Volume 5
Issue 1
Summer 2012

European Journal of Legal Studies

www.ejls.eu

Submissions

The European Journal of Legal Studies invites submissions from professors, practitioners, and students. In particular, we welcome articles demonstrating a comparative, contextual and interdisciplinary approach to legal scholarship. Articles should be submitted in electronic format to subejls@gmail.com. If electronic submission is not practicable, please send a hard copy to:

The Editor-in-Chief
European Journal of Legal Studies
c/o EUI Department of Law
Villa Schifanoia, Via Boccaccio 121
50133 Florence – Italy

Footnotes should conform to The Oxford Standard for Citation of Legal Authorities, available for downloading from **www.law.ox.ac.uk/publications/oscola.php**.
The name of the author and contact details should appear only on a separate cover sheet to facilitate objective, anonymous evaluation. The Journal is committed to multilingualism and accepts submissions in any European language within the competence of the Board.

EJLS

European University Institute
DEPARTMENT OF LAW

European Journal of Legal Studies
Vol. 5 No. 1 (Summer 2012)

TABLE OF CONTENTS

Book Reviews

EDITORIAL

Benedict S. Wray*

I begin my final editorial on a slightly different note to my previous offerings. I would like to say that it has been my privilege and pleasure to take the helm of this journal for the last eighteen months. It is an exciting time for academic publishing (indeed, as I opined in my first editorial, an exciting time for academia generally). Although my stay has been brief, it has been both challenging and enjoyable and I think represents the dynamic nature of the EJLS. Much has changed since the journal's ambitious launch five years ago, and despite the necessary process of fine-tuning its internal structure to cope with the pressures of online publishing, I believe it has maintained its core values of excellence in writing and training. It is only right that my stay should be fairly brief, given the focus which the EJLS places on providing opportunities to the next generation of academics to learn valuable reviewing and publishing skills, and I look forward with interest with watching its further development from the wings as I step down and leave the stage to Tiago Andreotti e Silva, who will take over from the next issue onwards.

LAW AS STRATEGIC CHOICE

A time of great change is afoot in many areas of law, despite the stubborn insistence of a few that the classic structures of law remain unaffected by globalization. Anthony Kronman wrote in 1993 that there was a crisis of identity in the legal profession, left by the collapse of what he called the 'lawyer-statesman' idea and the increasing commercialisation and division of labour amongst lawyers themselves.[1] Interestingly, he rubbished the distinction between those who practice law for money and those who seek to use the law to 'do good', suggesting that both parties view the law in the same way: in strictly instrumental terms.[2]

With the growth of large one-side-only bars, and the adoption of commercial tactics by many, if not all firms, it is worth asking what remains to legal practice beyond another battleground upon which various actors can seek furtherance of their goals and frustrate the goals of their opponents or competitors. Indeed, many in legal practice today would be surprised – perhaps even shocked – by the suggestion that it is anything

* European University Institute (Italy). Any errors or omissions are entirely my own. I am grateful to Andrea Talarico for her feedback.

else. Let me then explain why the question is worth asking. I begin with an example: that of the well-known phenomenon of the 'Italian torpedo' in international civil litigation. This derives from the fact that Italian courts are often slow-moving. In the Trasporti Castelletti v Hugo Trumpy case,[3] bills of lading were delivered which contained a choice-of-court clause in favour of England. However, the receiver of the cargo brought proceedings in Italy where the court took a full ten years (or eight years if the two years in the ECJ are discounted) to decide that it had no jurisdiction. What this enables is a 'torpedo' action by a bad-faith party to sue first and write letters later, in the full knowledge that it may take many years (and expense) for its opponent to obtain satisfaction, in the hope of forcing the latter to accept an unfavourable settlement.

This may seem pretty benign when the parties are both significant commercial players with well-resourced legal advisors. However, what about where one party is a wrongdoing multinational and on the other side are stacked its numerous victims? In the infamous Bhopal disaster, Union Carbide used their best efforts (successfully) to send the litigation back to India, where they settled the case for a sum that was manifestly inadequate, and nearly 30 years after the original explosion many victims remain unpaid.

Such examples are endemic in a culture which sees the law as a mere tool for litigating or contracting parties, and lawyers who see themselves as no more than mouthpieces of their clients' agendas. However, I would argue, it is the law itself that is cheapened by being instrumentalised in this fashion. It has not always been so, if Kronman's history of the lawyer-statesman is to be believed. And occasional examples bear witness to exceptions. In his autobiography, Gandhi relates the story of a certain case where he successfully represented the defendant in a commercial arbitration, but realized it would be impossible for the plaintiff to pay without declaring bankruptcy. He therefore persuaded his client to accept payment in instalments spread out over a 'very long period'. 'My joy was boundless', he wrote, 'I realized that the true function of a lawyer was to unite parties riven asunder'.[4]

When we sent out the original call for papers for this issue I received an interesting email from a Professor who was shocked that we had adopted the term 'lawfare' at all. To conceive of law in this way was an affront, in his mind. Perhaps he is right: in an age of ruthless competition in various fields, it could be argued that the last thing that is needed is the reduction of law to an all-purpose battering ram. If, as Kronman feared and globalization has proved, law is seen increasingly as a mere tool and those who practise it as technicians, perhaps the time has come for a new

identity for lawyers which reifies the ideal of seeing beyond the parties to a dispute to wider societal needs for a global age, be that in crafting new constitutional approaches, human rights or justice regimes, or simply updating professional ethics to take account of the tendency to instrumentalise. Perhaps the word which law should be married to is not war, but welfare.

IN THIS ISSUE

This issue has a very balanced feel. Cambien and Moorhead provide two very different perspectives on EU law, Cambien by revisiting the concept of EU citizenship and its effect on the immigration law of member states, and Moorhead by challenging the generally accepted view of EU law as sui generis, arguing that it is an international legal based on the same underlying rationales as general international law. Meanwhile, Marín García offers a classic comparative examination of penalty clauses, updated for the transnational age, suggesting we need clear transnational rules to manage the friction between the civil and common law traditional approaches.

In international law, Zivkovic, Pervou and Perez de la Fuente offer general articles covering the recognition of contracts as international investments, the Convention on Enforced Disappearance, and culturally motivated crimes, respectively. Pervou's article in particular provides a topical contribution to the debates surrounding the regulation of terrorism, arguing that many practices being adopted in the war on terror constitute enforced disappearance according to the convention. Last but not least, de Hoon gives us our lawfare article, arguing that law is ultimately not up to the task of regulating warfare at all, but creates a false presumption that law can resolve 'fundamental disagreements'.

REFERENCES

[1] Anthony T. Kronman, The Lost Lawyer, (1995, Cambridge, Massachusetts: Belknap Harvard)
[2] Ibid, at 366
[3] Case C-159/97, [1999] ECR I-1597 (ECJ).
[4] M. K. Gandhi, The Story of My Experiments with Truth: An Autobiography, (2007, London: Penguin), Part II, Ch. 14 esp. pp.132-133

UNION CITIZENSHIP AND IMMIGRATION: RETHINKING THE CLASSICS?

Dr. Nathan Cambien*

The free movement of Union citizens hinges on three 'classic' requirements, namely the possession of Member State nationality, the inter-State element and the condition of self-sufficiency. Recent case law of the ECJ seems to shake the traditional conceptions of these requirements and, as a consequence, to widen the scope of application of the free movement rules. This in turn will have significant consequences for the immigration laws of the Member States. On the one hand, Union law will increasingly influence the Member States' rules on acquisition and loss of nationality. On the other hand, the Member States will have to accord residence rights to certain categories of Union citizens and their family members who would previously not have been entitled to invoke Union law. The resulting financial burdens for the Member States are potentially very significant, although it is not yet possible to ascertain the precise reach of the principles articulated by the ECJ.

TABLE OF CONTENTS

* Research fellow, Institute for European law, University of Leuven; Visiting Research Fellow, Columbia Law School. I am grateful to Janek T. Nowak and Timothy Roes for their valuable comments and suggestions. All mistakes remain my own.

I. INTRODUCTION: 'CLASSIC' ELEMENTS OF FREE MOVEMENT

In the most basic terms, the right to free movement enjoyed by Union citizens is rather straightforward: every Union citizen is entitled to move to another Member State and reside there if he can prove that he is either economically active or has sufficient financial resources at his disposal. This general sketch already reveals that the free movement of Union citizens is centred on three basic elements, which can be labelled 'classic' elements of free movement. First of all, it is clear that the right to free movement is only enjoyed by Union citizens, i.e. persons who have acquired and retained the nationality of a Member State in accordance with the nationality rules of that Member State. Second, it can only be invoked by Union citizens once they leave their Member State and move to another Member State. Static Union citizens, i.e. Union citizens who have never resided in a Member State other than that of their nationality, cannot normally invoke the benefits related to the right to free movement. Third, Union citizens can only reside in another Member State for longer periods of time if they are self-sufficient, i.e. if they have a job or can fall back on sufficient personal means.

These three classic elements are embedded in the Treaties and in secondary Union law, most notably Directive 2004/38,[1] and have been consistently confirmed by the ECJ. Nonetheless, recent case law of the ECJ seems to shake the traditional conceptions of these elements and to

[1] Directive 2004/38/EC of the European Parliament and the Council of 29 April 2004 on the right of the citizens of the Union and their family members to move and reside freely within the territory of the Member States amending Regulation (EEC) No 1612/68 and repealing Directives 64/221/EEC, 68/360/EEC, 72/194/EEC, 73/148/EEC, 75/34/EEC, 75/35/EEC, 90/364/EEC, 90/365/EEC and 93/96/EEC [2004] OJ L158/77.

considerably reduce their importance as requirements for the application of the free movement rules. As a consequence, the scope of these rules is widened. This in turn, I will demonstrate, will have significant consequences for the immigration laws of the Member States. In the following, I briefly discuss for each of the classic elements the traditional approach and its underlying reasons, before analysing the recent evolution in the case law and its likely consequences for the immigration laws of the Member States.

II. MEMBER STATE NATIONALITY

1. *Traditional Approach*

It follows from Article 20(1) TFEU that every national of a Member State is also a citizen of the Union. Traditionally, it is assumed therefore that the Member States autonomously determine the personal scope of Union citizenship, since the Member States have exclusive competence to regulate nationality. Union law, it is traditionally accepted, does not apply in the field of nationality legislation.

The competence to lay down the rules regarding acquisition and loss of nationality is a key competence of sovereign States. Understandably, the Member States have jealously guarded this competence and have never been prepared to transfer any competence in this field to the EU. Precisely for this reason, the Member States opted in the Maastricht Treaty to define Union citizenship by reference to Member State nationality. As such, the Member States arguably intended to prevent Union citizenship from competing with or even superseding Member State nationality. Furthermore, the Member States explicitly stipulated in a declaration annexed to the Treaties[2] and in a decision of the Heads of State or Government meeting within the European Council[3] that Member State nationality is to be determined solely by reference to the national law of the Member State concerned. Accordingly, there simply seemed to be no room for arguing that Union law applied in the field of nationality rules.

Still, it must be remarked that it has been accepted for some time that

[2] Declaration (No 2) on nationality of a Member State, annexed to the Treaty on European Union [1992] OJ C191/98.

[3] Decision of the Heads of State and Government, meeting within the European Council, concerning certain problems raised by Denmark on the Treaty on European Union [1992] OJ C348/1. For a discussion, see Deirdre Curtin and Ronald van Ooik, 'Denmark and the Edinburgh Summit: Maastricht without Tears' in David O'Keeffe and Patrick M. Twomey (eds), *Legal Issues of the Maastricht Treaty* (Chancery Law Publishing 1994) 349-365.

Union law, through the provisions on Union citizenship, indirectly influences the nationality rules of the Member States. Even before the introduction of Union citizenship, the Court had proclaimed that the Member States have to respect unconditionally nationality measures adopted by other Member States.[4] This duty of unconditional recognition can set in motion a subtle interplay between the Member States, whereby rules and practices regarding nationality in one Member State may have significant consequences for other Member States. Indeed, Member States with flexible nationality rules make it easy for third country nationals to acquire Union citizenship, which in turn entitles them to claim rights and benefits in all other Member States. For this reason, Member States with flexible rules may come under pressure from other Member States to restrict their rules. The most famous case in point is the restriction of Irish nationality rules in 2004 after the flexible Irish nationality laws had come under pressure in circumstances that gave rise to the *Zhu en Chen* case[5].

Besides, the Court had held in its *Micheletti* judgment, in a famous *obiter dictum*, that it is for each Member State, *having due regard to Union law*, to lay down the conditions for the acquisition and loss of nationality.[6] However, while in the almost 20 years following *Micheletti*, the ECJ repeated this dictum in a number of cases,[7] it had never clarified its meaning by stating what principles of Union law Member States must respect in this connection or found a Member State's nationality legislation to be in breach of Union law. This led to a vivid debate in legal literature about the possible meaning and significance of the dictum.[8] In its *Rottmann* judgment of 2 March 2010,[9] the ECJ for the first time

[4] This principle was articulated for the first time in Case C-369/90 *Micheletti* [1992] ECR I-4239.

[5] Case C-200/02 *Zhu and Chen* [2004] ECR I-9925, with a case note by Kristien Vanvoorden in (2005) Colum. J. Eur. L. 305-321. See the discussion in Bernard Ryan, 'The Celtic Cubs: The Controversy over Birthright Citizenship in Ireland' (2004) 6 Eur. J. Migration & L. 173-193.

[6] Case C-369/90 *Micheletti* [1992] ECR I-4239, para 10 (emphasis added).

[7] Case C-179/98 *Mesbah* [1999] ECR I-7955, para 29; Case C-192/99 *Kaur* [2001] ECR I-1237, para 19; Case C-200/02 *Zhu and Chen* [2004] ECR I-9925, para 37.

[8] See, for instance, the longstanding debate in the Netherlands between De Groot (see *inter alia* Gerard-René De Groot, 'Towards a European Nationality Law' (2004) 8.3 EJCL <http://www.ejcl.org/83/art83-4.PDF> and the literature cited therein) and Jessurun d'Oliveira (see *inter alia* Hans Ulrich Jessurun d'Oliveira, 'Nationality and the European Union after Amsterdam' in David O'Keeffe and Patrick M. Twomey (eds), *Legal Issues of the Amsterdam Treaty* (Hart Publishing 1999) 395-412).

[9] Case C-135/08 *Rottmann* [2010] ECR I-1449, with case notes by Dimitry Kochenov in (2010) 47 CML Rev. 1831-1846; Nathan Cambien in (2011) 17 Colum.

clarified to some extent the meaning of the phrase 'having due regard to Union law'.

2. *The Rottmann Case and its consequences*

a. The Case

The facts of the case are rather peculiar. Mr. Rottmann was an Austrian national who was prosecuted in Austria on account of suspected serious fraud in his profession. While the judicial investigation was ongoing, he moved to Germany and acquired the German nationality. When the competent German authority learned of the pending proceedings against Mr. Rottmann, it reacted by withdrawing his naturalisation with retroactive effect, considering that, by failing to disclose this relevant information, Mr. Rottmann had obtained the German nationality by deception. The withdrawal decision had rather disastrous consequences for Mr. Rottmann. As a consequence of his naturalisation he had lost his Austrian nationality, in accordance with both Austrian and German law. The withdrawal decision would strip him of his only remaining nationality, the German nationality. Consequently, the interplay of Austrian and German provisions on nationality in the circumstances of the case threatened[10] to render Mr. Rottmann stateless.

The question the ECJ had to answer was whether this outcome was in accordance with Union law, in particular with the provisions on Union citizenship. The Court started by tackling the question of admissibility, namely by determining whether Union law was applicable to the dispute at all.[11] It famously stated that the situation of Mr. Rottmann fell, 'by reason of its nature and its consequences', within the ambit of Union law.[12] This is a point of paramount importance to which I will come back in more detail below.[13] Next, the Court assessed whether the withdrawal decision was taken in accordance with Union law.[14] The Court accepted that withdrawal of nationality for reasons of deception could be compatible with Union law, since such corresponds to a reason relating to the public interest, namely the protection of the special relationship of solidarity and good faith between a Member State and its nationals. It added, however, that,

---

J. Eur. L. 375-394.

[10] The effects of the withdrawal decision were suspended by the appeal brought against it by Mr. Rottmann. Accordingly, the effects of the decision under Austrian and German law had not yet materialised when the ECJ delivered its judgment.

[11] Case C-135/08 *Rottmann* [2010] ECR I-1449, paras 37-45.

[12] Case C-135/08 *Rottmann* [2010] ECR I-1449, para 42.

[13] See under III.B.1, *infra*.

[14] Case C-135/08 *Rottmann* [2010] ECR I-1449, paras 46-59.

where the withdrawal of nationality has for a consequence that the person concerned loses his Union citizenship, this decision must respect the principle of proportionality. It was necessary, therefore, to balance the consequences of the withdrawal decision for the person concerned and his or her family members with regard to the loss of the rights enjoyed by every Union citizen against the gravity of the offence committed by that person, the lapse of time between the naturalisation decision and the withdrawal decision and the possibilities for that person to recover his original nationality.

b. Analysis

Nationality Rules within the Scope of Union Law

In *Rottmann* the Court for the very first time directly assessed Member State nationality rules in the light of Union law. The Court justified its competence for carrying out this validity assessment by pointing at the intrinsic link between Member State nationality and Union citizenship. Every national measure entailing the loss of Union citizenship automatically entails for the person concerned the loss of his most fundamental status under Union law[15] and has for a consequence that this person can no longer exercise his citizenship rights in the different Member States. Consequently, such a decision will fall 'by reason of its nature and its consequences' within the scope of Union law. This justification seems to confirm the expectation that the Court will henceforth be prepared to screen Member State rules and decisions on loss of Member State nationality. This will probably be different only in cases where the person losing his Member State nationality preserves or at the same time acquires the nationality of another Member State because under such circumstances there will be no impact on the Union citizen status of that person.

An important question left unanswered by the *Rottmann* judgment is whether the Court's reasoning should be confined to cases of loss of nationality or should be held equally applicable in cases of acquisition of nationality, or the refusal thereof.[16] By its very nature, the acquisition of Member State nationality confers upon an individual the most

[15] According to settled case law, Union citizenship is 'destined to be' or 'intended to be' the most fundamental status of nationals of the Member States (see eg Case C-184/99 *Grzelczyk* [2001] ECR I-6193, para 31).

[16] See the discussion in Gareth T. Davies, 'The entirely conventional supremacy of Union citizenship and rights' (2010) EUDO Citizenship Forum <http://eudo-citizenship.eu/citizenship-forum/254-has-the-european-court-of-justice-challenged-member-state-sovereignty-in-nationality-law?start=2>.

fundamental status of nationals of the Member States [17] and this automatically has Union wide consequences, in the sense that the person concerned will be entitled to exercise certain rights in all Member States. Accordingly, cases of acquisition of Member State nationality now arguably fall within the scope of Union law to the extent that they entail Union citizenship. Matters are less clear-cut in case of decisions refusing the grant of Member State nationality. On the one hand, such decisions clearly have consequences that go beyond the remit of the Member State concerned. Indeed, where an individual is denied the nationality of a Member State, this has for a consequence that he will not be able to enjoy the rights attached to Union citizenship. These are the very same rights that a Member State national would lose if his (only) Member State nationality were to be withdrawn. On the other hand, it is not possible to argue that a decision refusing nationality impacts negatively on the citizenship rights of the person concerned, for the simple reason that this person will never have enjoyed these rights in the first place. [18] Be that as it may, once it is agreed that the Member States' rules on acquisition of nationality come under the scrutiny of Union law, it would be somewhat illogical to distinguish between conferral and refusal of nationality, since the very same rules will embody the criteria that determine both.

In conclusion, it arguably follows from the *Rottmann* judgment that the Member States' competence regarding both acquisition and loss of nationality falls within the scope of Union law to the extent that it has an impact on the status of Union citizenship.[19] This broad interpretation of the scope of Union law finds some support in the wording of the *Rottmann* judgment. The famous *Micheletti* dictum, which is cited by the Court, refers to the conditions for the *acquisition and loss* of nationality. Besides, the Court in *Rottmann* held that the principles announced in the judgment apply to both the Member State of naturalisation and the Member State of the original nationality.[20] One could deduce from this that a possible refusal of the Austrian authorities to grant or revive Mr. Rottmann's Austrian nationality will only be valid if it is in accordance with fundamental principles of Union law. Future case law will have to clarify this point.

---

[17] Except, of course, where the person concerned already possessed the nationality of a Member State.

[18] This accords with the reasoning followed in Case C-192/99 *Kaur* [2001] ECR I-1237. In that case, the Court considered that the non-conferral of Union citizenship on the applicant was valid because there was no question of any deprivation of rights under Union law since those rights had never arisen in the first place.

[19] See in that sense Davies (n 16).

[20] Case C-135/08 *Rottmann* [2010] ECR I-1449, paras 60-64.

Limitations Deriving from Union Law

In *Rottmann* the Court for the very first time gave concrete guidance on the significance and scope of its *Micheletti* dictum, as far as the conditions for loss of Member State nationality are concerned. The Court explained that such conditions have to be in accordance with the principle of proportionality, which requires a delicate balancing act between the interests of the Member State and those of the individual concerned.[21] The scope and effect of this limitation will depend on how stringently it is applied, a task which pertains first and foremost to the national courts of the Member States. It is submitted that it is probably only in extreme cases, i.e. where the interests of the individual *manifestly* outweigh those of the Member State concerned, that the principle of proportionality can be considered to be violated. Such would seem necessary in order to safeguard the Member States' principled competence in the field of nationality. Safeguarding that competence is arguably also necessary to protect the national identities of the Member States,[22] given that nationality is without any doubt one of the elements central to that identity. In any event, a Member State is not obliged to refrain from withdrawing its nationality merely because the person concerned has not recovered the nationality of his Member State of origin. At the same time, the principle of proportionality may require the person concerned to be afforded a reasonable period of time in order to try to recover the nationality of his Member State of origin. [23]

In *Rottmann* the Court only explicitly mentioned the principle of proportionality as a limitation deriving from Union law. However, other general principles of Union law could equally serve as limitations to the Member States' competence regarding nationality, as was observed by AG Poiares Maduro in his Opinion in the case.[24] In particular, the duty to respect fundamental rights[25], the principle of legitimate expectations[26] and

[21] The principle of proportionality is a general principle of Union law which also figures in art 52(1) of the Charter of fundamental rights. See the discussion in Koen Lenaerts and Piet Van Nuffel (Robert Bray and Nathan Cambien (eds)), *European Union Law* (3rd edn, Sweet & Maxwell 2011) 141ff.

