, in general, its preference would be for an intervention decision by another actor rather than for action on its part. The UN Security Council, NATO, and any other regional international organization come to mind as such possible actors. These institutions have an accuracy interest because their policy decisions have a broader scope and affect various other actors in similar situations.

The U.S. attack on Afghanistan in 2001 was generally regarded as a clear case of self-defense, in the face of the continuing threat of terrorist activity emanating from Afghanistan.³⁴ Doctrinally, however, it remains difficult to explain the attack on the Taliban forces (rather than Al-Qaeda targets), given the fact that the Taliban regime was not directly involved in the September 11 attacks. From the doctrinal perspective, the U.S. was expected to try and negotiate with the Taliban the extradition of the Al-Qaeda leaders before resorting to the military option. But the overwhelming international support for the U.S. action can be justified not only by the devastating terrorist attack it suffered but also by the general understanding that the U.S. was attempting to eliminate a threat to the entire international community. This understanding in itself is not enough for several commentators who emphasize the implicit approval of the UN Security Council of *any* action taken by the US.³⁵ This

33 For an analysis of the United States' role in obtaining the global public good of security and the effect of this role on international law, see Eyal Benvenisti, *The US and the Use of Force: Double Edged Hegemony and the Management of Global Emergencies*, 15 Eur. J. Int'l L. 677 (2004).

34 See, e.g., Thomas Franck, *Recourse to Force: State Action Against Threats and Armed Attacks* (2002); Michael Byers, *Preemptive Self-Defense: Hegemony, Equality and Strategies of Legal Change*, 11 J. Pol. Phil. 171 (2003), Christopher Greenwood, *International Law And The Pre-Emptive Use Of Force: Afghanistan, Al-Qaida, And Iraq*, 4 San Diego Int'l L.J. 7 (2003).

35 This implicit approval was given in Security Council Resolution 1368 of September

approval by an institution that has an accuracy interest is a sufficient (and, for some, necessary) condition for approval of an action that is doctrinally suspect. By focusing on the impartiality factor, we can conclude that the acceptance of the U.S. action does not lend support to the claim that the law on the use of force has changed as a result of the Afghanistan "precedent." Rather, we can conclude that when the incentives of the agent conform to the general social interests and, particularly, when an authoritative impartial international actor supports the agent's action, then international law will not view that action as a violation of its norms. In such a case, the agent promotes individually the general interest.

Can the same rationale apply to the U.S. attack on Iraq in 2003? In this case, the U.S. failed to convince the UN of the necessity of the military operation. This raises a difficult question concerning the unilateral use of force by states in promoting global interests without the approval of the UN or other international institutions. This is a debate that rages nowadays among international lawyers and is beyond the scope of our Article. Our analysis can be used to address the questions of whether the U.S. can be credibly regarded as pursuing the general interest in preventing the concentration of weapons of mass destruction in the hands of rogue regimes and, if so, whether such outcome-impartiality can justify its actions without further eroding the doctrinal conditions regarding the prohibitions on the use of force.

A much less dramatic, but more systematic, question arises in the context of unilateral state attempts to raise global standards of environment protection and other practices aimed at more sustainable management of global resources. During the twentieth century, several coastal states acted unilaterally to assert jurisdiction over the fisheries near their shores or over the continental shelves. These unilateral acts were motivated by these states' self-interest, but they also reflected the collective interest in sustainable management of these resources. In such cases, the state asserting control over previously common property was not constrained by a reciprocity interest because retaliation against it was not an issue. An accuracy interest was lacking as well, and the likelihood of international enforcement was minimal. Nevertheless, there was a clear convergence of the states' outcome interest and the public interest.³⁶

12, 2001 ("*Recognizing the inherent right of individual or collective self-defence in accordance with the Charter*").

36 See the Fisheries Jurisdiction Case (U.K. v. Ice.), 1974 I.C.J. 2, where Iceland extended its maritime jurisdiction unilaterally by fifty miles to regulate offshore fishing. The International Court of Justice found this a violation of customary