[22] See art 4(2) TEU, which provides that the 'Union shall respect the equality of Member States before the Treaties as well as their national identities'. This provision was explicitly relied upon by the Court in order to justify a rather broad construction of Member State competence (see Case C-208/09 *Sayn-Wittgenstein* (ECJ, 22 December 2010), para 92).

[23] Case C-135/08 *Rottmann* [2010] ECR I-1449, paras 57-58.

[24] Case C-135/08 *Rottmann* [2010] ECR I-1449, Opinion of AG Poiares Maduro, paras 29-32.

[25] See art 6(3) TEU and art 51(1) of the Charter of Fundamental Rights. For a discussion, see Lenaerts and Van Nuffel (n 21), 821ff.

the freedom of movement and residence[27] could act as limitations in a way similar to the principle of proportionality. The principle of legitimate expectations and the duty to respect fundamental rights could even be said to 'feed' the principle of proportionality in the sense that a measure concerning nationality will be more likely to be disproportionate if it infringes one of them. The freedom of movement and residence, for its part, could be violated if a Member State's nationality law were to provide that nationals of that Member State would lose their nationality after having lived in another Member State during a certain period of time.[28] It will be for the Court to clarify the precise scope and meaning of these principles in this context.

Another principle which may be very important in this context is the principle of sincere cooperation.[29] That principle might require the Member States to take account of each other's nationality rules and the combined effect they may have for an individual in particular circumstances.[30] In this connection, it must be remarked that the referring court in the *Rottmann* case has in the meantime ruled that the German withdrawal of nationality was in accordance with the principle of proportionality and therefore valid.[31] One may wonder whether the Austrian authorities are under an obligation to take into account that Mr. Rottmann has now definitively lost his German nationality and are, on that ground, obliged to revive his Austrian nationality and his Union citizenship.[32] Sincere cooperation in this sense would enable Germany to apply the provisions of its nationality law while avoiding the definitive loss of Mr. Rottmann's Union citizenship[33] and, as such, be apt to further the

---

[26] There is a large body of case law mentioning the principle of legitimate expectations as a general principle of Union law. See eg Joined Cases C-181/04 to C-183/04, *Elmeka*, [2006] ECR I-8167, para 31. For a discussion, see Takis Tridimas, *The General Principles of EU Law* (2nd edn, OUP 2006) 251ff.

[27] See art 21(1) TFEU. For a discussion, see Lenaerts and Van Nuffel (n 21), 184-189.

[28] Gerard-René De Groot, 'The Relationship between the Nationality of the Member States of the European Union and European Citizenship' in Massimo La Torre (ed), *European Citizenship: an Institutional Challenge* (Kluwer Law International 1998) 136-139.

[29] See art 4(3) TEU. For a discussion, see Lenaerts and Van Nuffel (n 21), 147ff.

[30] See in this connection Case C-165/91 *van Munster* [1994] ECR I-4661, in which the Court held that the principle of sincere cooperation may require a Member State, when applying its legislation, to take into account the legislative provisions of another Member State.

[31] BVerwG 5 C 12.10., Judgment of 11 November 2010.

[32] The BundesverwaltungsGericht in fact explicitly made a suggestion in this sense (Case C-135/08 *Rottmann* [2010] ECR I-1449, para 34).

[33] Somewhat ironically, it would seem that better coordination between the Austrian

aims of the provisions on Union citizenship.[34]

3. *Consequences for the Member States' Immigration Laws*

The foregoing makes it clear that it can no longer be doubted that the nationality rules of the Member States have to be in accordance with a number of fundamental principles of Union law. This requirement evidently has consequences for the immigration laws and policies of the Member States, since the criteria for granting nationality to third country nationals now appear to fall within the scope of Union law. The precise scope of the requirement is at present, however, far from clear. Do the principle of proportionality and the principle of legitimate expectations require, for instance, that a Member State grant its nationality to third country nationals long time resident on its territory? And does the Commission have the power to bring an infringement action against a Member State whose criteria for the acquisition of nationality appear contrary to certain fundamental rights? An oft-discussed case in this connection is the nationality legislation of Estonia and Latvia, which makes it very hard for Russian-speaking minorities to acquire the nationalities of these countries.[35] The Commission has in the past expressed its concern over this situation, but it has never taken concrete action.[36] The Union institutions have so far adopted a low profile in nationality matters, given the traditional view that Union law had no say in these matters. The increasing importance of Union citizenship and the bold case law of the ECJ just discussed may lead to a more proactive approach on their part in the near future.

At the same time it must be emphasised that the *Rottmann* judgment in no way changes the fact that the Member States remain exclusively competent to adopt the rules on acquisition and loss of nationality. The Court in *Rottmann* only confirmed that this competence has to be exercised in accordance with Union law as far as situations falling within

---

and German authorities would have prevented the possibility of Mr. Rottmann losing his Union citizenship in the first place. If the Austrian authorities had been quicker to inform the German authorities about the pending criminal proceedings against Mr. Rottmann, the latter would presumably never have acquired the German nationality at all.

[34] AG Ruiz-Jarabo Colomer has remarked that 'Citizenship of the Union...must at least guarantee that it is possible to change nationality within the European Union without suffering any legal disadvantage' (Case C-386/02 *Baldinger* [2004] ECR I-8411, Opinion of AG Ruiz-Jarabo Colomer, para 47).

[35] See the discussion in Annelies Lottmann, 'No Direction Home: Nationalism and Statelessness in the Baltics' (2008) 43 Tex. Int'l L.J. 503-520.

[36] See Fifth Report from the European Commission of 15 February 2008 on Citizenship of the Union, COM(2008)85 final, 2.

the scope of Union law are concerned. This holding has nothing extraordinary in itself. The same duty to comply with Union law applies in other fields falling outside the Union's regulatory competence, such as *inter alia*, criminal legislation,[37] direct taxation,[38] rules governing a person's name,[39] and the organisation of social security schemes[40]. The Court's reasoning on the scope of Union law, by contrast, was very innovative. This will be discussed in more detail in the following point.

III. INTER-STATE MOVEMENT

1. *Traditional Approach*

It is trite law that Union law is only applicable to situations which fall within the scope of Union law. Traditionally, it has consistently been held that the situation of a Union citizen falls within the scope of Union law only where a link with two or more Member States –often referred to as a 'cross-border dimension' or 'inter-State element'– is present.[41] This link is most commonly provided by the fact that a Union citizen has exercised his right to free movement by moving from his home Member State to another Member State. In other words, movement between two Member States allows a Union citizen to bring his situation within the scope of Union law. This entitles the citizen concerned to invoke the right to equal treatment in the host Member State. Conversely, the home Member State is precluded from treating a national less favourably on ground of the fact that he has exercised his right to free movement.[42] However, the ECJ has been prepared to give a lenient interpretation to the inter-State element. Accordingly, movement between Member States was not always required by the Court. In *Zhu and Chen*, for instance, the fact that a Union citizen resided in a Member State other than her Member State of nationality sufficed to bring her situation within the scope of Union law.[43] In *Schempp*, the Court even considered that Union law was applicable in a situation in which not the Union citizen himself but his spouse had exercised her right to free movement.[44]

In all the situations mentioned, the Court considered a sufficient inter-State element

[37] *See* eg Case 186/87 *Cowan* [1989] ECR 195, para 19.
[38] See eg Case C-520/04 *Turpeinen* [2006] ECR I-10685, para 11.
[39] *See* eg Case C-353/06 *Grunkin and Paul* [2008] ECR I-7639, para 16.
[40] *See* eg Case C-135/99 *Elsen* [2000] ECR I-1049, para 33.
[41] For a discussion, see Spaventa, 'Seeing the Wood Despite the Trees? On the Scope of Union Citizenship and its Constitutional Effects' (2008) 45 CML Rev. 13-45.
[42] See eg Case C-192/05 *Tas-Hagen and Tas* [2006] ECR I-10451.
[43] Case C-200/02 *Zhu and Chen* [2004] ECR I-9925.
[44] Case C-403/03 *Schempp* [2005] ECR I-6421.

to be present because the applicant in each case could point at a link with two or more specific Member States. Cases where such a link is not present, by contrast, are traditionally considered to be 'purely internal' situations, in which no reliance on Union law is possible. Consequently, the traditional approach followed in the case law gives rise to instances of 'reverse discrimination', i.e. Union citizens who find themselves in a purely internal situation being treated less favourably than Union citizens who can demonstrate a sufficient link with Union law.[45] The reason is that Union citizens in a purely internal situation cannot rely on the rights conferred by Union free movement law, but only on the possibly less favourable rights conferred by the national law of their Member State of residence. Instances of reverse discrimination do not infringe the Union principle of non-discrimination because the latter is not applicable to purely internal situations.

2. *Developments regarding purely internal situations*

The traditional approach towards the required link with Union law has been fiercely criticised, in particular because it entails the possibility for reverse discrimination. It is sometimes argued that such is incompatible with the concept of the internal market as an 'area without internal frontiers'[46] because in a true internal market the crossing of a border between Member States should not be a relevant distinguishing factor for the application of Union law.[47] More broadly, the traditional approach can be said to be contrary, for the same reason, to the idea of the Union as an 'area of freedom, security and justice without internal frontiers, in which the free movement of persons is ensured'.[48] Besides, the traditional approach is sometimes said to be at odds with the provisions on Union citizenship.[49] In this connection it is argued that the distinction drawn in

[45]For a critical analysis of the doctrine, see, *inter alia*, Alina Tryfonidou, *Reverse Discrimination in EC Law* (Kluwer Law International 2009) 271 pp.; Niamh Nic Shuibhne, 'Free movement of persons and the wholly internal rule: time to move on?' (2002) 39 CML Rev. 731-771; Miguel Poiares Maduro, 'The Scope of European Remedies: The Case of Reverse Discrimination and Purely Internal Situations' in Claire Kilpatrick, Tonia Novitz and Paul Skidmore (eds), *The Future of Remedies in Europe* (Hart Publishing 2000) 117-140.

[46] See art 26(2) TFEU.

[47] This idea was cogently put forward, *inter alia*, in Hans Ulrich Jessurun d'Oliveira, 'Is Reverse Discrimination Still Permissible Under the Single European Act?' in Th. M. De Boer (ed), *Forty Years On: The Evolution of Postwar Private International Law in Europe* (Kluwer 1990) 71-86. See also Joined Cases 80/85 and 159/85 *Edah* [1986] ECR 3359, Opinion of AG Mischo.

[48] See art 3(2) TEU.

[49] See *inter alia*, Alina Tryfonidou, 'Reverse Discrimination in Purely Internal Situations: An Incongruity in a Citizens' Europe' (2008) 35 L.I.E.I. 43-67; Nic Shuibhne (n 45), 731-771. See also Francis G. Jacobs, 'Citizenship of the European Union - A Legal Analysis' (2007) 13 E.L.J. 591-610.

the case law between Union citizens who can demonstrate an even tenuous inter-State element and those who cannot is arbitrary and that Union citizenship should, as the most fundamental status of nationals of the Member States, embody a guarantee to equal treatment of Union citizens regardless of any further link with Union law. Accordingly, in the most extreme version of this argument, all instances of reverse discrimination of Union citizens should be held in violation of Union law.

Despite these longstanding criticisms, the ECJ has consistently repeated the 'wholly internal rule' and has refused to apply Union law to situations which do not present a link with two or more specific Member States. The Court's position is grounded on the need to respect the division of competences between the Union and the Member States. Union law has a limited scope of application and cannot be relied on, therefore, in situations that fall outside this scope. However, it does not necessarily follow that Union law, and the provisions on the free movement of Union citizens in particular, do not apply in situations lacking a link with two or more specific Member States.[50] It would be possible to accept a more abstract link with the Union Legal order as a sufficient link with Union law. A number of Advocate-Generals, most notably Advocate-General Sharpston, have made suggestions which go in that direction.[51] In a number of very recent cases, the Court seems to have adopted for the first time a similar reasoning, be it in a highly nuanced form and restricted to a limited set of circumstances. In these cases, the Court appears to have accepted Union citizenship in itself as a sufficient link with Union law, thereby applying Union law in situations hitherto considered to be purely internal.

a. Member State nationality

The first revolutionary case concerning the link required with Union law is the *Rottmann* case discussed above. In its judgment, the Court did not examine whether the traditional requirement of an inter-State element was satisfied, even though such appeared to be the case in the circumstances before the Court. As AG Poiares Maduro explained in his Opinion to the

[50] It must be remarked that, in any event, a number of provisions on Union citizenship apply regardless of such a link. This is the case for the right to petition the European Parliament, the right to apply to the Ombudsman and the right to write to any of the institutions or bodies of the Union in an official language and have an answer in the same language (see art 24 TFEU).

[51] See Case C-212/06 *Government of the French Community and Walloon Government v Flemish Government* [2008] ECR I-1683, Opinion of AG Sharpston, paras 112-157; Case C-34/09 *Ruiz Zambrano*, Opinion of AG Sharpston, paras 67-122. For an earlier example, see Case C-214/94 *Boukhalfa* [1996] ECR I-2253, Opinion of AG Léger, para 63.

case, Mr. Rottmann had exercised his right to free movement by moving from Austria to Germany and this exercise, indirectly, gave rise to the disadvantage suffered, namely the loss of the status of Union citizen and the attached rights.[52] The Court decided to take a different approach however, holding that the situation fell within the scope of Union law 'by reason of its nature and its consequences'. Accordingly, the Court accepted Member State nationality and the possible loss thereof, given the inextricable links with Union citizenship, as sufficient in itself to consider the withdrawal of nationality as falling within the scope of Union law. Any further connection with Union law appeared unnecessary for the situation to fall within the scope of Union law.

This approach is very innovative. Although the Court confirmed that Union law only applies to situations presenting a link with Union law, it conceptualised this link in a different way. It did not require a link with two specific Member States, but rather a more abstract link with the Union legal order. This abstract link was offered by Union citizenship. Indeed, each national measure affecting the Union citizenship status of an individual will automatically affect his most fundamental status under Union law and entail potential consequences for all the Member States.

It could be wondered to what extent the outcome of the judgment was predicated on the facts of the case. The fact should not be overlooked that in *Rottmann* the Court was confronted with a situation in which, due to the lack of coordination between the nationality laws of two Member States, a person risked becoming stateless and losing his Union citizenship for having committed an offence which was in many ways not extraordinary. It could be suggested that the Court was principally concerned with avoiding these negative consequences from happening all too readily, and that the Court's reasoning should not, therefore, be extrapolated to cases with a different set of circumstances. One could speculate that the fact that the nationality legislation of two specific Member States was at stake induced the Court to find that Union law was applicable.[53] Besides, it is clear that *Rottmann* was a dispute about nationality rules. Nationality rules are a particular set of national rules because they directly regulate the access to Union citizenship and therefore determine the applicability of a significant

[52] Case C-135/08 *Rottmann* [2010] ECR I-1449, Opinion of AG Poiares Maduro, paras 11-13.

[53] This would appear from a literal reading of para 42 of the *Rottmann* judgment and, more in particular, of the phrase 'after he has lost the nationality of another Member State that he originally possessed'. See the discussion in Gerard-René De Groot, 'Invloed van het Unierecht op het nationaliteitsrecht van de Lidstaten: Overwegingen over de Janko Rottmann-beslissing van het Europees Hof van Justitie' (2010) Asiel & Migratierecht 295-296.

part of the Union *acquis*. For that reason, one could argue that a dispute concerning nationality rules will by its very nature have a more significant link with the Union legal order than disputes concerning other sets of national rules.

In my view, these observations relating to the specific circumstances of the *Rottmann* case are rather beside the point. It is clear from the Court's reasoning that the crucial element in deciding that the situation fell within the ambit of Union law was the fact that the national measure threatened to cause the loss of the applicant's Union citizenship and the enjoyment of the attached rights. Hence, the Court's reasoning should be held to apply more broadly, even where only the legislation of one Member State is at stake. Moreover, it can apply to national rules outside the field of Member State nationality. This is clearly illustrated by the *Ruiz Zambrano* judgment.[54]

b. Genuine enjoyment of Union Citizenship Rights

Mr. Ruiz Zambrano was a Colombian national who came to Belgium together with his Colombian spouse and their first child. Although his request for asylum was rejected by the Belgian authorities, he nevertheless remained in the country and even managed to become gainfully employed. He did not, however, satisfy the conditions under Belgian law for obtaining a residence permit or a work permit. The question to be answered by the ECJ was whether Mr. Ruiz Zambrano could derive a right of residence in Belgium from Union law and whether Union law would exempt him from the obligation to hold a work permit. The crucial element in this regard was that, during his stay in Belgium, Mr. Ruiz Zambrano's spouse gave birth to a second and third child, who acquired the Belgian nationality on grounds of their birth in Belgium.[55] In *Zhu and Chen* the Court had ruled that a young Union citizen was entitled to be accompanied in the host Member State by the parent who is his or her primary carer.[56] It seemed problematic, however, to apply an analogous reasoning to the facts of the *Ruiz Zambrano* case since, in contrast with baby Chen, the children of Mr. Ruiz Zambrano had never resided in a Member State other than that of their nationality. For that reason, it seemed that the situation of Mr. Ruiz Zambrano was a purely internal one, in which no reliance on Union law

[54] Case C-34/09 *Ruiz Zambrano* (ECJ, 8 March 2011), with case notes by Kay Hailbronner and Daniel Thym in (2011) 48 CML Rev. 1253-1270; Janek T. Nowak in (2011) Colum. J. Eur. L. 673-704.

[55] Pursuant to art 10(1) of the Belgian Nationality Code, in the version applicable at that time, children born in Belgium acquired the Belgian nationality if they would otherwise be stateless.

[56] Case C-200/02 *Zhu and Chen* [2004] ECR I-9925.

was possible. This point of view was defended before the ECJ by no less than eight Member States and by the Commission.

The ECJ disagreed and held that Union law was applicable to the circumstances of the case. In a remarkably short judgment, the Court pointed out that the children of Mr. Ruiz Zambrano were undeniably Union citizens and that Union citizenship was, according to settled case law, the fundamental status of nationals of the Member States.[57] Referring to paragraph 42 of the *Rottmann* judgment,[58] the Court stated that Article 20 TFEU precludes national measures which have the effect of depriving Union citizens of the 'genuine enjoyment of the substance of the rights conferred by virtue of their status as citizens of the Union'.[59] The Court held that the refusal of a residence permit and of a work permit to a person in a situation like Mr. Ruiz Zambrano had precisely this effect. The reason was that a refusal of a residence permit would require Ruiz Zambrano's children to accompany their parents to a third country. Similarly, the refusal of a work permit would entail the risk that Ruiz Zambrano would not have sufficient resources to provide for himself and his family, which would also result in the children having to leave the territory of the Union. In both circumstances, the children would, as a result, be unable to exercise the 'substance of the rights conferred on them by virtue of their status as Union citizens'.[60] That outcome would be at variance with Article 20 TFEU.

c. But not in all circumstances?

Although the *Ruiz Zambrano* judgment was remarkably short and lacking in elaborate reasoning,[61] it did appear to confirm the landslide in the Court's case law which was initiated with the *Rottmann* judgment. Indeed, the Court found Union law to be applicable despite the fact that the traditional requirement of an inter-State element was not satisfied. The Court accepted the fact that the national measure deprived a Union citizen of the genuine enjoyment of his citizenship rights as a sufficient connection with Union law, regardless of any further such connection. The question which arose immediately after the judgment was how broadly the new approach of the Court will apply. The precise scope of the judgment

[57] Case C-34/09 *Ruiz Zambrano* (ECJ, 8 March 2011), paras 40-41.
[58] Case C-135/08 *Rottmann* [2010] ECR I-1449, para 42 (in which the Court held that the withdrawal of the applicant's nationality in the circumstances of the case fell 'by reason of its nature and its consequences' within the ambit of Union law).
[59] Case C-34/09 *Ruiz Zambrano* (ECJ, 8 March 2011), para 42.
[60] Case C-34/09 *Ruiz Zambrano* (ECJ, 8 March 2011), paras 43-44.
[61] See Niamh Nic Shuibhne, 'Seven questions for seven paragraphs' (2011) 36 EL Rev. 162.

was impossible to infer from its succinct wording. Some clarity was restored by the subsequent *McCarthy* case and *Dereci and Others* cases, in which the Court appears to have given a rather narrow interpretation to the reasoning followed in *Ruiz Zambrano*.

The applicant in the *McCarthy* case, Mrs. McCarthy, held both the Irish and the UK nationality, but had lived her whole life in the UK. In 2002, she married a Jamaican national, who was not, however, entitled to reside in the UK in accordance with the British immigration rules. Relying on her Irish nationality, Mrs. McCarthy and her husband argued that they were entitled to residence on the basis of Union law, namely in their capacity of Union citizen and husband of a Union citizen, respectively. Mrs. McCarthy had never exercised her right to free movement and, consequently, her situation seemed to amount to a purely internal situation. Yet, such was far from certain after the Court's judgment in *Ruiz Zambrano*. Moreover, the question arose whether the fact that Mrs. McCarthy possessed the nationality of another Member State than the Member State in which she resided could provide a sufficient link with Union law. Some earlier cases, the *Garcia Avello* case[62] in particular, appeared to confirm that the possession of the nationality of two Member States was sufficient in order to enable a Union citizen to invoke Union law.

Contrary to what some commentators had expected in view of the recent *Ruiz Zambrano* judgment, the Court ruled that Union law was not applicable in the circumstances of the case. According to the Court, Mrs. McCarthy could not invoke Article 21 TFEU because the contested national measure did not have the effect of depriving her of the genuine enjoyment of the substance of her citizenship rights or of impeding the exercise of her right of free movement and residence.[63] The Court explicitly distinguished the circumstances of the *McCarthy* case from those at stake in *Ruiz Zambrano.* It held that, in contrast to the case of *Ruiz Zambrano*, the contested national measure did not have the effect of obliging Mrs McCarthy to leave the territory of the European Union. The fact that Mrs. McCarthy possessed the nationality of two Member States could not change anything with regard to these findings, as it did not trigger the application of national measures depriving her of the genuine enjoyment of the substance of her citizenship rights or impeding the exercise of her right of free movement and residence.[64]

[62] Case C-148/02 *Garcia Avello* [2003] ECR I-11613.
[63] Case C-434/09 *McCarthy* (ECJ, 5 May 2011), paras 44-56.
[64] Case C-434/09 *McCarthy* (ECJ, 5 May 2011), para 54.

The *Dereci and Others* case provided the ECJ with an ideal opportunity to further clarify the scope of its holdings in *Ruiz Zambrano* and *McCarthy*.[65] The reference of the Austrian V*erwaltungsgerichtshof* in fact concerned five cases in which a third country family member[66] of a static adult Austrian national were refused a right of residence in Austria. The referring court wanted to know, essentially, whether these refusal decisions were precluded under Article 20 TFEU. This required the ECJ to clarify whether such decisions were to be considered as having the effect of depriving the EU citizens concerned of the genuine enjoyment of the substance of their citizenship rights. The ECJ firmly stated that this criterion is only satisfied in situations in which the EU citizen has, in fact, "to leave not only the territory of the Member State of which he is a national but also the territory of the EU as a whole".[67] It emphasised that this criterion would only under exceptional circumstances preclude a refusal of a right of residence. In this connection, the Court explained that the mere fact that it might appear desirable to an EU citizen, for economic reasons or in order to keep his family together, for his third country family members to be able to reside with him in the territory of the EU, is not sufficient in itself to support the view that the EU citizen will be forced to leave EU territory if such a right is not granted.[68]

The picture resulting from the judgments just discussed is rather nuanced. In *Ruiz Zambrano* the Court departed from its traditional Union citizenship case law, which was centred on the presence or absence of an inter-State element. As a consequence, a large number of situations could seem to fall henceforth within the scope of Union law which would previously have fallen outside that scope. That would have drastic consequences for the vertical division of competences between the Union and the Member States. On a closer look, however, it seems that the judgment does not entail such wide consequences. As the Court clarified in *McCarthy* and *Dereci and Others*, it is willing to apply Union law only where the 'substance of' citizenship rights is at stake. In such circumstances an inter-State element will no longer be required. In essence, the Court is merely drawing the consequences from its *Rottmann* judgment. If a measure taking away one's Union citizen status falls within the scope of Union law in the absence of a cross-border dimension, the same should be the case for a national measure completely rendering it impossible for someone to exercise the rights attached to that status. Put

[65] Case C-256/11 *Dereci and Others* (ECJ, 15 November 2011).
[66] Namely the spouse of an EU citizen in three cases and the adult children of an EU citizen in the two other cases.
[67] Case C-256/11 *Dereci and Others* (ECJ, 15 November 2011), para. 66.
[68] Case C-256/11 *Dereci and Others* (ECJ, 15 November 2011), paras 67-68.

differently, national measures which *de iure* or *de facto* annihilate one's Union citizenship should be treated equally and be held to fall within the scope of Union law even in the absence of a cross-border dimension.[69]

It appears from these cases that the Court accepts that a refusal of a right of residence to the parent of a minor Union citizen makes it impossible for that citizen to exercise the substance of his citizenship rights. The impossibility for Mrs. McCarthy to be joined by her husband, by contrast, did not have this consequence because it did not oblige her to leave the territory of the Union.[70] The same was true, presumably, for the applicants in *Dereci and Others*.[71] Consequently, the Court appears to limit its extensive interpretation of the scope of Union law to children who face the impossibility to be joined by their parent(s).

Two observations may explain the Court's narrow interpretation. First, it must be pointed out that Union law has traditionally paid special attention to the position of young children.[72] Already in previous cases, the impossibility for children to reside independently in a Member State appears to have inspired the Court to recognise for their family members more extensive rights than those enjoyed by family members of other Union citizens.[73] Second, the Court's holding in *Ruiz Zambrano* is arguably implicitly based on considerations relating to the need to respect fundamental rights, the right to respect for family life in particular. The relevant case law of the ECtHR is also more restrictive as far as minors are concerned.[74]

3. *Consequences for the Member States' Immigration Laws*

The foregoing makes it clear that the classic inter-State requirement will

[69] Nathan Cambien, 'Case Note: Case C-34/09 *Ruiz Zambrano*' (2011) Sociaal-economische wetgeving 410-43.
[70] Case C-434/09 *McCarthy* (ECJ, 5 May 2011), para 50.
[71] Somewhat curiously the Court in *Dereci and Others* did not make a final assessment of compliance with Article 20 TFEU, explicitly leaving this to the referring court (Case C-256/11 *Dereci and Others* (ECJ, 15 November 2011), para. 74). Yet the Court's emphasis on the limited applicability of Article 20 TFEU vis-à-vis static EU citizens and on the fact that *Ruiz Zambrano* concerned the right of residence of a third country national with dependent minor children clearly indicate that it was of the opinion that the applicants' argument under EU law would not succeed.
[72] See eg art 12 of Regulation 1612/68 and, more recently, art 24 of the Charter of Fundamental Rights of the EU.
[73] See, most recently, Case C-310/08 *Ibrahim* [2010] ECR I-1065 and Case C-480/08 *Teixeira* [2010] ECR I-1107 (see also the discussion under IV.B., *infra*).
[74] Hailbronner and Thym (n 54), 1268, referring to *Maslov vs. Austria* ECHR 2008 1638-03, paras 70-72.

no longer apply in some cases involving Union citizenship. Union citizens who are confronted with a national measure *de iure* or *de facto* taking away the genuine enjoyment of their Union citizenship rights will be able to invoke Union law against their home Member State, even in the absence of a link with any other Member State. This development in the case law can have significant consequences for the immigration laws of the Member States.

On the one hand, Member States will have to accord some categories of their static nationals exactly the same rights regarding family reunification as are conferred by Union law on moving Union citizens. At present, this duty applies with certainty to young Union citizens. They should be accorded the right to reside in their home Member State together with their parent primary carer. Consequently, Member States which deny this right to static nationals will have to change their legislation. This is important in those Member States which accord their static nationals a right to family reunifications under more burdensome conditions than those applicable to other Union citizens.[75] In Belgium, for instance, a recent legislative proposal introduces more burdensome rights for static nationals when compared to other Union citizens, but makes an exception for minor children and their parents.[76] However, many uncertainties remain concerning the exact scope of the new case law. First of all, it is not exactly clear precisely what categories of static nationals have to be accorded the said residence rights. For instance, must adult Union citizens who are dependent on a primary carer be equalled with young dependent Union citizens in this connection?[77] Besides, it is not clear whether Member States may, under the circumstances mentioned, refuse to accord a right to a Union citizen and their primary carer by relying on a legitimate Member State interest[78] and whether they may impose conditions relating to self-sufficiency[79]. These uncertainties will hopefully be settled by future cases.[80]

---

[75] See the overview in Anne Walter, *Reverse Discrimination and Family Reunification* (Wolf Legal Publishers 2008) 78 pp.

[76] See the discussion in Nathan Cambien, 'Mogen statische Unieburgers worden gediscrimineerd? Enkele beschouwingen bij *Ruiz Zambrano* en *McCarthy*' (2011) Tijdschrift voor Vreemdelingenrecht 242-253.

[77] One can think, for instance, of disabled persons who need the presence of a primary carer (see in this context: Case C-303/06 *Coleman* [2008] ECR I-05603).

[78] The Court did not explicitly consider this point in *Ruiz Zambrano*, *McCarthy* or *Dereci and Others*, but one could argue that the parallel drawn by the Court with the *Rottmann* judgment leaves open the possibility that a refusal to accord a right of residence might in certain circumstances be justified.

[79] See the discussion under IV, *infra*

[80] A substantial number of references have already been made to the Court, asking for further clarification of the *Ruiz Zambrano* and *McCarthy* judgments.

On the other hand, this development will likely influence the criteria for the acquisition of nationality in the Member States. As was pointed out higher, it arguably follows from the *Rottmann* judgment that these criteria have to comply with Union law, even in situations in which no inter-State element is present. Moreover, the generous interpretation by the Court of the rights accruing to certain categories of static Union citizens and their third country family members, creates an incentive to restrict the criteria for the acquisition of nationality. It appears from the interventions of a large number of Member States in high profile cases before the Court that most Member States resist a wide interpretation of the Union citizenship provisions because they fear that such will render it impossible for them to control immigration, resulting in significant and uncontrollable financial burdens.[81] After the *Ruiz Zambrano* judgment, third country nationals who manage to obtain the nationality of a Member State for their child can claim a right of residence in that State as the primary carers of that child. As such, the judgments could result in an enormous increase in claims for residence permits. In order to prevent this scenario from happening too easily, Member States will probably restrict the possibilities for acquiring their nationality, thereby restricting the possibilities for Union citizenship based residence claims. A case in point is the Belgian nationality legislation, which was restricted in the context of a number of claims similar to the one in *Ruiz Zambrano*.[82] At the same time, it must be remarked that such restrictions are only valid as long as they do not contravene certain fundamental principles of Union law, such as the principle of legitimate expectations.

The potentially significant consequences of the wide interpretation of the Union citizenship provisions for the immigration laws and policies of the Member States is also cogently illustrated by the cases discussed under the next title, relating to the self-sufficiency requirement.

IV. REQUIREMENT OF SELF-SUFFICIENCY

See, for instance, pending cases C-356/11 *O and S* and C-357/11 *L*, lodged on 7 July 2011.

[81] See eg Case C-127/08 *Metock and Others* [2008] ECR I-6241, paras 71-72.

[82] Whereas traditionally the Belgian Nationality Code provided that a child born in Belgium acquired the Belgian nationality if it would otherwise be stateless, after an amendment in 2006 such is no longer the case 'if, by appropriate administrative action instituted with the diplomatic or consular authorities of the country of nationality of the child's parent(s), the child's legal representative(s) can obtain a different nationality for it'.

1. *Traditional Residence Requirements*

As was remarked higher, the conditions surrounding the right to free movement and residence are now comprehensively laid down in Directive 2004/38, which repeals earlier directives governing the free movement and residence rights of specific categories of Union citizens and their family members.[83] The Directive also replaces a number of provisions of Regulation 1612/68 on the free movement of workers.[84] It must be emphasised, however, that the latter Regulation was not repealed. Some of its key provisions remain in force. This is the case, for instance, for Article 12, which grants children of a migrant worker the right to access to education in the host Member State under the same conditions as nationals of that State. That provision has been interpreted by the ECJ as granting a right of residence to school-going children of migrant workers which is independent from the right of residence of their parents.[85] I will come back to this point below.

The central conditions stated in Directive 2004/38 are those relating to the financial situation of the Union citizen. Union citizens are only entitled to reside in the host Member State for more than three months if they are either economically active or have sufficient resources for themselves and their family members not to become a burden on the social assistance system of the host Member State and have comprehensive sickness insurance cover in that State.[86] The underlying idea is that Union citizens can only move to another Member State if they are financially independent, in order to avoid moving Union citizens and their family members becoming a burden for the social assistance system of the host Member State. This condition of self-sufficiency can, like the possession of Member State nationality and the inter-State element, be labelled a classic element of the free movement of Union citizens. However, this element too has to be nuanced in view of recent case law of in which the ECJ has recognised a right of residence for certain categories of Union citizens, despite the fact that they did not at all fulfil the requirements regarding self-sufficiency.[87]

---

[83] See the discussion in Anastasia Iliopoulou, 'Le nouveau droit de séjour des citoyens de l'Union et des membres de leur famille: la directive 2004/38/CE' (2004) RDUE 523-557.

[84] Regulation (EEC) No 1612/68 of the Council of 15 October 1968 on freedom of movement for workers within the Community [1968] JO L257/2.

[85] *Inter alia* Joined Cases 389/87 and 390/87 *Echternach and Moritz* [1989] ECR 723; Case C-413/99 *Baumbast and R* [2002] ECR I-7091.

[86] See art 7 of Directive 2004/38. Students only have to 'assure' the host Member State that they have sufficient resources for themselves and their family members (see art 7(1)c).

[87] In earlier cases the Court had already adopted a restrictive interpretation of the

2. *The Ibrahim and Teixeira Cases and their Consequences*

a. The Cases

The facts of the *Ibrahim* and *Teixeira* cases are very similar.[88] The applicants in both cases entered the UK as the spouse of a Union migrant worker, together with their children. Consequently, both women separated from their husband and continued to live in the UK independently, together with their children. At some point in time, both women applied for housing assistance for themselves and for their children. Their application was rejected because, according to the competent UK authority, they were not entitled to reside in the UK under Union law.[89] This view was based on the fact that the applicants were not self-sufficient or covered by comprehensive sickness insurance and depended on social assistance to cover the living expenses of themselves and their children. In both cases, the applicants submitted, however, that they did derive a right of residence under Union law from the fact that they were the primary carer of school-going children.

The ECJ essentially confirmed its holding in *Baumbast and R* and held that the primary carer of school-going children was entitled to reside in the host Member State for the period of his or her children's education. It confirmed that school-going children of a (former) migrant worker derive an independent right of residence from Article 12 of Regulation 1612/68.[90] Furthermore, it explained that, precisely in order to guarantee the

---

possibilities for Member States to impose these requirements, *inter alia* by holding that they have to be interpreted in accordance with the principle of proportionality (Case C-413/99 *Baumbast and R* [2002] ECR I-7091, para 91).

[88] Case C-310/08 *Ibrahim* [2010] ECR I-1065; Case C-480/08 *Teixeira* [2010] ECR I-1107, with case notes by Matthew Starup and Peter Elsmore in (2010) 35 EL Rev. *571-1160; Charlotte* O'Brien in (2011) 48 CML Rev. 203-225. The most important factual difference between the two cases was that Ms. Ibrahim was a third country national, whereas Ms. Texeira was a Union citizen who had previously been employed in the UK. Furthermore, Teixeira's daughter was over 18 years old, whereas Ibrahim's children were young minors.

[89] It follows from the *Housing Act 1996* and the *Allocation of Housing and Homelessness (Eligibility) Regulations 2006* that a person is not eligible for housing assistance unless he has a right of residence in the United Kingdom conferred by Union law (Case C-310/08 *Ibrahim* [2010] ECR I-1065, para 14).

[90] Accordingly, the Court held that that right is not lost where the parents of the children concerned have meanwhile divorced, and that the fact that only one parent is a Union citizen and the fact that that parent has ceased to be a migrant worker in the host Member State are irrelevant in this regard (Case C-310/08 *Ibrahim* [2010] ECR I-1065, para 29 and Case C-480/08 *Teixeira* [2010] ECR I-1107, para 37, referring to Case C-413/99 *Baumbast and R* [2002] ECR I-7091, para 63).

effectiveness of this independent right of residence, residence rights must be extended to the primary carer of these children, without whom the latter cannot realistically exercise this right. The Court was prepared to go far in its protection of the *effet utile* of the residence rights of school-going children and their primary carer by holding, first, that these residence rights were not subject to the conditions regarding self-sufficiency. The Court held, moreover, that these residence rights cannot be made subject to a condition of age. Accordingly, the primary carer of a school-going child is entitled to reside in the host Member State even after that child reaches the age of majority for as long as the child continues to need his presence and care in order to be able to pursue and complete his or her education.

b. Analysis

The *Ibrahim* and *Texeira* cases confirm that school-going children of a (former) migrant worker and their primary carer derive a right of residence in the host Member State, even if they are not self-sufficient. This outcome is surprising in view of the fact that Directive 2004/38, which codifies the rules on the free movement of persons, mentions no such right.[91] One could have expected the Court to hold that the Directive, which to a large extent incorporates predating ECJ case law[92], has implicitly overruled *Baumbast and R*. Still, as the Court correctly pointed out, one cannot ignore the fact that the Directive did not repeal Article 12 of Regulation 1612/68, in contrast with Articles 10 and 11 of that Regulation. This probably means that the Union legislator did not intend to change the meaning and consequences of that provision. As the Court pointed out in *Ibrahim* and *Teixeira*, if Article 12 of Regulation 1612/68 could no longer be interpreted as conferring a right of residence on school-going children and their primary carer but only as conferring the right to equal treatment with regard to access to education, it would have become superfluous with the entry into force of Directive 2004/38, which lays down in its Article 24(1) a general right to equal treatment, which is applicable to access to education. Besides, one can agree with the Court that the aim of that Directive is *inter alia* to simplify and strengthen the right of free movement and residence of all Union citizens[93] and that it

[91] Except under the exceptional circumstances foreseen by art 12(3) of the Directive (see also text to n 104).

[92] This appears from a number of recitals in the preamble to the Directive (see *inter alia* recitals 9 and 27). See also Samantha Currie, 'EU Migrant Children, their Primary Carers and the European Court of Justice: Access to Education as a Precursor to Residence under Community Law' (2009) 16 Journal of Social Security Law 81.

[93] See recital 3 in the preamble to the Directive.

must not be interpreted therefore as a 'step back' as far as the rights of school-going children and their primary carer are concerned.

The Court based the residence rights for school-going children and their primary carer on Article 12 of Regulation 1612/68. Since that Article does not explicitly confer such rights, the question arises, again, exactly what categories of persons can invoke the said rights. The Court's case law gives a number of important clues.

c. Residence Rights for School-going Children

The first category of persons deriving a right of residence from Article 12 of Regulation 1612/68 are children of a migrant worker who have resided with him in the host Member State for a certain period of time. These children will normally derive their initial residence right in that Member State from their status of family member of a migrant worker. Once they start schooling in the host Member State, however, they acquire an independent right of residence for the duration of their education. In this regard it is not required that their parent had the status of migrant worker on the date on which the child started its education.[94] Moreover, the children's right of residence does not end when their parent stops working or when they no longer live together with this parent. Article 12 of Regulation 1612/68 applies to all types of education, including higher education and university education. Accordingly, school-going children continue to enjoy a right of residence when they attain the age of majority and for as long as their schooling lasts, even if they are not financially dependent on their parents. Consequently, their right of residence extends even further than the residence rights enjoyed by children in the host Member State in their capacity as family members of a Union citizen. The latter category only enjoys a right of residence if they are under 21 years old or dependent.[95]

d. Residence Rights for the Primary Carer

The second category deriving a right of residence from Article 12 of Regulation 1612/68 are primary carers of school-going children. Naturally, this right too only lasts as long as the child's schooling continues. Moreover, the Court has ruled that this right will normally end when the child reaches the age of majority, unless the child continues to need the presence and care of that parent in order to be able to pursue and

[94] Case C-480/08 *Teixeira* [2010] ECR I-1107, paras 71-75.

[95] See art 2(2) of Directive 2004/38. The dependency condition is interpreted by the Court as referring to financial dependency (see Case C-1/05 *Jia* [2007] ECR I-1, para 35 and the case law cited).

complete his or her education.[96]

The Court recognised the residence right just mentioned only for the *parent* who is the primary carer of his children. The Court's reasoning could presumably, however, also apply to other categories of primary carers, for instance where the primary carer of a child is not his parent but another family member or even a non-family member like a legal guardian. [97] Indeed, the very reason for which the Court recognised a right of residence on behalf of the primary carer is that such is necessary in order to guarantee the *effet utile* of the right to education of the children concerned. *Prima facie* it cannot be seen why this *effet utile* should not be guaranteed if the primary carer is not the parent of the child concerned.[98] A broad interpretation can further be based on the need to comply with the right to respect for family life[99]. Furthermore, the possibility cannot be excluded that a child has multiple primary carers, who should be accorded a right of residence on the basis of Article 12.[100]

An interesting question is whether the primary carer is entitled to equal treatment in the host Member State.[101] Directive 2004/38 confers the right to equal treatment on both Union citizens and their family members, with certain derogations for *inter alia* students (see Article 24). However, this right only applies to persons 'residing on the basis of this Directive in the territory of the host Member State' (see Article 24(1)). The residence right of the primary carer is not, however, based on the Directive, but on Regulation 1612/68. One could assume therefore that primary carers cannot invoke the Union principle of equal treatment in order to claim access to social assistance, for instance. However, that view is not tenable in the light of the reasoning followed by the Court in *Ibrahim* and *Teixeira*. The conferral of a right of residence on the primary carer without at the same time conferring a right to equal treatment would put in peril the *effet utile* of the residence rights of the children, as is clearly illustrated by the facts of these cases. In both cases the mother primary carer was jobless and

[96] Case C-480/08 *Teixeira* [2010] ECR I-1107, paras 84-87.
[97] Interesting to note is that the Court in some paragraphs refers to the *person* who is the primary carer of children (eg Case C-310/08 *Ibrahim* [2010] ECR I-1065, para 31).
[98] Annette Schrauwen, 'Zelfstandig verblijfsrecht van schoolgaande kinderen van werknemers en hun verzorgers: ontbreken van bestaansmiddelen niet relevant' (2010) Nederlands tijdschrift voor Europees recht 236.
[99] See in particular Case C-60/00 *Carpenter* [2002] ECR I-6279 (in which the Court recognised a right of residence for the stepparent who was the primary carer of children of a Union citizen).
[100] Starup and Elsmore mention the possibility of recognising a 'secondary carer' besides the 'primary carer' (Starup and Elsmore (n 88), at 584).
[101] See also Schrauwen (n 98), at 236-237.

fully dependent on welfare benefits. It should be clear that, in the absence of a right to claim equal access to social assistance, the mother could not realistically reside with her children in the UK. Besides, it must be remarked that legal residence in the host Member State in accordance with the national laws of that State may enable a person to invoke the general principle of equal treatment laid down in Article 18 TFEU.[102]

e. Children of Non-economically Active Persons and Their Primary Carer?