But the lack of accuracy or reciprocity interests should not suggest that these states violated international law, although technically, they did act in breach of their customary international law obligations. International law must be careful to offer incentives to states to act unilaterally to promote global standards, because this is perhaps the main vehicle for keeping customary international law efficient. The evolution of the law depends on countries that breach inefficient customary norms and unilaterally adopt new norms that are more efficient. Technically, states that adopt higher or more efficient norms may be considered in violation of international law. But they may be initiating a transition to a more efficient norm. There is a strong collective interest not to impose too high a burden on states that technically violate the existing law but, in fact, proclaim what is likely to become the new standard. Thus, for example, the first states that made claims to unilateral control over the continental shelf of the oceans and the 200-mile exclusive economic zone not only breached the then-prevailing customary international norms, but also infringed on the right of access and exploitation of other states to those common resources. But because these claims made sense from the perspective of the efficient allocation of resources, the outcome interest of these states was identical to the general interest, and hence they could be regarded as acting as effective agents in the progressive development of the law. What we can suggest in this context is that when states and international actors such as international tribunals and commissions address actions by other states that unilaterally infringe customary international law but that are justified as motivated by an outcome interest that conforms with a public interest, then such actions are entitled to the most serious consideration. Scientific evidence supporting the justificatory claims of the acting states can be relied upon by those considering the actions of the state as sufficient basis to declare a change in the law. In fact, it has been argued elsewhere that this is what international adjudicators have been doing all along.³⁷

international law. It refused to recognize that the law has evolved given the fact that multilateral negotiations were being held over a new convention on the law of the sea. On this case and its significance, see Eyal Benvenisti, *Customary International Law as a Judicial Tool for Promoting Efficiency*, in *The Impact of Norms and Institutions on International Cooperation* 85 (Eyal Benvenisti & Moshe Hirsch eds., 2004).

³⁷ Benvenisti, *supra* note 36.

Category 5: The Agent Balances among Its and Another State's Converging Interests

Example 10: State A decides to adopt standards for the emission of fumes. The standards impact both State A's territory and the territory of State B, due to the fumes borne by the wind into the latter's territory. International law does not prescribe any standards for fume emission, but there is a general duty to prevent "significant harm" from pollution.

This example involves actors who cannot retaliate against each other. Hence, unlike in Category 1, State A has no reciprocity interest that can motivate it to comply with international law. Also, in general, State A has no accuracy interest in this matter. Therefore, such situations call for outside intervention to implement the norm against significant harm. But when an international organ, such as an international commission or court, takes up this issue, a question will arise as to the compatibility of State A's standards with international law. In this context, the tribunal will have to examine whether the standards adopted by State A cause "significant harm."³⁸ Because there is no clear-cut objective definition of significant harm, the tribunal will ask itself whether the fact that State A's standards also affect its own territory is indication that it adopted a proper standard.³⁹ Our focus on impartiality would suggest that at least some deference to State A's standards would be appropriate only if these standards are the outcome of a domestic process in which conflicting interests were weighed either by a legislature (which has an accuracy interest due to the internal reciprocity interest of the various political actors) or a court (which also has an accuracy interest). The more State A internalizes the environmental damage, the more deference is due to State A's fume emission standards. On the other hand, the more damage suffered by State B, the less the international tribunal can rely on the assumption that State A's institutions acted impartially.

38 See the famous arbitral award in the Trail Smelter (U.S. v. Can.), 3 R.I.A.A. 1905, 1938 (1949).

39 This is the "golden rule" suggested by Thomas Merrill. Thomas W. Merrill, *Golden Rules for Transboundary Pollution*, 46 Duke L.J. 931 (suggesting that adjudication of such disputes should be governed by the "golden rule" — do unto other states as you do unto your own citizens — and the "reverse golden rule" — do not ask of other states what you do not ask of your own citizens).

CONCLUSION

Legal analysis usually focuses on the actions and reactions of formal decision-makers such as legislatures and courts as the prime candidates for resolving conflicts of interest. This Article throws light on the individual actor who, in certain situations, can prove to be a better societal agent than the formal institutions in resolving such conflicts. The Article develops the conditions for impartiality that qualify the agent to pursue social goals and then offers five categories of conflicts of interest, which serve to identify the agent's impartiality. The Article argues that due to an agent's impartiality (assuming the agent is skillful and enforcement is costly), enforcement of the law by formal decision-makers is unnecessary and can often even be detrimental to the resolution of the conflict and promotion of the social interest.

This analysis suggests that formal decision-makers, when reviewing and evaluating the actions of different actors, should take into consideration the potential for enforcement of the law through individual agents. Courts, in particular, should develop a sensitivity to the variety of conditions that ensure the impartiality of agents and defer to their judgment when they are impartial. Our Article explores the validity and the promise of focusing on the agents' impartiality. We believe that further analysis in many other areas of law will demonstrate that agent impartiality is a valid factor that is, or should be, taken into account in positive law.