It can be wondered whether the residence rights enjoyed by school-going children and their primary carer only pertain to children of (former) migrant workers. Does the case law discussed also apply to school-going children of other categories of Union citizens –like self employed persons or, more importantly still, non-economically active Union citizens– and their primary carer? This question has no obvious answer.

On the one hand, a strict reading of the *Ibrahim* and *Teixeira* cases leads to the conclusion that the question must be answered in the negative. Since Article 12 of Regulation 1612/68 only applies to migrant workers and their families, it seems not possible for children of other categories of Union citizens and their primary carer to rely on this case law. On the other hand, the reasoning followed by the Court provides some support for a positive answer. As was pointed out above, the Court's main concern was to preserve the *effet utile* of the right of access to education for children of migrant workers. It should be clear that children of non-economically active Union citizens equally enjoy a right of access to education in the host Member State.[103] It could be argued that, once such children have obtained a right of residence in the host Member State and attend school there, they should similarly obtain an independent right of residence for themselves and for their primary carer which cannot be made subject to restrictive conditions such as those concerning self-sufficiency.

In fact, the Union legislator has (partially) confirmed this point of view in Article 12(3) of Directive 2004/38, to which the Court explicitly referred to support its reasoning in *Ibrahim* and *Teixeira*.[104] That Article states:

The Union citizen's departure from the host Member State or his/her death shall not entail loss of the right of residence of his/her children or of the parent who has actual custody of the children, irrespective of

[102] See Case C-456/02 *Trojani* [2004] ECR I-7573.
[103] See art 24 of Directive 2004/38, which also applies to access to education.
[104] See Case C-310/08 *Ibrahim* [2010] ECR I-1065, paras 57-58: Case C-480/08 *Teixeira* [2010] ECR I-1107, paras 68-69.

nationality, if the children reside in the host Member State and are enrolled at an educational establishment, for the purpose of studying there, until the completion of their studies.

The right laid down in Article 12(3) applies to school-going children of all categories of Union citizens covered by Directive 2004/38 – i.e. economically active and non-economically active Union citizens – and is not subject to the classic residence conditions regarding self-sufficiency.[105] However, Article 12(3) is only applicable in the event of death or departure of a Union citizen from the host Member State. It does not, on its face, apply in the case of a non-economically active Union citizen who continues to reside in the host Member State after no longer fulfilling the requisite conditions.[106] All the same, it could be argued that the non-application of the substance of Article 12(3) in such circumstances would undermine the aim pursued by that provision, namely safeguarding the right of access to education for school-going children of a Union citizen in the host Member State. One could argue, therefore, that the Court should adopt a wide interpretation of Article 12(3), going beyond its literal wording, and finding application in all circumstances where the Union citizen whose children attend an educational establishment in the host Member State loses his entitlement to residence in that State.[107] In all such circumstances, the right of residence in the host Member State for the children concerned would continue until they finish their education. The same would be true for their primary carer, at least until they reach the age of majority.

This reasoning is, of course, merely speculative. Strictly speaking, the *Ibrahim* and *Texeira* cases only apply to children of migrant workers. It is very well possible that the Court will refuse the wide interpretation suggested of Article 12(3) of Directive 2004/38 for it goes against the apparent will of the legislator, who limited Article 12(3) to cases of death or departure of the Union citizen concerned. Besides, the wider interpretation of Article 12(3) would take away much of the added value of Article 12 of Regulation 1612/68, a provision which was preserved by the Union legislator even after the adoption of Directive 2004/38. Still, as was pointed out, it cannot be fully excluded that in future cases the Court will enlarge its holding to children of other categories of Union citizens and their primary carer. The Court has in past cases already been prepared to

[105] This clearly ensues when art 12(3) is contrasted with arts 12(1) and (2) of the Directive.
[106] The most obvious example is that of a Union citizen who initially had sufficient resources and a comprehensive sickness insurance cover in the host Member State, but later lost one of these.
[107] See Starup and Elsmore (n 88), 583-584.

go past the strict wording of secondary Union law in order to guarantee the rights of Union citizens and their family members.[108]

3. *Consequences for the Member States' Immigration Laws*

While the *Ibrahim* and *Teixeira* cases are perhaps less discussed than the *Rottmann* and *Ruiz Zambrano* cases, their consequences for the immigration policies of the Member States could be more far-reaching. Indeed, vis-à-vis school-going children of migrant workers and their primary carer(s) Member States cannot impose the traditional requirements regarding self-sufficiency. Such persons have to be accorded a right of residence and equal access to social benefits. The resulting financial burden for the Member States could potentially be enormous, as can be gathered when one looks at the facts of the *Ibrahim* case. Ms. Ibrahim's husband had only worked for a very brief period in the host Member State.[109] This sufficed for him to qualify as migrant worker and hence for his family members to invoke the provisions of Regulation 1612/68. Given the young age of his children, this could imply a right of residence for a substantial period of time for Ms. Ibrahim and her children, during which they could fully rely for their subsistence on welfare benefits.

These consequences would be even more drastic if the reasoning is extended to school-going children of non-economically active Union citizens and their primary carer. As was pointed out higher, it cannot be totally excluded that the Court will do so in future case law. In that case all categories of Union citizens could, by moving to another Member State and enrolling their children in an educational establishment there, secure a right of continued residence in that State, even when they no longer satisfy the conditions regarding self-sufficiency.[110] This would result in a much greater number of potential 'welfare tourists'. Besides, granting an unconditional residence right to children of non-economically active Union citizens is more problematic from a financial point of view than granting such rights to children of migrant workers because the latter will in the past have contributed to the host Member State's social assistance systems by paying taxes and social assistance contributions.[111] This may be

[108] See eg Case C-200/02 *Zhu and Chen* [2004] ECR I-9925.

[109] He had only been employed in the UK for a total of about eight months and had claimed incapacity benefits in the UK for an additional nine months.

[110] Non-economically active Union citizens should, in order to obtain an initial right of residence, demonstrate to have sufficient financial resources. Once such a right obtained, however, they could then derive a right of residence even when these conditions would no longer be fulfilled.

[111] See, in this sense, explicitly, Case C-480/08 *Teixeira* [2010] ECR I-1107, Opinion of AG Kokott, para 81. It must be remarked, however, that this contribution appears to have been very limited on the facts of, in particular, the

a further reason why the Court might in the future not be prepared to extend similar rights to school-going children of economically active Union citizens and their primary carer. If the Court should do so nevertheless, I submit, the ensuing financial consequences could be tempered by allowing sufficient scope to the Member States to tackle abuse of residence rights and by allowing Member States, in certain circumstances, to restrict the residence rights discussed to school-going children who are sufficiently integrated in their society.[112]

V. Conclusion

As this article has demonstrated, recent ECJ case law brings important changes to three basic elements of the free movement of Union citizens, which I have labelled the 'classics' of free movement. In the first place, it seems that Union law will increasingly influence the nationality rules of the Member States. The traditional assumption that Member States are exclusively competent to regulate the personal scope of Union citizenship can no longer be maintained therefore. In the second place, the benefits of Union free movement law, in particular those relating to family reunification, can now in some circumstances be invoked by Union citizens even if they have never resided in a Member State other than that of their nationality. Consequently, the traditional assumption that Union law can only be invoked by Union citizens who have moved between Member States is no longer valid. Lastly, recent case law appears to diminish the importance of self-sufficiency as a condition for legal residence in another Member State for longer periods of time.

There can be no doubt that the developments outlined will have significant consequences for the immigration laws and policies of the Member States. First, the Member States' rules on acquisition and loss of nationality will increasingly be tested on their compliance with certain fundamental principles of Union law. This might reduce their margin of discretion, for instance, for refusing to confer their nationality on third country nationals. Second, Member States will have to accord certain categories of static nationals a right of residence together with certain of their close family members. Third country family members of a Member State national thereby derive greater claims for residence than were

---

Ibrahim case. This was not problematic according to AG Kokott since financial contributions are made by migrant workers 'viewed as a group' (*Ibid.*).

[112] The Court could draw inspiration from a line of cases in which it held such 'integration' requirements to be valid (see eg Case C-209/03 *Bidar* [2005] ECR I-2119 and the discussion in Charlotte O'Brien, 'Real Links, Abstract Rights and False Alarms: The Relationship between the ECJ's "Real Link" Case Law and National Solidarity' (2008) 33 EL Rev. 643-665.

traditionally conferred on them under the immigration laws of some Member States. Lastly, Member States will have to accord certain categories of Union citizens from other Member States residence rights even if they are fully dependent for their subsistence on welfare benefits. The resulting financial burdens for the Member States are potentially very significant.

It must be emphasised, however, that these remarkable evolutions in the ECJ case law on Union citizenship are in full development and that it is not yet possible therefore to ascertain their precise scope. Future case law will have to clarify the reach of the principles articulated by the Court in the recent cases discussed and make clear to what precise extent they reduce the discretion of the Member States in crafting their immigration policies.

Collateral Damage from Criminalizing Aggression?

Lawfare Through Aggression Accusations in the Nagorno Karabakh Conflict

Marieke de Hoon*

The regulation of war through the prohibition and criminalization of the act of aggression has provided a common legal language for denouncing an opponent for committing aggressive war that is used between warring sides and is understood globally. Nevertheless, due to the indeterminate nature of aggression, different actors may invoke different interpretations and conceptual frameworks related to the legality of war to accuse the other of aggression. This article asserts that the notion of aggression can be used as a weapon of lawfare because the laws of war can be interpreted differently by different actors. The article explores how this is done by analyzing the Nagorno Karabakh conflict as a case study. A deconstruction of both sides' arguments where they accuse each other of committing aggressive war shows that even though both sides speak the same 'language' of law, they rely on contradictory underlying assumptions, both in their internal argumentative structure as well as between both sides' legal argumentations. The article furthermore asserts that, strengthened by the criminalization of aggression, the indeterminacy of the notion of aggression provides conflicting parties with another weapon to battle with, and another battlefield to fight on. Despite its aim to monopolize and prevent war, the regulation of war and criminalization of aggression thereby provides new ways to continue a conflict, allows law to be used as a strategic tool of lawfare, and creates false presumptions of the ability of law to resolve fundamental disagreement.

Table of Contents

I. INTRODUCTION

On 11 June 2010, delegations from the member states of the International Criminal Court (ICC) came to a consensus agreement on the definition of the crime of aggression and a jurisdictional regime.[1] This outcome of the 2-week long Review Conference of the Rome Statute, taking place in Kampala, Uganda, was unexpected for most followers prior to the conference, and was celebrated widely as a historic achievement. The inclusion of the crime of aggression in the Rome Statute was seen as the capstone of a century-long process to prohibit and criminalize aggressive war. The reason for this celebratory atmosphere was that what had been worked towards but not yet fully achieved at Versailles, in the interwar period with bilateral and multilateral treaties, in San Francisco in 1945, in the special working groups in the 1950s, 1960s and 1970s, in the International Law Commission in the 1990s, or in Rome in 1998, was finally on the verge of culmination.[2] The crime of aggression would enter into force in 2017 or soon thereafter, and with the crime of aggression, the world not only renounced aggressive war as an instrument of national policy, but agreed that it is a crime, for which individuals can be prosecuted. Moreover, it provided the norm with a hierarchically superior status, because forming part of the jurisdiction of the ICC denounces it as one of 'the most serious crimes of concern to the international community as a whole."[3] In short, this enthusiasm in Kampala celebrated progress. With the crime of aggression, the world had come together to climb the barricades and take a collective stand against those considering aggressive war.

---

* Doctoral Candidate at Vrije Universiteit (Amsterdam), Faculty of Law; Senior Counsel at the Public International Law & Policy Group (PILPG) and Co-Director of PILPG's Netherlands Office. The author traveled to Nagorno Karabakh in 2010 to serve as an election monitor for the parliamentary elections, and in February 2012 to speak on aggression at an international conference held in Stepanakert on the legal and (geo-)political status of Nagorno Karabakh.

[1] *The Crime of Aggression* (adopted 11 June 2010) Resolution RC-Res.6.

[2] 'On the verge' because the Kampala amendment provides that the exercise of jurisdiction is pending a decision to be taken by a majority of the states parties after 1 January 2017 (Rome Statute of the International Criminal Court (adopted 17 July 1998, entered into force 1 July 2002) U.N. Doc. A/CONF.183/9 (Rome Statute), available at http://www.icc-cpi.int/NR/rdonlyres/ADD16852-AEE9-4757-ABE7-9CDC7CF02886/283503/RomeStatutEng1.pdf, art. 15*bis*(3) and 15*ter*(3)).

[3] Rome Statute, *supra* note 2, Preamble.

Problematic with aggression is, however, that due to its indeterminate character, the notion of aggression can be applied as a useful weapon of lawfare. Its indeterminacy is caused by the fact that different actors can hold different conceptual frameworks on at least three levels that are associated with an actor's perspective on the limiting power of international law for a particular use of force. Namely, related to i) fundamentally differing world views, ii) fundamentally differing perspectives on the function of the use of force, and iii) fundamentally differing views about the source and binding nature of international law. Since these differing conceptual frameworks are based on differing and often contradictory underlying assumptions, they may lead to fundamental disagreements on the aggressiveness of a particular use of force. Despite the recent consensus agreement in Kampala, the definition of the crime of aggression's 'manifest' criterion, which provides that a prohibited use of force is only a crime of aggression if it is a 'manifest' violation of the UN Charter,[4] does not overcome the fundamentality of this disagreement. It does not, nor does any other part in the amendment on the crime of aggression, provide for a meta-criterion to choose between these fundamentally opposing conceptual frameworks.

Consequently, the legal concept of aggression is flexible enough to provide for different interpretations of aggression so that, in many situations, both sides in a conflict can accuse each other of aggression by both relying on legal argumentation. This method to accuse one's opponent of aggression has become increasingly powerful with the regulation of war, i.e. encapsulating the previously non-legal realm of war with legal norms, and particularly with its criminalization. As David Kennedy explained, with the development of law as vernacular of political judgment, and with the fading of distinctions of war from peace and of law from morality and politics, war has become 'the continuation of law by other means.'[5] Kennedy asserts that Clausewitz was right in his assessment in his time that war is a continuation of politics with other means, but that the notion of law, and particularly its separation from politics, has changed fundamentally, as well as the separation of war from peace.[6] Law has become a strategic tool for military and humanitarian actors to frame a situation to their advantage. It has become a way to communicate a message, to argue for the legitimacy of one's actions and to delegitimize the opponent's actions.[7]

Lawfare is understood by Charles Dunlap and Kennedy as a concept that

[4] Rome Statute, *supra* note 2, art. 8*bis*(1).
[5] David Kennedy, *Of War and Law* (Princeton University Press 2006) 46-47.
[6] Kennedy, *supra* note 5.
[7] Kennedy, *supra* note 5, 125-126.

helps understand how '[l]egal arrangements not only put limits on warfare, but also provide venues to legitimize the use of military force, to de-legitimize the enemy, and to supplement the use of force with less destructive – and less costly – means.'[8] The idea that the regulation of war limits politics, is outdated. Rather than limiting power, the regulation of war, and particularly the criminalization of aggression, provides for a potential intensification of the battle, by enlarging the arsenal of warring sides with the weapon of lawfare. The effect of this regulation is the continuation of the struggle not only on the battlefield, but also in the arena of the law.

How the notion of aggression is usable as a weapon of lawfare becomes apparent when looking at particular conflicts to see how the arguments accusing an opponent of aggression are structured. This article analyzes the Nagorno Karabakh conflict as a case study to illustrate how the regulation of war can allow for this strategic use of law. To do this, the article deconstructs the argumentative structure of both sides in the conflict.

Nagorno Karabakh is a mountainous area in the South Caucasus, predominantly inhabited by a people of Armenian descent, but internationally recognized as part of neighboring Azerbaijan, yet *de facto* independent for over 20 years. When the Soviet Union was on the verge of collapsing, the smoldering conflict about the status of Nagorno Karabakh blazed into open warfare between Azerbaijan and Armenia between 1991 and 1994, and is to date an ongoing 'frozen conflict' with a fragile cease-fire. What began[9] with a resolution adopted on 20 February 1988 by the local Soviet of the Nagorno Karabakh Autonomous region of Azerbaijan to leave the Azerbaijani Soviet and join the Armenian Soviet, led to open demonstrations in Nagorno Karabakh's Armenian-dominated capital Stepanakert, followed by counter-demonstrations in Azeri-dominated cities in the area, violent incidents in the following days, deportations of Azeris from Armenia and Armenians from Azerbaijan, further intensification of violence on both sides, and eventually the outbreak of

[8] Wouter G. Werner, 'The Curious Career of Lawfare' (2010) 43 Case W. Res. J. Int'l L. 61, 66.

[9] The roots of the conflict date back much longer than 1988, and in the months before February 1988, there were other violent inter-communal incidents as well. However, according to Thomas de Waal, who wrote an authoritative, balanced and insightful account on the Nagorno Karabakh conflict, the beginnings of the armed conflict in the early 1990s is usually connected to the uprisings in February 1988 following the resolutions to join the Armenian Soviet, Thomas de Waal, *Black Garden. Armenia and Azerbaijan Through Peace and War* (New York University Press 2003) 18.

war in 1991. The resolution to join the Armenian Soviet was denied by General Secretary Mikheil Gorbachev because he feared setting a precedent by making concessions in this dispute when there were at least nineteen other potential territorial disputes in the Soviet Union that were feared to erupt if he decided in favor of the resolution.[10] Due to the collapse of the Soviet Union and international recognition of Azerbaijan and Armenia as independent states, the inter-communal conflict on the status of Nagorno Karabakh became an inter-state conflict between Armenia and Azerbaijan, and in the eyes of the Armenian-dominated Nagorno Karabakh that declared independence in 1991, between the proclaimed independent state of Nagorno Karabakh and Azerbaijan.[11] Accusations of unlawful violence thereby arose to the status of accusations of unlawful use of inter-state force; accusations of committing aggressive war.

Both sides in the Nagorno Karabakh conflict accuse each other of committing aggression. Both sides dress their positions in legal terminology to convince the world that *they* are law-abiding but that the other is the violator, the aggressor. By these accusations and by invoking law to convince the world of the wrongfulness of the other's behavior, legal argument is used not so much to convince the *opponent* of the rightness of one's claims, but rather for the purpose of winning in the forum of global public opinion, which has come to form the other battlefield, called international law.

This article aims to demonstrate how the notion of aggression is invoked and becomes usable for lawfare.[12] The main argument is that, due to its indeterminacy, the notion of aggression can be used as a weapon of lawfare, which becomes increasingly powerful with its criminalization due to an inherent morality attached to (international) criminal law. The regulation of the legality of war and particularly the criminalization of aggression increases the powerfulness of this means of lawfare by providing a common legal 'language' that is understood globally and that allows for a strong denunciation of the opponent as an aggressor, i.e. a criminal and enemy of mankind. As an illustration of how accusations of aggression can be used for lawfare, the arguments that both sides in the Nagorno Karabakh conflict use to accuse the other of aggression are deconstructed to analyze the assumptions upon which they are based. This analysis shows

[10] De Waal, *supra* note 9, 13.

[11] Even though Nagorno Karabakh has declared its independence in 1991 and has been *de facto* independent from Azerbaijan since, no state has recognized it as an independent state, not even Armenia.

[12] For this interpretation of lawfare, I rely particularly on the works of Charles J. Dunlap, Jr., 'Lawfare Today: A Perspective' (2008) 3 Yale J. Int'l. Aff. 146, Kennedy, *supra* note 5, and Werner, *supra* note 8.

that even though they both speak the 'language' of law,[13] they rely on differing and often contradictory underlying assumptions with regard to the notion of aggression that they employ when posing their arguments against each other. Different conceptual frameworks are applied depending on the type of argument that is being rebutted. The result is that their arguments not only contradict when put opposite one another, but are also internally contradictory. This is a symptom of the indeterminacy of the notion of aggression. *Because* of its indeterminacy, the notion of aggression lends itself to being used as a weapon of lawfare in conflicts where the legality of the use of force is disagreed on.

The reasons to choose this particular conflict as an illustration of this argument for the purpose of this article include that by spending time in the area recently and having the opportunity to listen to discussions on the conflict and the accusations of aggression, I found this a particularly telling example of how aggression as a legal concept can be invoked by both sides to argue their case. However, in other conflicts, similar patterns to those described in this article can be seen. Even though the crime of aggression, as included in the ICC's Rome Statute, is not directly applicable to this conflict,[14] the relevance of discussing the criminalization of aggression in light of the Nagorno Karabakh conflict lies not in the direct applicability of the legal norm for adjudication, but in the way that the process of regulating and criminalizing creates other, counterproductive effects in the usability of the legal norm for lawfare.

Where David Kennedy explores the interrelatedness of law and war in the context of the Charter of the United Nations and the collective security system, this article seeks to further develop this analysis of the regulation of war by exploring this in the context of the crime of aggression, the capstone of the regulation of war. Moreover, where Kennedy insightfully describes the fluidity and diversity of the legal context to assert that the laws of war allow for diverse interpretations, this article adds to this by demonstrating *how* this occurs. Furthermore, the article also aims to add to the discussions on the crime of aggression by applying the understandings of the lawfare discussion to the aggression debate and pointing to the potential collateral effect of criminalizing aggression by

[13] See on the international laws about war as a common legal vocabulary for assessing the legitimacy of war, Kennedy, *supra* note 5.

[14] Neither Armenia nor Azerbaijan is a state party of the International Criminal Court, and there is no reason to assume that the Security Council would be interested in referring this situation as such, besides the fact that the crime of aggression is not operative, and won't be until at least 2017. And even then, there needs to be new aggression in order to meet the temporal requirement of Article 11 Rome Statute.

providing conflicting parties with another weapon to battle with, and another battlefield to fight on.

II. The Notion of Aggression as Weapon of Lawfare

The Nagorno Karabakh conflict finds itself militarily at a stalemate, even though this may be changing with the way Azerbaijan is currently building its military. For over 20 years, the Armenian Karabakhis, helped by Armenia, have been effectively controlling the area and have formed a *de facto* independent state.[15] The 'international community' is not interested enough to intervene or actively push for a sustainable solution, partly because both sides need to be willing to compromise, which they are currently not. And Azerbaijan is militarily not powerful enough (yet) to invade and violently retake the area. Aside from the sniper fire in the border area, the current battle is therefore predominantly fought in the legal arena. Interestingly, both sides use the legal paradigm to add to their violent, military struggles on the ground, a legal fight on the battlefield of global public opinion and international law. In this context, Charles Dunlap defined lawfare as the 'strategy of using – or misusing – law as a substitute for traditional military means to achieve an operational objective."[16]

The regulation of war has turned the idea that war is a right for states – to advance objectives and interests – into a prohibition and a crime. With the term 'aggression', international law refers to interstate use of force that is not authorized by the United Nations Security Council and is not in self-defense, following an 'armed attack'. The allegedly illegal use of force may be considered as aggression if it is directed against the sovereignty, territorial integrity or political independence of another state, or in any other manner inconsistent with the Charter of the United Nations.[17] This definition of aggression was adopted by the General Assembly in 1974. It also forms the basis of the definition of aggression for the purpose of the crime of aggression. In 2010, the Assembly of States Parties to the International Criminal Court (ICC) came to a consensus agreement during the Review Conference for the Statute of the ICC, to expand the jurisdiction of the ICC with the crime of aggression. For the purpose of the *crime* of aggression, the states parties added to the 1974 definition that an *individual* can be held criminally responsible for the crime of aggression if that person planned, prepared, initiated or executed an act of aggression,

[15] However, even though clearly independent from Azerbaijan, one can wonder to what extent it is independent from Armenian influence and control.
[16] Dunlap, *supra* note 12, 146.
[17] *Definition of Aggression*, UNGA Res 3314 (XXIX) (14 December 1974), art. 1.

and was in a position effectively to exercise control over or direct the political or military action of a state, provided that this act of aggression constitutes a *manifest* violation of the UN Charter by its character, gravity and scale.[18] However, even though it is called a 'definition' of aggression, neither the 1974 resolution nor the 2010 Kampala amendment provide definitive answers where fundamental disagreement exists on a particular use of force related to the indeterminacy of the notion of aggression.

The regulation of war and criminalization of the notion of aggression has provided a common legal language to denounce an opponent of committing aggressive war.[19] This language is used between parties and understood globally.[20] However, despite the similarity of the language, the analysis in this article shows that even though parties to a conflict can speak the same 'language' of law, they rely on differing and often contradictory underlying assumptions. This is possible because the notion of aggression is indeterminate. The disagreement about whether a particular use of force constitutes aggression and which side is the actual aggressor is more than merely a disagreement on the scope of the legal provision that prohibits and criminalizes aggressive war. This disagreement stems from differing conceptual frameworks with regard to the nature of international relations, the function of the use of force, and the source and binding nature of international law.

For instance, with regard to the nature of international relations, a worldview that sees the world comprised of a system of states, each in a legitimate struggle to further national interests and policy objectives, may lead to a different perspective on the legality of the use of force than a view that sees the world as a society of states, based on cooperation and interdependence. And this is very different still from a perspective on international relations based on the idea that individuals and civil society groups are relevant international actors.[21] Who are the main victims of aggressive war? States, the stability of a region or world, a community or people, individuals, or humanity even? One of the above, some of the above, all of the above?

---

[18] Rome Statute, *supra* note 2, Article 8*bis*.
[19] Kennedy, *supra* note 5.
[20] See Kennedy, *supra* note 5; and Werner, *supra* note 8, 67.
[21] These distinctions are drawn from the English School in international relations theory. For analysis on the structure and normatively progressivist understanding of contemporary international relations see, for instance, Barry Buzan, *From International to World Society. English School Theory and the Social Structure of Globalisation* (Cambridge University Press 2004) and Andrew Linklater & Hidemi Suganami, *The English School of International Relations. A Contemporary Reassessment* (Cambridge University Press 2006).

The function of the use of force is another aspect of the legality of the use of force that fundamentally disagreed upon. For instance, perspectives may differ on whether force can be used to protect human rights, or, instead, to protect territory, or to protect the interests of a state, or of a community, or when the survival of a state (or of a community or a people) is at stake, or whether force can be used in the interest of stability in a region, or for protecting a *status quo*, or whether there are no circumstances conceivable in which force can or should be resorted to outside the limited scope of self-defense or Security Council authorization. Different actors apply different reasons to determine the legitimacy of a particular use of force, and they often rely on differing or even contradictory assumptions related to the function of war. Where, for example, ideas derived from the just war tradition assume war to be an instrument of law enforcement,[22] the 'war as institution of law'-concept is based on the idea of war as an instrument of furthering policy objectives. [23] The two are mutually

[22] The just war tradition arose in the European Middle Ages as a combination of the ideas of Christian thinkers, who abandoned absolute pacifism, and natural law thinkers. St. Augustine was one of the founders of just war thinking and wrote that even though wars were in principle unfavorable, a wrong inflicted by an adversary required the waging of just wars (St. Augustine, *De Civitate Dei Contra Paganos,* Book XIX, para. VII (6 Loeb Classical ed., W.C. Greene transl. 1960) 150-151). St. Thomas Aquinas developed this idea further and focused particularly on the justice of the causes of war and the rightful intention of the war maker (St. Thomas Aquinas, *Summa Theologiae*, Secunda Secundae, Quaestio 40, 1 (35 Blackfriars ed. 1972) 80-83). War was only just under certain circumstances and was a mechanism to punish wrongs. These criteria include that only a sovereign authority can wage a just war (*auctoritas*), that only certain categories of persons are allowed to engage in the use of force (*personae*), that the war has a well-defined objective (*res*), that the force has to be waged in pursuit of a valid legal claim (*justa causa*), and with the rightful intention (*animus*) (Stephen C. Neff, *War and the Law of Nations. A General History* (Cambridge University Press 2005) 49-68).

[23] See Neff, *supra* note 22, 177. The 'war as institution of law'-concept assumes that it belongs to the prerogative of states to decide on whether or not to go to war, since any other decision-making body is unable to make universally just or 'right' decisions on whether or not to engage in warfare. It regards resort to force as an accepted instrument of foreign policy and international business. War is seen as a rule-governed resort to armed force for the settlement of disputes. The war as institution of international law concept is to certain extent inspired by the writings of Thomas Hobbes. Hobbes argued that because the state of nature is a ruthlessly competitive world in which each individual and each state rightfully seeks to safeguard their own self-preservation, perpetual war is the natural condition between states (Thomas Hobbes, *Leviathan* (first published 1951, Penguin 1985)). According to Hobbes, peace can be *created* through the skillful drafting of treaties and agreements between states, but peace is not the state of nature between states. In pursuit of their own safety, two states can be in conflict with each other with each having right on its side. This idea broke decidedly with

exclusive, since the former relies on the presumption that war is legitimate to further a universal notion of justice, the latter assumes that war is legitimate if furthering particular interests.

The third aspect on differing conceptual frameworks related to the question of legality of the use of force is on the source and binding nature of international law. Is the law derived from the will of the state or from nature? If derived from state will, why would any state be limited by law? If it changes its will, it will no longer be bound. But if the source of international law is nature, divinity, morality or another 'higher' source detached from state will or interests, you stumble upon the problem of sovereign equality due to the question of whose interpretation of the law is leading.[24]

It is due to these fundamentally different perspectives related to the legality of war that the question of how to define aggression has been a source of disagreement amongst legal, political and philosophical scholars and practitioners ever since legal thinking about war started many centuries ago.[25] Each of these differing conceptual frameworks can rely on one out of a number of differing underlying assumptions. Any actor involved in a conflict or judging from the public opinion forum may hold any combination of the above assumptions or others on which it bases its assessment on the legality or legitimacy of a certain use of force. To make matters more complex, each actor may well apply different conceptual frameworks today than it did yesterday or will tomorrow. Moreover, it is different for a situation in one part of the world than in another, and different where other circumstances are involved. In other words, the dimensions of time, space and circumstance added to the differing underlying assumptions attached to worldviews, function of war, and source of international law, provide for an inherently complex net of conceptual frameworks that can be applied and differed upon fundamentally. Consequently, the underlying assumptions of the actors involved in or commenting on a particular conflict, and their approach to the legality of a particular resort to force, may very well differ.

---

the just war tradition. According to Hobbes, opposing sides could both be lawfully entitled to use force since both were exercising their natural right to survival. Hobbes held that if necessary for self-preservation, states are allowed to break treaty obligations and resort to armed force lawfully (Thomas Hobbes, *De Cive* (first published 1642, Clarendon Press 1983)).

[24] This problem has been explained in much detail by Martti Koskenniemi in *From Apology to Utopia. The Structure of International Legal Argument (*Reissue with new Epilogue, Cambridge University Press 2005).

[25] See for an insightful account of the history of the legal argument on war Neff, *supra* note 22.

It is this complexity of potential underlying assumptions and the corresponding complexity of conceptual frameworks regarding the use of force that complicates the application of law to the issue of aggression. The indeterminacy of the notion of aggression leads to fundamental disagreement on a determination of aggression in particular situations. What one actor may find lawful self-defense, another sees as aggressive war; what one may find a heroic intervention, or even a responsibility to protect, another sees as aggressive war; what one may find a rightful protection of a people's land, another sees as aggressive war; etcetera. Even though they speak the same 'language' of law, this language has a different meaning depending on the conceptual frame through which one looks at the law. They disagree and will continue to disagree because they hold different, and often contradicting, underlying assumptions.

However, despite the indeterminacy as a legal notion, the consequential difficulty of construing abstract legal rules to distinguish between legitimate war and aggression, and the accompanying reluctance of the Security Council or international tribunal to determine a situation as aggression, parties in a conflict are not shy to use the notion to accuse their opponent of committing aggression. With the increasing importance attached to the concept in international law (from renunciation[26] to prohibition[27] to crime[28]), the force of the accusation has also increased. To understand better how aggression can be used as a weapon of lawfare, the next sections analyze this dynamic in the context of the Nagorno Karabakh conflict. This conflict is particularly interesting because it demonstrates the indeterminacy of the notion of aggression so clearly, thereby raising questions about the limitations of law in the context of the legality of war.

III. LAWFARE IN THE NAGORNO KARABAKH CONFLICT

Upon visiting Nagorno Karabakh, the incongruity between the *de jure* and *de facto* situation of the status of this mountainous area is striking. When looking at a conventional map, one would presume that the area forms part of Azerbaijan, and is not even directly bordering Armenia. However, when

[26] General Treaty for Renunciation of War as an Instrument of National Policy, also known as the Kellogg-Briand Treaty or Pact of Paris (adopted 27 August 1928) 46 Stat. 2343, T.S. No. 796, 94 L.N.T.S. 57.
[27] Charter of the United Nations (adopted 26 June 1945, entered into force 24 October 1945) 1 UNTS XVI, art 2(4).
[28] Charter of the International Military Tribunal (adopted 8 August 1945) Nuremberg Trial Proceedings Vol. I, art 6(a); Rome Statute, *supra* note 2, art. 5 and 8*bis*.

actually visiting, the world of reality opens and one enters a *de facto* independent country,[29] that is only accessible from Armenia. It has a parliament, a government, a justice system that includes a court of first instance, a court of appeal, and a supreme court, democratic elections supervised by an electoral commission, schools, universities, civil society organizations, a television station and newspapers, and an airport waiting to be opened. And all this with a population of less than 150.000 citizens, in a state of war, and in a decidedly poor economic situation. It is a world of contradictions: in a state of war with Baku, the Armenian Karabakhis hold effective control over an area of at least 7.000 square km,[30] for over 20 years now, even though no state has recognized this independence, not even Armenia.

IV. LAWFARE THROUGH RESOLUTIONS IN THE BUILD UP TO WAR

When war broke out in 1991 shortly after the collapse of the Soviet Union and the independence of former soviet republics, Azerbaijan and Armenia were without strong military forces. This is one of the reasons that the war was fought through law at the same time as on the ground. Prior to the actual military operations, a form of lawfare on the status of Nagorno Karabakh was fought through resolutions by the local parliaments. Following the resolution of 20 February 1988 in which the local parliament of the Nagorno Karabakh Autonomous Region of Azerbaijan requested to leave Azerbaijan and join Armenia, the Armenian Supreme Soviet – the local parliament – adopted a resolution on 15 June 1988 in which it formally gave its approval to Nagorno Karabakh joining Armenia. Two days later, the Azerbaijani Supreme Soviet passed a counter-resolution to reaffirm that Nagorno Karabakh was part of Azerbaijan.[31] This so-called 'war of laws'[32] was furthered by another incendiary resolution by the local parliament of Nagorno Karabakh on 12 July to secede unilaterally from Azerbaijan and rename Nagorno Karabakh 'the Artsakh Armenian

[29] At least independent from Azerbaijan, but one can wonder about the independence with regard to the amount of influence that Yerevan (Armenia) has over Nagorno Karabakh.

[30] Encyclopedia Britannica, available at http://www.britannica.com/EBchecked/topic/401669/Nagorno-Karabakh (last visited 18 June 2012). However, the Permanent Representation of Nagorno Karabakh in the United States estimates the area that is under the jurisdiction of the Nagorno Karabakh Republic at approximately 11.500 square km, in accordance with its Constitution of 2006, see http://www.nkrusa.org/country_profile/overview.shtml (last visited on 18 June 2012).

[31] De Waal, *supra* note 9, 61.

[32] De Waal, *supra* note 9, 60-61.

Autonomous Region'.[33] On 18 July, this campaign of lawfare[34] was temporarily quelled when the full Soviet parliament reconfirmed that Nagorno Karabakh was staying with Azerbaijan.[35]

When the Soviet Union collapsed, the conflict increased in intensity, not only through increasing inter-communal violence, and ethnic cleansings through deportations and persecutions on both sides, but also through further legal actions on both sides. With the independence of Armenia and Azerbaijan, formally, the borders of the former soviets were maintained and became international borders. Nagorno Karabakh was internationally recognized as part of Azerbaijan. Armenia was now in danger of international condemnation for violating Azerbaijan's territorial integrity by laying claims on its territory. This was circumvented, or at least attempted, by Nagorno Karabakh's declaration of independence on 2 September 1991, three days after Azerbaijan had declared independence. The declaration of independence allowed Armenia to say that it was only an interested observer to the conflict, not a party to it.[36] After some failed efforts at peace negotiations, not least because Azerbaijan accused Armenians of shooting down an Azerbaijani helicopter with 22 prominent individuals on board, Azerbaijan responded by revoking the autonomous status of Nagorno Karabakh and renaming its capital Khankendi. Nagorno Karabakh responded to that with a referendum overwhelmingly supporting the independence.

The Armenians were more successful in their military strategy and more powerful because they got hold of the left-behind Soviet weaponry and were the better fighters in Soviet times. Eventually, in 1994, the mass violence ended with a *de facto* independent Nagorno Karabakh and an impassable line of contact between Nagorno Karabakh and Azerbaijan, but the political dispute remained unresolved. No international force was deployed to monitor the frontline, or line of contact, which stretched

[33] Artsakh is a name that nationalist Karabakhi also use to refer to Nagorno Karabakh.
[34] This strategy falls within the description of lawfare made by Qiao Liang and Wang Xiangsui where they state that the use of law can be one of many strategies where the dividing line between war and non-war becomes impossible to draw. Wouter Werner interprets their view as turning the relationship between war and politics upside down. '[W]ar is no longer the continuation of politics with the inclusion of other means; it is politics that has become the continuation – or even just one of the manifestations – of war. Werner, *supra* note 8, 65, citing Qiao Liang & Wang Xiangsui, *Unrestricted Warfare* (PLA Literature and Arts Publishing House 1999).
[35] De Waal, *supra* note 9, 61.
[36] De Waal, *supra* note 9, 161.

approximately 180 kms,[37] and even though the May 1994 cease-fire agreement is still effective, a peace has never been achieved.

On the Armenian side, the area is claimed as either part of Armenia or as an independent state, historically and currently inhabited by the Karabakhis, ethnically and culturally related to the Armenian people. The Azeris, however, claim it as part of Azerbaijan, at least since Stalin made the region a part of the Azerbaijani Soviet in 1921. Furthermore, the Azeris claim that its historical ties with the region are stronger, and that, in any event, the declaration of independence of Nagorno Karabakh and claimed secession from Azerbaijan were illegal because it was enforced militarily and achieved through illegal use of armed force.[38]

V. Accusations of Aggression in the Nagorno Karabakh Conflict

Both sides accuse each other of committing aggression. Even though the notion of aggression has arisen to the status of an 'international crime', importantly for the notion of aggression is that each side is usually convinced of the correctness of its own assertion. The accused party usually believes that its use of force is legitimate, if not lawful.[39] As David Kennedy explained, '[i]t is hard to think of a use of force that could not be legitimated in the Charter's terms. It is a rare statesman who launches a war simply to be aggressive. There is almost always something else to be said – the province is actually ours, our rights have been violated, our enemy is not, in fact, a state, we were invited to help, they were about to attack us, we are promoting the purposes and principles of the United Nations. Something.'[40]

---

[37] International Crisis Group, *Tackling Azerbaijan's IDP Burden*, Policy Briefing (27 February 2012), p. 1, available at http://www.crisisgroup.org/~/media/Files/europe/caucasus/azerbaijan/b067-tackling-azerbaijans-idp-burden.pdf.

[38] My account of the positions and arguments of the Armenians and Azeris are drawn mainly from De Waal, *supra* note 9, conversations with individuals from both sides, and news articles from local and international media.

[39] There are many examples in which force is claimed to be lawfully resorted to even though the legal basis was deemed absent by many others. See for instance the discussions on the scope of self-defense, particularly anticipatory self-defense, for instance regarding the current situation with Iran's nuclear build-up. Also, the discussions on Security Council authorizations. For instance the argument that the US/UK invasion in Iraq was lawful on the basis of Resolution 678 (1990), brought forward by, amongst others, Yoram Dinstein, *War, Aggression and Self-Defence* (4th ed., Cambridge University Press 2005), 294-304. A third example are the arguments put forward to claim that humanitarian intervention without Security Council authorization is legal.

[40] Kennedy, *supra* note 5, 80.

But what does it say if both sides can provide a *prima facie* sound legal argument to accuse the other side of aggression? The simple answer to this is that one side must be plainly wrong. And this type of approach is usually chosen. However, in this article, I would like to take the analysis beyond this stalemate and consider the implications of mutual accusation of aggressive use of force by deconstructing both sides' arguments regarding the other's aggression. This deconstruction demonstrates that the concept of aggression is indeterminate. The consequence of that conclusion is that the application of law, and particularly of criminal law through the crime of aggression and jurisdiction over this crime by the ICC, is highly problematic. But before turning to implications, it is interesting to explore *how* the notion of aggression is used to accuse the other of committing aggression.

The Nagorno Karabakh conflict provides an interesting case study for the notion of aggression. But before analyzing the argumentative structures of both sides, it is important to justify which actors I use for this analysis. On the one hand is the Azerbaijan side. On the other, I discuss the arguments of Armenia and Nagorno Karabakh together, mainly because, even though their interests and (foreign) policy objectives may differ on several issues, with regard to the arguments regarding the use of force, they do not differ significantly. It therefore seems appropriate to discuss the aggression arguments in terms of the Azeri arguments that accuse the Armenians of aggression and the Armenian and Karabakhi Armenian arguments accusing Azerbaijan of aggression. The dominant understanding of the concept of aggression is that it applies only to the use of force between states,[41] even though an argument can be made that the concept of aggression should also apply to *de facto* independent states.[42] Since Nagorno Karabakh is not recognized as a sovereign state by any other state, one could criticize the exercise in this article on the basis of the argument that Nagorno Karabakh may therefore not be regarded as a potential perpetrator nor victim of aggression. However, since the purpose of this article is not to examine the legal application of aggression to this conflict, and therefore not to draw conclusions about whether or not Azerbaijan, Nagorno Karabakh and/or Armenia could be considered aggressors, but instead to

[41] See UN Charter, *supra* note 27, Article 2(4); UNGA Resolution 3314, *supra* note 17; Rome Statute, *supra* note 2, Article 8*bis*; Declaration on Principles of International Law concerning Friendly Relations and Co-operation among States in accordance with the Charter of the United Nations, UNGA Res 2625 (XXV) (24 October 1970).

[42] See for this argument Alexander G. Wills, 'The Crime of Aggression and the Resort to Force against Entities in Statu Nascendi' (2012) 10 Journal of International Criminal Justice 83.

study the structure of the arguments, I find it irrelevant to include the question of the status as state or quasi-state with regard to aggression into this analysis. This would only make matters more complex, without providing analytical benefits.

Regarding the method of selecting the arguments, I want to stress that this analysis is aimed to be explorative in order to illustrate a logic, rather than empirical. I do not understand Armenian, Azeri nor Russian and am therefore unable to study primary sources. The analysis is based on the many discussions over the past few years with those involved in the conflict and descriptions of experts. I have presented and discussed this analysis to regional experts and in Nagorno Karabakh. Therefore, even though the sections below may come across as rather hypothetical due to the lack of primary sources, I believe that the logic and argumentative dynamic described is a helpful illustration of this article's argument that the notion of aggression is fluid enough to be interpreted in what seem like mutually exclusive ways.
This analysis is not aimed to criticize or discuss the substance of the arguments, of either side. Instead, the focus turns to the structure of the arguments, and particularly, the structure of the legal arguments regarding aggressive use of force. There are two levels of contradictions between the underlying assumptions of the legal arguments that are used by both sides when accusing the other of aggression. First, there are contradictions between the different assumptions *within* each side's argumentative structure, and second, there are contradictions between the assumptions on the function of the use of force *between* the two sides to this conflict. In asserting that the argumentative structure of the legal argument is contradictory, I do not mean to criticize either side in the way they choose to formulate their arguments. This can be highly effective for, for instance, political purposes. Rather, it is a critique on the ability of law to deal with situations like these, or, put differently, a study of the limitation of international law when it is applied to the issue of war.

Arguments regarding aggression are directed towards the alleged unlawfulness of the use of force. For the Nagorno Karabakh conflict, one would look at the grounds on which the use of force started in 1991. The arguments that are used can be subdivided in arguments i) based on territory and ii) based on human rights discourse. Arguments that connect the lawfulness of the use of force to territorial claims are either based on (a) the idea that territory belongs to a community or a people, or (b) that it belongs to a state. The human rights-type arguments to legitimate the use of force either rely on (a) the individual's rights or on (b) the right to self-determination and self-preservation attached to a community.

1. *Territorial Arguments*

Both Azeri and Armenian Karabakhis claim historical ties to the area. The Armenian argument comes down to the claim that the Armenian Karabakhis have the strongest historical ties because they have lived there since before the time when the mountains formed part of the great Armenian kingdom. They buttress this, amongst others, by pointing to Armenian inscriptions and Armenian religious buildings that were built in various historical periods and are found throughout the territory. The Azerbaijani argument, on the other hand, claims that the territory belonged to Azeris and point to sources that state that the people claiming to be Armenian Karabakhis are actually 19th century settlers from Persia, or even Albanian descendants, making them in fact Azeris. Both these arguments are based on assumptions that connect a people to territory, that a territory belongs to a community of people or a nation. I will therefore refer to these as community-based territorial arguments.

These arguments that refer to historical claims are often taken into, what I call, state-based arguments, asserting that, because of stronger historical ties, the territory belongs to a *state*: either to Armenia or to an independent Nagorno Karabakh, or to Azerbaijan. The Armenian argument puts forward that the area belonged to the great Armenian kingdom and should therefore naturally belong to the current Armenian state, or, alternatively, be an independent Karabakhi state, Nagorno Karabakh. They buttress their state-based claim by asserting that Stalin's reassignment of Nagorno Karabakh from Armenia to the Azerbaijani Soviet in 1921 was unjustified and illegal, merely an appeasement towards Turkey. International law cannot accept this illegality and injustice, and the use of force to protect Nagorno Karabakh from illegal annexation is therefore justified. The Azerbaijani state-based argument is that Nagorno Karabakh was part of the Azerbaijani Soviet before the collapse of the Soviet Union, and, pointing to the *uti possidetis* principle, that therefore the contested area belongs to the state Azerbaijan. Any use of force against this territory is therefore a violation of Article 2(4) of the UN Charter, which protects the territorial integrity of (*de jure*) states.[43]

2. *Human Rights Arguments*

However, because neither of these community-based and state-based arguments that are grounded on the legality of using force for territoriality

[43] Even if a state has become a 'failed state', international law is still widely believed to regard the entity as a state under international law, as long as it at one point met the Montevideo criteria of statehood and was recognized as an independent state by other states (which and how many suffices remains unclear and part of the political domain, as was reiterated by the Kosovo situation).

provide a legal solution to the dispute and do not convince the opponent of revoking its claims, they are supplemented by human rights arguments – usually individual-based arguments. The Armenian argument is that force is to be used because the human rights of the individuals from Armenian descent are violated by Azeris through persecution, and these rights can be vindicated and protected by using force. The individual-based counterargument of the Azeri side is that it is they, not the Armenians, that are entitled to use force, in order to protect the human rights of Azeri individuals, who are persecuted and ethnically cleansed by Armenians. They point to the now almost entirely Armenian population of Nagorno Karabakh as evidence, since during the Soviet period Azeris made up approximately a quarter of the population.

In addition, there are two other human rights based arguments from the Armenian side, but they are community-based arguments. First is an argument based on self-preservation. The argument is that their resort to force is lawful because the Armenian Karabakhis as a 'people', as opposed to individuals, are under threat of extermination. They argue that living under Azerbaijani rule will cause (further) persecution. For self-preservation, the community is entitled to defend itself, and therefore can lawfully resort to force. And another state, Armenia, is entitled to assist a people that would otherwise be in danger of extinction. This resembles a Hobbesian approach. Hobbes wrote that for self-preservation, there would always be the right to resort to all means. Hobbes derived the idea of self-defense of a state from the right to self-preservation of any individual. Therefore, a community that fights to protect itself from extermination would in the opinion of Hobbes and many thinkers in the 'war as institution of law'-concept of the use of force, be a resort to force in accordance with the law.[44] Hobbes therefore held that if necessary for self-preservation, states are allowed to break treaty obligations and resort to armed force lawfully.[45]

A second community-based human rights argument that the Armenian side invokes is a self-determination argument. Namely, that resort to force by Armenia is justified to assist secession because the internal right to self-determination is denied to the Karabakh people, due to violence against the Karabakhis and the lack of ability to pursue political, economic, social and cultural development. They argue that they should be granted independence by asserting that they are 'subject to alien subjugation, domination or exploitation outside a colonial context', which is declared a violation of the right to self-determination in the Declaration on Friendly

[44] See Neff, *supra* note 22, 177 and further.
[45] Hobbes, *De Cive*, *supra* note 23. See also Hobbes, *Leviathan*, *supra* note 23..

Relations[46] and which was recognized as grounds for external self-determination by the Canadian Supreme Court in the *Quebec* case.[47] In addition, and particularly since the independence of Kosovo, they argue that the abrogation of Nagorno Karabakh as an autonomous region by Azerbaijan in 1991 and the violation of fundamental human rights, block the meaningful exercise of their self-determination, leaving no other means than secession. The Armenian use of force to ward off violators of this right and enable the exercise of self-determination is therefore justified, according to the Armenians. Particularly since Kosovo was recognized as an independent state by some states, especially Western states,[48] because the secession of the Kosovar Albanians was deemed lawful in their opinion due to human rights abuses by Serbia,[49] the argument that such a 'remedial right' to independence exists in international law has been invoked increasingly.

VI. Contradictory Assumptions Within and Between Argumentative Structures

Neither of these bases for arguments (state-based, community-based or individual-based) have led to a resolution of the conflict. They have been incapable of convincing the other side. But besides being unconvincing for the opponent, if they are analyzed as *legal* arguments, relying on underlying assumptions regarding the nature and function of law, they are also contradictory. For example, the arguments described above hold contradictory assumptions i) within both sides' argumentative structure on territoriality, namely between the state-based and community-based arguments, ii) between the territorial and human rights arguments within each side's argumentative structure, and iii) between the Azerbaijani and Armenian arguments with regard to worldviews or nature of international relations.

First, applying a state-based approach to argue that use of force is legitimate because the territory belongs to this or that state, such as set out above, assumes that the world is comprised of states that have claimed parts of territory and have agreed amongst each other on a certain division

[46] Friendly Relations Declaration, *supra* note 41.
[47] *Reference re Secession of Quebec* [1998] 2 S.C.R. 217, para 133.
[48] Kosovo is recognized by 91 states, most of which are Western states (including 22 EU member states), but also including few from Africa, Latin-America and Asia. See for an overview of recognitions of Kosovo: http://www.kosovothanksyou.com/ (last visited on 18 June 2012).
[49] See for instance the letter from the government of the Netherlands, recognizing the independence of Kosovo, *Brief van de Minister van Buitenlandse Zaken* (4 March 2008) Tweede Kamer, vergaderjaar 2007-2008, 29478, nr. 8.

thereof, and that, consequently, a certain territory belongs to a certain state. However, the community-based arguments assume the contrary, namely that it is communities, nations or peoples that have a claim over a territory, not states. In this view, statehood is derived from nation-hood, not from territorial acquisition and division amongst organizational structures called states. They often go together, but not necessarily, as numerous conflicts in the world demonstrate. Conceptually, to argue on the basis that a territory belongs to a state in the way of the state-based argument is to discredit the argument that the territory belongs to a people of the community-based argument: a territorial claim is either derived from the idea of sovereign states' rights over territory or from the idea of peoples' rights over territory, but not a combination of the two. The state- and community-based arguments are therefore mutually exclusive.

Second, the human rights-type arguments also contradict the state-based arguments. For example, to invoke an individual-based argument that armed force can be resorted to in order to protect individuals assumes that there are universal values that are applicable to any individual that belongs to humanity, that the state is a mere production of these individual rights, and that international law is to be understood as based on objective assessments of universal truths and values. This contradicts with a state-based approach, which assumes that the world consists (exclusively) of states, particular (as opposed to universal) interests, and that international law is derived from state will instead of a natural or universal law and cosmopolitan values.

Third, in addition to these internal contradictions within each side's argumentative structures, both sides also argue with each other from different and inconsistent conceptual frameworks regarding the nature of international relations. The argumentative strategy often chosen is to oppose the other's state-based argument by putting forward a community- or individuals-based argument or vice-versa. For example, against the Azerbaijani state-based reference to the UN Charter and claim that the Armenians violate Azerbaijan's territorial integrity, Armenians put forward the argument that the territory actually belongs to Armenian Karabakhi people and/or that their right to self-determination and/or individual human rights are violated. Switching between these argumentative bases is a strategy to provide for a (politically) more convincing and compelling position than the other side, and presents the conflict as a legal conflict between conflicting norms. It adds to the persuasive power of the arguments, which, in 'the court of world public opinion' is often more

important than its validity.[50] As asserted by Kennedy, '[w]hether a norm is or is not legal is a function not of its origin or pedigree, but of its effects. Law has an effect – is law – when it persuades an audience with political clout that something someone else did, or plans to do, is or is not legitimate. The point is no longer the validity of distinctions, but the persuasiveness of arguments.'[51] However, in a similar way as described above for the internal contradictions, this search for persuasiveness merely places these contradictory assumptions opposite each other.

Both sides are convinced of the legitimacy of their own resort to force and claim the other's use of force as aggressive. Interestingly enough, however, both sides use conceptual frameworks that are legal frameworks, i.e. a limitation of unbridled power. For example, if the argument for Azerbaijan is that they are entitled to use force because their territorial integrity is violated, that they are merely defending against aggression, the argument applied to argue the Nagorno Karabakh side is that they are entitled to use force because they are denied their self-determination, and are in the course of a legitimate struggle for freedom from alien subjugation. They may invoke the Kosovo Advisory Opinion of the ICJ that stated that territorial integrity is a principle that only applies in interstate relations and cannot be invoked against a self-determination movement.[52] For Armenia, the argument goes that they are entitled to use force to protect Karabakhi individuals from human rights abuses – a responsibility to protect-type argument. A state-based argument is responded to with a community-based and an individual-based argument, all sides also applying conceptual frameworks that rest on contradictory underlying assumptions with regard to the nature of international relations, the function of the use of force, and the source and binding nature of international law.

The legal dispute is only resolved by making a choice between one or another concept of the use of force. In the examples above, I have purposely simplified the positions of the parties to this conflict to show the contradictions between their perspectives. This does not alter the dynamic of the argumentative process. When one side jumps to another use of force concept and conceptual framework to argue their position, the opponent counters this by switching as well, in order to provide a counterargument, maintaining the contradiction.[53]

---

[50] Kennedy, *supra* note 5, 96.
[51] Kennedy, *supra* note 5, 96.
[52] *Accordance with international law of the unilateral declaration of independence in respect of Kosovo* (Advisory Opinion) 2010, available at http://www.icj-cij.org/docket/files/141/15987.pdf, para. 80.
[53] See for a profound analysis of the structure of the legal argument Koskenniemi, *supra* note 24.

In the arguments exchange between Armenia and Nagorno Karabakh on one side and Azerbaijan on the other, where each side accuses the other of aggression, state-based, individual-based and community-based arguments, each with their own underlying assumptions, are used at the same time, with the purpose to reinforce the other arguments. However, at the same time, they create conceptual contradictions; within each argumentative structure and between the two opposing positions. Neither of these bases for arguments have in itself led to a resolution of the conflict. They have shown incapable of convincing the other side. But besides unconvincing for the opponent, if they are analyzed as *legal* arguments, relying on underlying assumptions regarding the nature and function of law, they may also be contradictory.

VII. The Crime of Aggression: Indeterminacy Meets Moral Repudiation

This analysis demonstrates what Kennedy described as 'the *fluidity* and *diversity* of the legal context,' in which '[o]ften more than one law might apply, or one law might be thought to apply in quite different ways.'[54] He continues, '[a]lthough any of us might well disagree with one or another interpretation, we must recognize that the legal materials are elastic enough to enable diverse interpretations. Harnessing law as a strategic asset to strengthen or restrain the military requires the creative use of *legal pluralism* – and a careful assessment of the power those with different interpretations may have to influence the context for operations. The astonishing thing is that these are differences in *perspective* on a quite similar set of legal doctrines and political considerations.'[55] The notion of aggression is a particularly elastic legal concept that is therefore ever so useful for creative argumentation.

The indeterminacy of the notion of aggression allows parties to invoke differing interpretations and switch between differing underlying conceptual frameworks, while speaking the same 'language' of law, making it an instrument that is flexible enough to be a felicitous weapon in the struggle for global public opinion. Besides *well-suited* for lawfare, it has moreover become an increasingly *powerful* tool for lawfare because the regulation of war and the criminalization of aggression has provided the accusation an opponent of committing aggressive war with the connotation of it being a *jus cogens* violator and a criminal, guilty of committing an international crime, which is according to the Rome Statute of the ICC one of 'the most serious crimes of concern to the

[54] Kennedy, *supra* note 5, 38 (emphasis in original text).
[55] Kennedy, *supra* note 5, 39 (emphasis in original text).

international community as a whole.'[56] It has thereby added a layer of moral repudiation to the concept. Parallel to the development of law as inseparable from politics, and the development of war as inseparable from peace, this moral condemnation has made the *jus ad bellum* a strategic partner for both sides in a conflict.[57]

It is difficult to assess this development as being 'good' or 'bad' as such. Providing instruments in terms of law rather than military means seems a less harmful manner of fighting a war: who will die or lose a limb from that? But changing the battlefield to the arena of law, does not solve the conflict. The positions and objectives of the parties to a conflict remain opposed. As Paul Williams noted, '[t]hey may switch their guns for their pens, but they are still engaged in a very aggressive action to accomplish those same political objectives that led them to the battlefield. Most often, when agreeing to a peace process, the parties have not changed their positions but have simply changed the venue of the battle.'[58] This is highly visible in the Nagorno Karabakh example. Economic circumstances and (geo-)political considerations may have brought parties to verbalize their positions in legal terms rather than only through military power, but a solution to the deep-rooted conflict is ever so far away, and ready to burst out again beyond the current near-weekly casualties.

Despite the presumption that the prohibition to use force in the UN Charter provides legal mechanisms to monopolize war and 'to save succeeding generations from the scourge of war,'[59] it has in fact created a constitutional regime of legitimate justifications for warfare.[60] The Nagorno Karabakh example provides some insight, but not nearly an exhaustive one, into the wide variety of possible justifications for the presumed *jus cogens* norm.[61] By providing the vocabulary not only for restricting the occurrence of war but also for legitimizing the decision to go to war, the prohibition of the use of force has not banned war, and certainly not prevented war as several post-WWII examples have shown. Rather, it has provided a new strategic tool for the continuation of the conflict. As Kennedy put it, 'this bold new vocabulary beats ploughshares into swords as often as the reverse.'[62]

---

[56] Rome Statute, *supra* note 2, Preamble.
[57] See Kennedy, *supra* note 5, 41.
[58] Paul R. Williams, 'Lawfare: A War Worth Fighting' (2010) 43 Case W. Res. J. Int'l L. 145, at 147.
[59] UN Charter, *supra* note 27, Preamble.
[60] Kennedy, *supra* note 5, 79.
[61] Presumed, because one may wonder whether it is not a *contradictio in terminis* to have justifications for breaking a jus cogens norm.
[62] Kennedy, *supra* note 5, 167.

Furthermore, the integrity of law, and in particular of the notion of aggression, may well be at stake. This was noted by Wouter Werner, who raised the question 'What is left of the integrity of law and the responsibility of lawyers if legal provisions are turned into strategic tools to fight an enemy?'[63] Werner thereby refers to Kennedy, who pointed out that an increasingly strategic use of the modern law in war would eventually undermine the normative force of law.[64] One can raise concerns, as Kennedy does, regarding the responsibility of lawyers, both in the military and in the humanitarian professions. By presenting law as a mechanism through which all relevant factors are taken into account and justice is its only outcome, lawyers create an image of law that it cannot deliver.

This is not necessarily or solely due to parties' and their lawyers' use of the notion of aggression (because if the law allows for it, why shouldn't they?), but due to the indeterminacy of aggression and the limitation of law in providing answers to every question. Regulation, and especially criminalization, raises the presumption that law can prevent war or resolve its underlying conflict, and that it can always provide the answer to which side is good and which is bad, which party has the law on its side (if any), and which does not. Understanding the law in this way, as a means to give 'right' answers and solutions to any conflict that arises, may well be an overstatement of the law's capabilities. Such an interpretation of the ability and function of international law falsely provides the idea of the legal language as necessarily encapsulating truths. But it is precisely this perception that allows an indeterminate notion as aggression to be used as weapon of lawfare.

VIII. CONCLUSION

While wars of aggression are believed to be one of 'the most serious crimes of concern to the international community as a whole',[65] at the same time, the meaning of aggression is stretched and molded to mean almost anything one wants it to mean because the indeterminacy and variety of frames through which it can be interpreted allow it so. In Kennedy's words, '[w]here it is clear, the law in war will have winners and losers. Where the law is open and plural, it will be pulled and pushed in different directions, articulated in conflicting ways, by those with different strategic

[63] Werner, *supra* note 8, 67.
[64] Kennedy, *supra* note 5, 135; Werner, *supra* note 8, 67.
[65] Rome Statute, *supra* note 2, Preamble.

objectives."[66] War is the topic *par excellence* where the law is open and plural, and where a state will reject law to trump politics, but rather uses law to support its politics.

The Nagorno Karabakh conflict shows how the notion of aggression can be used as a weapon of lawfare, to fight the disagreements in the arena of law as well as on the ground. It also demonstrates the limits of the law's ability to provide universal truths and solutions to deep-rooted conflicts. Ultimately, even though parties can agree to speak 'law' to one another, different underlying assumptions lead to different interpretations of that law, that may well contradict each other. To think that the law can provide for answers where disagreement stems from fundamental disagreement on which conceptual frame to look through with regard to world views, the function of war, and the source and binding nature of international law, is an overestimation of the law's capabilities. By presenting law as a solution-bringing instrument, and by attaching to aggression the label of 'international crime' despite its indeterminacy, what is in fact put at stake is the *integrity* of the law, and particularly of the notion of aggression. It creates the illusion that law can deliver solutions where it cannot. It adds a layer of moral repudiation to the already heated mix of disagreements in existing conflicts. And it provides warring parties with yet another weapon to fight with, and yet another battlefield to fight on. As with the incongruity between the *de jure* and *de facto* situation in Nagorno Karabakh, the misfit between the presumed and the actual capabilities of the crime of aggression is conspicuous.

[66] Kennedy, *supra* note 5, 127.

DELITOS CULTURALMENTE MOTIVADOS. DIVERSIDAD CULTUREDERECHO E INMIGRACIÓN[1]

Oscar Pérez de la Fuente[*]

Este artículo analiza la defensa cultural o delitos culturalmente motivados. En concreto, se estudia la relación entre Derecho y diversidad cultural desde el enfoque de la aplicación judicial del derecho. La inmigración supone que existen nuevas dimensiones asociadas a la noción de pluralismo. Se explicarán los argumentos en contra del universalismo, del Feminismo y de la seguridad jurídica e igualdad en la ley y los argumentos a favor de que la diversidad cultural es inescapable y enriquecedora, frente al localismo y la no "neutralidad cultural" del Derecho Penal y la aplicación del Derecho sensible a la diferencia cultural.

This article analyses culturally motivated crimes and the use of culture as means of criminal defense. In particular, it studies the relationship between law and cultural diversity with a focus on the judicial application of law, and examines emergent dimensions to the notion of pluralism due to immigration. The article therefore explains arguments against universalism, feminism, and legal certainty and equal application of the law; and argues rather that cultural diversity is inescapable and enriching, and that criminal law must confront local specificities and take a non-"culturally-neutral" position, as well as provide for a culturally sensitive application of the law.

TABLE OF CONTENTS

[1] Este trabajo se enmarca dentro del *Proyecto Consolider-Ingenio 2010 "El tiempo de los derechos"* (CSD2008-00007).

[*] Departamento de Derecho Internacional, Eclesiástico y Filosofía del Derecho. Instituto de Derechos Humanos Bartolomé de las Casas. Universidad Carlos III de Madrid. E-mail:oscar.perez@uc3m.es Universidad Carlos III de Madrid

I. CUANDO LA CULTURA DE UNA PERSONA ES DELITO EN OTRA[2]

Uno de los efectos de la globalización y de los crecientes flujos migratorios ha sido que las sociedades occidentales son más diversas culturalmente. La gestión del pluralismo es una de las tareas actuales de las sociedades democráticas y la acomodación de los inmigrantes implica una serie de políticas específicas en diferentes niveles y, en ocasiones, un cambio de mentalidad de anteriores visiones homogeneizadoras y uniformizadoras. Una de las consecuencias de vivir sociedades más diversas es que en, ocasiones, la diferencia produce perplejidad. Una de las cuestiones que suelen ser objeto de debate en las sociedades occidentales, y quizá donde más genuinamente se han establecido las polémicas, es qué límites debe tener el papel punitivo del Estado en una sociedad multicultural.

La cuestión que subyace es que una de las decisiones de mayor calado de una sociedad es determinar la lista de los delitos y determinar los mecanismos de aplicación judicial del Derecho Penal. Desde este punto de vista, parece relevante analizar si la creación y aplicación del Derecho varía y de qué modo por la inmigración y la diversidad cultural. Estas cuestiones suelen conocerse como *delitos culturales*, *delitos culturalmente motivados o defensa cultural*.[3] Como explica Van Broek, existe una gran diferencia entre la literatura del Derecho continental y la *Common Law* tratando el problema de los delitos que son culturalmente motivados. Mientras la doctrina europea tiende a centrarse en el acto en sí mismo, llamándolos *delitos culturales*, la doctrina americana tiende a tratar el problema desde el

[2] M.M. Sheybani, 'Cultural defense: One person's is another crime' (1986-1987) 9 Loyola Los Angeles International and Comparative Law Journal 751.
[3] J. Van Broeck, 'Cultural defence and culturally motivated crimes (cultural offences)' (2001) 9/11 European Journal of Crime, Criminal Law and Criminal Justice 1.

punto de vista de la defensa del acusado, llamándolo *defensa cultural*.[4]

Para delimitar la cuestión, se podría definir defensa cultural, según Lyman, cuando se *negará o mitigará la responsabilidad penal donde los actos son cometidos bajo una creencia de buena fe, razonable en su propiedad, basada en la herencia o tradición cultural del actor*.[5] Otra definición la ofrece Van Broeck cuando considera que un delito cultural *es un acto de un miembro de una cultura minoritaria, que es considerado un delito por el sistema jurídico de la cultura dominante. El mismo acto es, sin embargo, dentro de un grupo cultural del delincuente, condonado, aceptado como un comportamiento normal y aprobado o, incluso, promovido en una situación dada.*"[6]

Los delitos culturales son un fenómeno donde, a nivel penal, revela una situación de *pluralismo jurídico*. Explica Foblets que esta situación – en que las concepciones son inconmensurables, si no contradictorias- puede producirse en dos circunstancias típicas. Primero, varios grupos normativos simultáneamente influencian a un individuo: aquí el conflicto de normas forma un elemento constitutivo de la personalidad del delincuente. Es el caso entre ciertos inmigrantes jóvenes de la segunda generación que están pobremente aculturados y pobremente "deculturados". Segundo, el individuo puede simplemente desconocer la norma del grupo que está intentado regular este comportamiento. En este caso, los elementos primarios de su personalidad derivan de un grupo cultural que no es familiar con la normas de comportamiento de la sociedad que establece el castigo. Es la experiencia de poblaciones colonizadas y de inmigrantes de primera generación.[7] Como señala Van Broek, el aspecto esencial se basa en el hecho de que un delito cultural está causado por la adherencia a normas morales o jurídicas discrepantes.[8]

Para analizar la cuestión de los *delitos culturalmente motivados* se estudiarán los argumentos en contra del universalismo, del Feminismo y de la seguridad jurídica e igualdad en la ley. Los argumentos a favor de que la diversidad cultural es inescapable y enriquecedora, frente al localismo y la

---

[4] idid. 31.
[5] J.C. Lyman, 'Cultural defense: Viable doctrine or wishful thinking?' (1986) 9 Criminal Justice Journal 88.
[6] Van Broeck, 'Cultural defence and culturally motivated crimes (cultural offences)' (n 4) 5.
[7] M.C. Foblets, 'Cultural delicts: the repercussion of cultural conflicts on delinquent behaviour. Reflections on the contribution of legal anthropology to contemporany debate' (1998) 6/3 European Journal of Crime, Criminal Law and Criminal Justice 191-192.
[8] J. Van Broeck, 'Cultural defence and culturally motivated crimes (cultural offences)' (n 4) 21.

no "neutralidad cultural" del Derecho Penal y la aplicación del Derecho sensible a la diferencia cultural.

II. Argumentos en contra de los *delitos culturalmente motivados*

A continuación de analizará *los argumentos en contra* del universalismo, Feminismo y de la seguridad jurídica e igualdad en la ley.

1. *Universalismo*

La primera línea de argumentación contra la noción de los delitos culturalmente motivados tiene que ver con la plena vigencia de los derechos humanos. El Derecho Penal tiene como objetivo la protección de determinados bienes jurídicos a través de sanciones. La explicitación de esos bienes jurídicos es una decisión de gran calado que toman las diversas sociedades y que depende de aquello que se considere socialmente relevante o moralmente justificado. En suma, la lista de delitos en una sociedad determinada requiere de un ejercicio especial de justificación que puede tener diversos itinerarios: a) Aquello socialmente relevante puede responder a los valores de una mayoría parlamentaria –*argumento democrático*- o a los valores vigentes de la moral social –*argumento de moralismo jurídico*- b) Aquello moralmente justificado puede responder a determinada concepción de la justicia, en general basada en alguna visión de la naturaleza humana –*argumento iusnaturalista*- o basado en los derechos humanos –*argumento de la universalidad*-.

Los argumentos fuertes en contra de los *delitos culturalmente motivados* son el argumento democrático y el argumento de la universalidad. La delimitación conceptual frente al argumento de moralismo jurídico y al argumento iusnaturalista es necesaria ya que, en el debate, se producen algunas confusiones que merecen clarificación. Una de las preguntas más importantes de toda Filosofía del Derecho es "¿Por qué debo obedecer? La respuesta a esta cuestión está vinculada, en parte, por los criterios de justificación de las decisiones públicas. Lo cual suele relacionarse con términos con legitimidad o justicia. El debate sobre los *delitos culturalmente motivados* esta estrechamente vinculado a los argumentos que se utilizan para justificar una de las decisiones públicas más relevantes: cuál es la lista de delitos y cómo deben aplicarla los jueces. La clave es que no es lo mismo condenar por trasgredir un delito natural –*argumento iusnaturalista*-, la moral social –*argumento del moralismo jurídico*-, los derechos humanos –*argumento de la universalidad*- o un delito establecido por la ley positiva –*argumento democrático*-.

Según este último, la mayoría parlamentaria delimita una serie de bienes jurídicos que deben ser protegidos por el *ius puniendi* del Estado, la máxima capacidad de intervención jurídica. Los motivos del legislador para desarrollar su política criminológica pueden ser variados y responden a diversas concepciones ideológicas. El argumento que aquí es relevante es que admitir los delitos culturalmente motivados va contra la vigencia de los derechos humanos, a cuya protección el legislador ha consagrado, en gran medida, la legislación penal.

Según Laporta, los derechos humanos se caracterizan por la *universalidad*, *absolutibilidad* e *inalienabilidad*. Para este autor, los derechos humanos serían derechos morales, que no dependen de circunstancias contextuales, que son exigencias fuertes, en forma de bienes o necesidades básicas, y son irrenunciables para el titular.[9]

La noción de derechos humanos suele recibir la crítica de la *abstracción*. La apelación a la universalidad implica un esfuerzo de justificación metaética que puede recibir una crítica similar -a la que Sandel realizaba a Rawls- sobre el mecanismo de abstracción de las circunstancias contextuales, que para el primero responderían a la parte más valiosa del mundo moral.[10] Una crítica similar frente a la abstracción en las visiones universalistas de base kantiana, está presente en algunas visiones feministas[11]. Como sostiene Pérez Luño sobre los derechos humanos, "no se puede hacer abstracción de su trama real y concreta, es decir contextualizada." Esa aproximación, desde la abstracción, sería *ajurídica*, *amoral* y *apolítica*.[12] En el contexto de los delitos culturales, es particularmente relevante esta interpretación *contextualizada* de los derechos humanos ya que un derecho humano, en concreto, podría interpretarse, de forma diversa, en relación con circunstancias particulares, pero habría un núcleo de certeza –un contenido esencial- que debería respetarse para poder hablar de ese derecho propiamente. Lo relevante es que el problema, como se verá, se traslada a una cuestión de interpretación y aplicación intercultural de los derechos.

[9] F.J. Laporta, 'El concepto de derechos humanos' (1987) 4, Doxa 43.
[10] He analizado la crítica de Sandel a Rawls en O. Pérez de la Fuente, *La polémica liberal comunitarista. Paisajes después de la batalla* (Dykinson, 2.005).
[11] Vid., O. Pérez de la Fuente, 'Feminismo, Liberalismo y Comunitarismo. Una aproximación a sus relaciones, implicaciones y dilemas' (2010) 217 Sistema 77-98.
[12] Pérez Luño explica que sería "*ajurídica:* por prescindir no sólo de los cauces normativos, sino de los propios avatares jurisdiccionales que dan la medida efectiva de la tutela de los derechos humanos; *amoral*, porque crecería de la fuerte convicción y el compromiso con los valores éticos de liberación y la emancipación humanas; y *apolítica,* por extrañar, a lo sumo, un concepto, pero no una concepción de los derechos humanos al no explicitar la voluntad de realizarlos." A.E. Pérez Luño, 'Concepto y concepción de los derechos humanos' (1987) 4 Doxa 62.

Y, desde otro punto de vista, que el contexto es relevante al analizar los derechos humanos

Otra crítica habitual a la noción de derechos humanos es la crítica del *etnocentrismo*. En concreto, su origen occidental se correspondería con la visión de una parte limitada de la humanidad –debido a este origen- y pondría en duda que fueran realmente universales. Frente a esta crítica del etnocentrismo, se pueden desarrollar el argumento *pragmático* y el argumento de la *interpretación intercultural* sobre la universalidad de los derechos humanos.

El argumento *pragmático* sería un una visión interesante contra la crítica del etnocentrismo de los derechos humanos que desarrolla Pareck. Este autor afirma la importancia de la Declaración de Derechos Humanos de la ONU de 1948, que "se ha convertido con el tiempo en una parte importante de la moralidad doméstica e internacional"[13]. La legitimidad del texto vendría del hecho de que la declaración está firmada por un gran número de gobiernos representando diferentes culturas, áreas geográficas y sistemas políticos, y que, además, las personas de todo el mundo han apelado frecuentemente a esos principios en sus luchas contra gobiernos represivos[14]. Pareck considera que la Declaración de Derechos Humanos "provee la base más valiosa para un consenso libremente negociado y constantemente evolucionado de los principios universalmente válidos de buen gobierno"[15]. La posición de Pareck no niega que existan valores universales, lo que pone en cuestión es que hayan de coincidir necesariamente con los valores liberales. La idea clave que desarrolla el argumento *pragmático* sería que, en la práctica, los derechos humanos son universales, son un referente necesario de la moralidad nacional e internacional. Como afirma Ansuategui Roig, "la Declaración –Universal de Derechos Humanos de 1948-, en su sentido de expresión de un determinado consenso histórico en relación con la moralidad de los derechos, constituye un punto de referencia inexcusable en la configuración de la moralidad que se expresa a través de los derechos."[16] El argumento pragmático es relevante para delimitar la diferencias entre el argumento iusnaturalista y el argumento de la universalidad. Como sostienen algunos, parte de tradicional contenido del Derecho Natural se ha positivizado. Existe una Carta Universal de

[13] B. Pareck, 'The cultural particularity of Liberal Democracy', en D. Held (ed.), *Prospects for Democracy: North, South, East, West* (Polity Press 1.994) 173. He analizado las relaciones entre universalismo y particularismo en O. Pérez de la Fuente, *Pluralismo cultural y derechos de las minorías* (Dykinson, 2005).
[14] Pareck, 'The cultural particularity of Liberal Democracy', (n 14) 174.
[15] ibid. 175.
[16] J. Ansuátegui Roig, *De los derechos y el Estado de Derecho. Aportaciones a una teoría jurídica de los derecho* (Universidad Externado de Colombia, 2007) 236.

Derechos Humanos desde 1948, que es una norma vigente y la forma de justificar su universalidad es *pragmática*, no apelando a una común naturaleza humana o planteamientos similares.

El argumento de la *interpretación intercultural* proviene de determinada lectura de Walzer. Este autor habla de la moralidad tenue –*thick*- y moralidad densa –*thin*- donde la universalidad es posible en el primer nivel, desarrollándose en el segundo. De esta forma, afirma: "un (tenue) conjunto de principios universales (densamente) adaptado a estas o esas circunstancias históricas. He sugerido anteriormente la imagen de un núcleo de moralidad elaborado de formas distintas en diferentes culturas."[17] De esta manera, se podría hablar de la universalidad de los derechos humanos, que serían interpretados de forma intercultural en las diversas culturas. Así, la cuestión de los *delitos culturalmente motivados* sería una cuestión de interpretación intercultural. Tras lo que subyace si existen límites interpretativos entre las diferentes culturas. Se puede analizar el artículo 3 de la Carta de Derechos Humanos donde se establece: "Todo individuo tiene derecho a la vida, a la libertad y a la seguridad de su persona." ¿Cabe una interpretación intercultural de derecho a la vida, que justifique o atenúe determinados asesinatos por motivos rituales o culturales? ¿Debe prevalecer una interpretación occidental de la libertad o caben otras alternativas?
Si se analiza el artículo 5 de Carta de Derechos Humanos, se establece "Nadie será sometido a torturas ni a penas o tratos crueles, inhumanos o degradantes." La Corte Constitucional Colombiana ha considerado que la pena comunal del fuete –castigo por latigazos- no es un una pena cruel ni degradante[18]. Lo cual, según algunas visiones occidentales, podría ser criticado pero sería un ejemplo de interpretación intercultural. Este es un caso interesante donde se pone en evidencia que detrás de las teorías de la interpretación existen diversas teorías metaéticas, epistemológicas y semánticas implicadas. Decantarse por un enfoque tendrá consecuencias prácticas en los casos de *delitos culturalmente motivados.*

Por tanto la interpretación intercultural de los derechos humanos es una algo abierta. Sin embargo, Coleman considera que "el uso de pruebas culturales tiende al riesgo a la balcanización peligrosa del Derecho penal, donde los americanos no inmigrantes son sujetos de un conjunto de leyes y los inmigrantes americanos a otro."[19]

---

[17] M. Walzer, *Moralidad en el ámbito local e internacional* (Rafael del Aguila, tr., Alianza, 1996) 36.
[18] La figura simbólica del fuete no constituye tortura ni pena degradante. S. T-523/97 Corte Constitucional Colombiana.
[19] D.L. Coleman, 'Individualizing justice through multiculturalism: The liberals' dilemma' (1996) 96/5 Columbia Law Review 1098.

2. *Feminismo*

No existe un Feminismo, sino más propiamente Feminismos. Pero si algo puede caracterizar una visión como feminista es la crítica de la subordinación de las mujeres. El argumento sería que aceptar los delitos culturalmente motivados supondría aumentar la discriminación de las mujeres y los niños, que suelen ser su víctimas, ya que, de esta forma, se aprobarían o condonarían tradiciones culturales patriarcales. Esta sería una faceta más de las tensiones entre Feminismo y multiculturalismo.[20] Desde esta perspectiva, es relevante la visión de Okin que considera que el reconocimiento de derechos colectivos de las minorías culturales no es parte de la solución, sino que refuerza el problema de la igualdad de las mujeres

La réplica desde otras visiones feministas trata de mostrar como la situación de las mujeres de las minorías culturales tiene una doble vulnerabilidad, en base al género y en base a la identidad. No existe una esencia de mujer, sino mujeres en diferentes contextos. En ocasiones, las visiones de las mujeres occidentales frente a otros colectivos de mujeres se basan en la superioridad y eso crea algunas tensiones.

Sin embargo, el caso de los *delitos culturalmente motivados* que suelen afectar a bienes jurídicos, tan relevantes como la vida o la libertad sexual, parece claro que el argumento contra la subordinación de las mujeres debe ser especialmente considerado. Se denominará el *argumento de la igual dignidad.* Según el art. 1 de la Carta Universal de Derechos Humanos de la ONU, "todos los seres humanos nacen libres e iguales en dignidad y derechos y, dotados como están de razón y conciencia, deben comportarse fraternalmente los unos con los otros."

Muchos de estos delitos tienen que ver con *códigos de honor* que, en diversas culturas, representan el control tradicional de la sexualidad de las mujeres

[20] He tratado las relaciones entre multiculturalismo y Feminismo en: O.Pérez de la Fuente.,"Mujeres gitanas. De la exclusión a la esperanza" (2008) 7 Universitas 109-146; O.Pérez de la Fuente, 'Indígenas y derechos colectivos. ¿es el multiculturalismo malo para las mujeres?' (2004) 13 Derechos y Libertades 399-430 O.Pérez de la Fuente 'Mujeres musulmanas, velo islámico y valores de la esfera pública' en O.Pérez de la Fuente (ed.) *Mujeres: Luchando por la igualdad, Reivindicando la diferencia* (Dykinson 2010) 255-284; O.Pérez de la Fuente "La polémica del velo islámico: algunas estrategias feministas en la laberinto de las identidades", Working Paper el Tiempo de los Derechos, (2010) 19; O.Pérez de la Fuente, "Feminismo y multiculturalismo. Una versión de Ariadna en el laberinto de las identidades" en AA.VV., *Perspectivas sobre feminismo y Derecho* (Dykinson, 2012).

por maridos, padres y familiares. Un cuerno del dilema con la *defensa cultural* es si la cultura puede servir para explicar o justificar la conducta del agresor en estos casos. El otro cuerno es, desde el punto de vista de la víctima, si se acepta esa *defensa cultural*, esto va contra su igual dignidad.

El argumento sería que la mayoría de las víctimas de los *delitos culturalmente motivados* son mujeres y su regulación especial redundaría en la discriminación por razón de género. Sería aceptar, de alguna forma, en las sociedades occidentales costumbres patriarcales de origen cultural diverso. Como afirma Coleman, "el interés del acusado de utilizar pruebas culturales que incorporan normas y comportamientos discriminatorios debe ser ponderada contra los intereses de las víctimas y potenciales victimas en obtener protección y desagravio a través de una aplicación no discriminatoria del Derecho Penal."[21]

3. *Seguridad jurídica e igualdad ante la ley*

Es una característica que suele predicarse del Derecho ofrecer un horizonte de publicidad y claridad para aquellas conductas que se consideran prohibidas y, por tanto, son susceptibles de recibir una sanción. Es algo que está implícito en la noción de seguridad jurídica. Dentro de este principio, también se suele incluir la certeza de que ante supuestos de hecho, iguales o parecidos, se recibirá una respuesta jurídica similar. El argumento sería que aceptar los delitos culturalmente motivados iría contra la seguridad jurídica ya que no estaría claro lo que está prohibido en Derecho y tampoco estaría claro cómo se aplica ya que dependería de *quién* comete el acto. Como sostiene Coleman, esto daría lugar a en "un tratamiento disparatado entre inmigrantes y otros miembros de la sociedad." [22]

Este argumento se desarrolla dentro de la ideología del *centralismo jurídico*, que se podría sintetizar en que el Derecho es y debe ser el Derecho del Estado, uniforme para todas las personas, exclusivo de todo otro derecho, y administrado por un conjunto sencillo de instituciones estatales.[23] Como representantes de esta visión menciona a Bodino, Hobbes, Austin, Kelsen y Hart. Es relevante, para comprender las dimensiones de esta compleja cuestión, que no es lo mismo el sistema europeo, donde algunos, de forma incipiente, podrían defender un pluralismo jurídico de tipo subjetivo -Facchi, Banderlbieden-, que el sistema de la *Common Law*, basado en

[21] Coleman, 'Individualizing justice through multiculturalism: The liberals' dilemma' (n 20) 1097.
[22] ibid. 1097.
[23] J Griffins, 'What is legal pluralism?' (1986) 24 Journal of legal pluralism and unofficial law 3.

precedentes y costumbres, o los sistemas jurídicos latinoamericanos donde existen crecientes reivindicaciones sobre el pluralismo jurídico, basado en el reconocimiento de la ley tradicional indígena.

Un argumento relevante, que tendrá desarrollo posterior, es la noción de delito de Kelsen que se opone al argumento del moralismo legal y al argumento iusnaturalista. Según este autor, el delito o acto ilícito es la conducta humana que, exclusiva y únicamente, el orden jurídico positivo convierte en condición de un acto coactivo –sanción-, sin que pueda predicarse de ninguna propiedad inmanente, ni relación con ninguna norma metajurídica, natural o divina. De esta forma, Kelsen concluye "no hay *mala in se*, sino solamente *mala prohibita*."[24]

El argumento de la *seguridad jurídica e igualdad ante la* ley sostendría que es un principio del Estado de Derecho que las leyes son iguales para todos y deben aplicarse de forma igual. Sería una consecuencia del principio de legalidad. Admitir los *delitos culturalmente motivados* sería pretender una excepción en la ley para los extranjeros. Sería admitir que la *cultura es una excusa* que exculpa o mitiga la condena. Esto redundaría en falta de certeza sobre el ordenamiento jurídico –seguridad jurídica- y supondría un agravio para quienes sí cumplen la ley.

III. Argumentos a favor de los delitos culturalmente motivados

A continuación se expondrán los *argumentos a favor* de que la diversidad cultural es inescapable y enriquecedora, frente al localismo y la no "neutralidad cultural" del Derecho Penal y la aplicación del Derecho sensible a la diferencia cultural.

1. *La diversidad cultural es inescapable y enriquecedora*

El argumento a favor de considerar los *delitos culturalmente motivados* está vinculado, con mayor o menor intensidad, con la tesis de que la diferencia cultural es positiva y debe ser tenida en cuenta en el ámbito de la moral y del Derecho. Esta visión está asociada habitualmente al multiculturalismo. Sin embargo, es necesario hacer algunas precisiones para tratar esta

[24] Según Kelsen, no se trata de ninguna propiedad "La doctrina predominante en la jurisprudencia tradicional, según la cual los conceptos de ilicitud y de sanción consecuente contienen un elemento axiológico moral; de que la ilicitud es necesariamente algo inmoral; que la pena tiene que significar algo difamante, es insostenible, aunque no fuera más por el carácter altamente relativo de los juicios axiológicos que entrarían en juego." H. Kelsen, *Teoría Pura del Derecho* (tr. Roberto Vernengo, Porrúa, 1998) 126.

cuestión. Existen versiones del multiculturalismo, tras las que subyace una diferente metaética y una forma distinta de relacionar moral y cultura.

Frente a la tentación de mostrar al multiculturalismo como esencialmente relativista, donde el valor de una práctica viene dado por la comunidad, cabe una versión pluralista del multiculturalismo, donde se acepte la universalidad de los derechos humanos y una interpretación intercultural de éstos. Esta es la visión de Pareck, Taylor y Young. Y en su versión del culturalismo liberal se podría incluir a Kymlicka y Raz. Lo relevante es que el argumento culturalista ha sido habitualmente utilizado en los actuales debates de Filosofía política. Es uno de los tópicos comunes sobre los que posicionarse.

El principal argumento del multiculturalismo, según Pareck, afirma que la visión que puede aportar una cultura es siempre limitada y, por tanto, la diversidad cultural es una realidad enriquecedora que muestra la pluralidad de las posibilidades de la condición humana. Esta diversidad es algo positivo de lo que no se puede escapar. Las culturas son procesos dinámicos, plurales internamente, y se produce un diálogo entre las culturas que es provechoso, ya que permite ser consciente de la propia particularidad, de las similitudes y diferencias entre culturas y de que están abiertas a influencias, cambios y crítica[25]. Los seres humanos están *incardinados* en una cultura y esto es central para su comprensión del mundo y de las relaciones entre individuo y comunidad.

En relación con los delitos culturalmente motivados, son relevantes dos argumentos relacionados con la visión multicultural: el que tiene que ver con el *pluralismo jurídico* y el relacionado con el *determinismo cultural*.

El *argumento del pluralismo jurídico* suele utilizarse en el contexto de países de pasado colonial donde existen normas de las autoridades legales y normas tradicionales de las comunidades. Sin embargo, en un esfuerzo de análisis, Facchi propone la categoría de *pluralismo jurídico de tipo subjetivo* para la realidad de que viven los inmigrantes en Europa. Desde esta perspectiva, esta autora afirma "la locución "pluralismo jurídico" indica al mismo tiempo, que estamos frente una pluralidad de normas, más que de ordenamientos o sistemas, y que estas normas tienen varios orígenes, no siempre calificables como jurídicos."[26]

Los fenómenos de pluralismo normativo ligados a la inmigración, que

[25] B. Pareck, *Rethinking multiculturalism* (Harvard University Press, 2.000) 338.
[26] A. Facchi, *Los derechos en la Europa multicultural. Pluralismo normativo e inmigración* (Ana Aliverti tr., La Ley, 2005) 37.

tienen mayor correspondencia con el plano judicial y administrativo que el legislativo, deben ser abordados a partir de dos puntos de vista según Facchi: Al primero – centrado en el individuo- le corresponde un enfoque de tipo antropológico orientado a identificar las costumbres, las creencias, los mecanismos jurídicos de los inmigrantes; al segundo –vinculado a las instituciones- uno de política legislativa y sociología jurídica dirigido a analizar la interacción social y normativa y a proponer directivas de acción pública. Naturalmente los dos enfoques se integran.[27]

Esta perspectiva es una forma de tomar en serio la existencia de diversos códigos morales y culturales en una misma sociedad, que reflejan la realidad de los inmigrantes. Es una forma de gestión de la diversidad, sensible a la diferencia cultural. En este sentido, son interesantes las reflexiones de Javier de Lucas sobre el pluralismo normativo:

No es que sostenga que la existencia de hecho de una diversidad de culturas comporte como modelo, como aspiración, o sencillamente, como deber ser la exigencia de la diversidad de códigos valorativos, sino muy simplemente que puesto que hay pluralidad cultural no puede no haber pluralidad normativa. Y es de ahí de donde se da el paso siguientes ¿Cuál es el criterio para defender la superioridad de un código sobre otro?¿qué es lo que permite erigir uno de ellos en criterio, de modelo, en supracódigo al que han de ajustarse los demás, que, por consiguiente, han de ser evaluados por aquél? La respuesta no hace más que reforzar las críticas de los relativistas: es mejor aquel que recoge más fielmente las exigencias de dignidad, la autonomía, la libertad, la emancipación o el progreso humano.[28]

El argumento del *pluralismo normativo de tipo subjetivo* parte de valorar las diversas presiones normativas que reciben los individuos –provenientes de otras culturas- en las sociedades occidentales. Esta cuestión tiene diversas dimensiones de interés filosófico. Sin entrar en un esquema complejo de la teoría de las razones para la acción, como el que utiliza Raz, se podría sintetizar como un posible conflicto de una obligación legal y una obligación moral. La moral que se debería obedecer se compondría del conjunto de valores, creencias y principios de una cultura que determinaría la adopción de un comportamiento concreto. La cuestión entonces es fijar cuáles son los elementos clave para decidir un comportamiento y si el peso de las razones para actuar tras las creencias de una determinada cultura es

[27] ibid 38.
[28] J de Lucas, 'Para un discusión de la nota de la universalidad de los derechos (a propósito de la crítica del relativismo ético y cultural' (1994) 3 Derechos y Libertades 275-276.

más decisivo que cumplir con el precepto legal.

El panorama es más complejo. En ocasiones, el individuo no conoce que exista obligación legal –*error de prohibición directo*- o cree que la trasgresión de la norma legal está justificada o excusada – *error de prohibición indirecto*-. La clave en estos casos es que el Derecho penal exige una voluntad consciente de que se comete un delito, para que se origine un castigo.

Otra cuestión relevante asociada es qué es una 'norma obligatoria' y, en última instancia,
cuál es realmente el concepto de Derecho. Desde la Sociología del Derecho, algunas visiones reivindican que determinadas normas tradicionales tienen carácter jurídicamente obligatorio para los pertenecientes a esa cultura. Este es un punto clave de los *delitos culturalmente motivados*. La idea que subyace es si una tradición o una costumbre implica obligación y en qué sentido y si realmente son normas vigentes y efectivas en la comunidad de origen.

Esta última consideración se relaciona con el *argumento del determinismo cultural*. Esta visión sostendría que los valores y creencias de una cultura determinan el comportamiento de sus miembros de tal forma que no pueden ser culpabilizados, ni sancionados por acciones que sigan esos valores. No se podría condenar a los griegos por practicar la esclavitud porque pertenecía a su concepción del mundo. De esta forma, las culturas establecerían los horizontes de significación que permiten orientarse moralmente a los individuos, pero qué ocurre si esos individuos son juzgados por los horizontes de significación de otra cultura.

El argumento del *determinismo cultural* puede someterse a algunas críticas. La primera es la relevancia de la libertad individual que puede transcender marcos culturales dados y adaptarse a unos nuevos. Sin embargo, para comprender el significado de los *delitos culturalmente motivados* es necesario analizar el grado de obligatoriedad para el individuo de las normas de una determinada cultura de origen. En este sentido, Basile propone un esquema que se basa en los siguientes elementos: a) Acción típica cometida; b) Norma cultural "observada"; c) Grado de vinculación de esa norma cultural; d) Grado de adhesión del sujeto agente a la propia cultura; e) Grado de integración del sujeto agente en la cultura del país de llegada.[29]

En el mismo sentido, Levine considera que las determinaciones de razonabilidad (o habilidad de resistir un esquema cultural dado) deben

[29] F. Basile., *Immigrazione e reati 'culturalmente motivati'.Il diritto penale nelle società multiculturali europee* (CUEM, 2008) 275-276.

responder, entre otros temas, a: *grado de asimilación*, la extensión del acusado del tiempo en residencia, educación y empleo en los Estados Unidos; *identificabilidad*, la existencia de atributos culturales que dan a un grupo étnico una identidad que lo sitúan aparte de otros grupos; *auto-contención*, la extensión en que la comunidad étnica está físicamente segregada de otras comunidades. Finalmente, un acusado demuestra que su comportamiento es consistente con unas creencias o prácticas culturales existentes y vigentes: él no puede inventar nuevas prácticas culturales o recurrir a prácticas culturales que previamente se han extinguido.[30]

Estos factores pueden determinar el grado de adhesión y vinculatoriedad de un individuo a una norma cultural determinada y también el grado de conocimiento del carácter delictivo de la acción. Los supuestos que se pueden dar son múltiples y, por tanto, estos casos de *delitos culturalmente motivados* tienen un componente importante de aplicación en el caso concreto.

2. *Frente al localismo y la no "neutralidad cultural" del Derecho Penal*

La primera aproximación desde este argumento afirma que el Derecho Penal es local y refleja una cultura. La segunda aproximación sostiene que la legislación penal debería justificarse en valores distintos de la apelación a la cultura y tendiera a ser más neutral culturalmente y, además, la aplicación al caso concreto del Derecho Penal fuera sensible a la diferencia cultural.

En su análisis, Basile parte de considerar el *localismo* del Derecho Penal. En suma, la situación parece tal que, parafraseando un dicho popular, se podría decir sin más: "*país al que vas, delito que encuentras*".[31] En esta línea, Silva Sánchez sostiene que la *supranacionalidad* de la ciencia del Derecho penal, de la dogmática y de la Política criminal, no debe identificarse con la *supraculturalidad*.[32]

Realmente, -sostiene Ambos- para el Derecho Penal, así como para cualquier otra ciencia socialmente relevante, resulta aplicable el famoso proverbio propio de las políticas de desarrollo "toda política es local, todo

[30] K.L. Levine, 'Negotiating the boundaries of crime and culture: A sociolegal perspective on cultural defense strategies' (2003) 28 Law and Social Inquiry 48.
[31] F. Basile, *Immigrazione e reati 'culturalmente motivati'.Il diritto penale nelle società multiculturali europee* (n 30) 72.
[32] J.M Silva Sanchez, 'Retos científicos y resto políticos de la ciencia del Derecho Penal', en L. Arroyo Zapatero, U. Neumann, A. Nieto Martín (Coords.) *Crítica y justificación del Derecho Penal en el cambio de siglo* (Ediciones de la Universidad Castilla La Mancha, 2003) 26.

gobierno es local."[33] Como sostiene, Basile, Radbruch ya proclamaba que "el derecho es manifestación de cultura". Entre los criminólogos más recientes, Gian Luigi Ponti escribe que "la norma penal es una de las expresiones más explícitas de los valores prevalecientes en una cierta área cultural."[34]

En esta línea, Silva Sánchez sostiene que "la dogmática, en la representación más generalizada de la misma, no se reduce a espacios lógicos estructurales, sino que entra en cuestiones de contenido, de las que es imposible excluir la valoración o, en todo caso, su vinculación con una determinada forma del ver el mundo."[35]

Para poder desarrollar la segunda parte de este argumento, es relevante antes hacer mención de los tipos de *delitos culturalmente motivados*. De esta forma, expone Basile, los *sectores* del ordenamiento penal, dentro de los cuales se encuentran intersecciones significativas entre el círculo de las normas penales y el círculo de las normas culturales son:
a.-) el sector de los llamados *delitos naturales* ("*mala in se*"), es decir, los delitos que reflejan valoraciones y convicciones radicadas en la cultura, y que sancionan hechos considerados odiosos o detestables por los ciudadanos, incluso en base a una valoración pre-jurídica.
b.-) el sector de los tipos penales construidos mediante el recurso a los llamados *elementos normativos culturales*, es decir, los elementos típicos que "se refieren a datos que pueden ser pensados y representados sólo bajo el presupuesto lógico de una norma" y, particularmente, de una norma *cultural.* Piénsese, por ejemplo, en los delitos de actos obscenos y de publicaciones y espectáculos obscenos, que giran en torno al concepto de "*sentido común del pudor*". Piénsese también en los conceptos de "*decencia pública*", "*escándalo público*" y "*orden y moral de las familias*", los cuales aparecen en otros tantos tipos penales: a través de tales elementos, ¡el legislador penal se vale plenamente de la cultura!
c.-) el *sector* ocupado por las incriminaciones tan *impregnadas de cultura* que su inserción en la legislación penal italiana y su posterior permanencia, modificación o desaparición del derecho vigente se explica sólo en función de la evolución que, paralelamente, conocen las normas culturales correspondientes. Piénsese, por ejemplo, en el delito de duelo, los delitos de adulterio y concubinato, los viejos delitos contra la libertad sexual, originalmente encuadrados en los crímenes contra la moralidad pública y

[33] K. Ambos, 'Dogmática jurídico-penal y concepto universal de hecho punible' (2008) 5 Política Criminal 9.
[34] F. Basile, *Immigrazione e reati 'culturalmente motivati'.Il diritto penale nelle società multiculturali europee* (n 30) 72.
[35] J.M. Silva Sanchez, 'Retos científicos y resto políticos de la ciencia del Derecho Penal' (n. 33) 26.

las buenas costumbres, en la previsión del denominado *matrimonio reparador* y en la llamada *causa de honor* que aseguraba penas ridículas para los autores de delitos como el aborto y el homicidio.

Los delitos culturalmente motivados suelen tratar de:

> 1) *violencia familiar,* maltrato y secuestro de personas, realizada en contextos culturales caracterizados por una concepción de los poderes del *pater familia* distinta a la que hoy inspira a la cultura occidental prevalente.
> 2) *delitos en defensa del honor:* honor familiar o grupal; honor sexual; u honor personal.
> 3) *delitos de reducción a la esclavitud en perjuicio de menores.*
> 4) *delitos contra la libertad sexual*, cuyas víctimas son, a veces, muchachas menores que en la cultura de origen del imputado no gozarían de una protección particular en virtud de su edad. Otras veces las víctimas son mujeres adultas a las que la cultura de origen del inmigrante no reconoce una plena libertad de autodeterminación en el ámbito sexual.
> 5) hechos de lesión personal, consistentes en *mutilaciones genitales rituales o sacrificios estéticos.*
> 6) *delitos en materia de sustancias estupefacientes***,** relativos a drogas cuyo consumo es considerado absolutamente lícito y, a veces, nada menos que recomendado en el grupo cultural de pertenencia del inmigrante.
> 7) quizás también *delitos de terrorismo internacional*, por lo menos limitadamente a los casos en que una consideración atenta del *background* religioso-cultural del imputado permite que los jueces lleguen a una reconstrucción más correcta de los hechos en los que se basaba la imputación.
> 8) finalmente, una última macro-categoría, residual respecto de las anteriores, está compuesta de varios delitos, todos acomunados por el hecho de que el imputado –a causa de la diversidad cultural que lo distingue de la sociedad hospedante- cae en una situación de *error* (sobre el hecho que constituye delito o sobre la ley que prevé el hecho como delito).[36]

El argumento del localismo y no neutralidad cultural del Derecho Penal en los términos expuestos parece dar la razón a Devlin y su visión del moralismo legal frente a Hart y su defensa del liberalismo. Es llamativo que aunque sea de forma moderada el argumento del localismo, que es

[36]Basile, *Immigrazione e reati 'culturalmente motivati'.Il diritto penale nelle società multiculturali europee* (n 30) 159-161.

reivindicado por eminentes penalistas, es una versión del moralismo legal. La posición que desarrolla Devlin es la de considerar que la preservación de los valores morales es una cuestión tan vital para la existencia de la comunidad que éstos deben ser impuestos por ley[37]. En esta línea, asimila los actos de traición y de sedición contra la sociedad con los actos inmorales que van contra el cemento social que es la moralidad[38]. Devlin afirma: "la sociedad no puede ignorar la moralidad del individuo más que pueda ignorar su lealtad; ambas florecen juntas y sin ellas la sociedad muere"[39]. Las pruebas que respaldan la ley moral de las que ninguna sociedad puede prescindir son la intolerancia, la indignación y la repugnancia[40]. Según su visión, el contenido de la moralidad ha de seguir los criterios del hombre de la calle[41]. En *Law, liberty and morality,* Hart realizó una crítica de los argumentos que Devlin sostenía sobre la base de la defensa del valor de la libertad individual, siguiendo el principio de Stuart Mill, que considera que la coerción sólo está justificada para prevenir el daño a terceros.[42] Y por tanto cabe establecer una separación entre las

[37] Devlin afirma que "Si la sociedad tiene el derecho de hacer un juicio y lo tiene sobre la base de que una moralidad reconocida es tan necesaria para la sociedad como un gobierno reconocido, entonces la sociedad debe usar la ley para preservar la moralidad en la misma forma como usa para salvaguardar cualquier cosa que sea esencial para su existencia". P. Devlin, *The enforcement of morals (*Oxford University Press, 1965) 11. Existe traducción en castellano P. Devlin, *La imposición de la moral* (Miguel Angel Ramiro tr., Dykinson, 2010).

[38] Devlin sostiene que "La supresión del vicio es tanto más asunto de la ley como la supresión de las actividades subversivas; no es posible definir una esfera de moralidad privada más que definir una actividad privada subversiva". Y más adelante afirma: "No hay límites teóricos al poder del Estado para legislar contra la traición y sedición, y de la misma forma creo que no hay límites teóricos a la legislación contra la inmoralidad". Devlin, *The enforcement of morals* (n 38) 14.

[39] ibid. 22.

[40] Devlin afirma que "No todo ha de tolerarse. Ninguna sociedad es capaz de prescindir de la intransigencia, la indignación y la repugnancia; son éstas las pruebas que respaldan la ley moral, y ciertamente puede argumentarse que, si no están presentes ellas u otras semejantes, los sentimientos de la sociedad no influirán lo bastante para privar al individuo de la libertad de elección". ibid. 17.

[41] Devlin también utiliza las expresiones 'el hombre del autobus' de Clapham, 'el hombre de mente recta' o 'el hombre razonable' para expresar esta idea. ibid. 15.

[42] Stuart Mill afirma que "este principio consiste en que el único fin que justifica la intervención de la especie humana, colectiva o individualmente, en la libertad de acción de cualquiera de sus semejantes, es su propia protección. Que el único propósito para el que puede ejercitarse legítimamente el poder sobre cualquier miembro de una comunidad civilizada, contra su voluntad, es evitar que perjudique a los demás. Su propio bien, sea físico o moral, no constituye justificación suficiente. Él no puede ser justificadamente forzado a actuar o a abstenerse de hacerlo porque sea mejor para él hacerlo así, porque ello le haga sentirse más feliz, porque en opinión de los demás hacerlo así sería de sentido común, o incluso justo. Éstas son buenas razones para amonestarle, para razonar

acciones que los individuos lleven a cabo en el ámbito público y en el ámbito privado. Hart afirma que "esto no es equivalente a castigar a personas simplemente porque otros objetan sobre lo que hacen"[43].

La visión de Devlin sostenía que la moral social constituía la moral crítica y que la sociedad debía estar cohesionada de acuerdo a sus valores morales, evitando, mediante la ley, que los individuos se comportaran contra estos valores. Hart afirma que es aceptable que sostener que *"alguna* moralidad compartida es esencial para la existencia de cualquier sociedad"[44]. Lo que sería equivalente a afirmar que todas las sociedades tienen alguna moral social. Sin embargo, Hart considera inaceptable y absurdo señalar que "la sociedad es idéntica con su moralidad como lo es en cualquier momento dado de su historia, por tanto, cualquier cambio en su moralidad es equivalente a la destrucción de la sociedad"[45]. Es una crítica a la visión de Devlin que afirma que los comportamientos que vayan contra la moralidad compartida traicionan la existencia de la sociedad. Lo que subyace a la persecución de cualquier disidencia de una determinada moral social es una valoración complaciente del *status quo* y un inmovilismo que no se justifica necesariamente.

Basar la moral crítica en la moral social no garantiza la corrección moral de los principios que la sostienen. Hart alega que todas las moralidades sociales hacen la previsión, en algún grado, de los valores universales de la libertad individual, la seguridad de la vida, y la protección por el daño deliberadamente causado. Seguramente, una sociedad que no reconociera estos valores no es una posibilidad lógica, ni empírica, ya que tal sociedad no tendría valor práctico para los seres humanos. Hart concluye que "podemos con Mill vivir en la verdad de que aunque estos valores esenciales universales deben estar asegurados, la sociedad no sólo sobrevivirá a las divergencias individuales en otros campos desde su moralidad prevalente, sino que sacará provecho de éstas" [46] . La argumentación de Hart consiste en afirmar que la moralidad crítica asegura

---

con él, para persuadirle o para suplicarle, pero no para obligarle o inflingirle cualquier mal en caso que actúe en forma diferente. Para justificar esto, la conducta de la que se desea disuadirle tendría que haber sido calculada para perjudicar a las otras personas. En la parte que le concierne meramente a él, a su independencia es, por derecho, absoluta. El individuo es soberano sobre sí mismo, sobre su propio cuerpo y sobre su mente". J. Stuart Mill, *Sobre la libertad* (Cristina García Cay tr., Espasa Calpe, 1991) 74-75.

[43] H.L.A. Hart, *Law, liberty and morality* (Oxford University Press, 1.981) 47-48. Existe traducción en castellano H.L.A. Hart, *Derecho, libertad y moralidad* (Miguel Angel Ramiro tr, Dykinson, 2006).

[44] Hart, *Law, liberty and morality* (n 44) 51.

[45] ibid. 51.

[46] ibid. 71.

la libertad individual como valor universal y que esto, en mayor o menor grado, debería estar reconocido en las diferentes morales sociales. Con lo cual, la imposición coercitiva del valor libertad individual debería suponer una consideración diferente de los comportamientos divergentes, incluso valorarlos positivamente como una muestra de la diversidad humana.

La cuestión es compleja: ¿Hasta qué punto el Código Penal debe reflejar los valores de la moral social? ¿Debería ser relevante la categoría de delitos naturales?

En este punto, como se ha visto, se oponen el argumento del *moralismo jurídico* y el argumento *iusnaturalista* frente al argumento del *daño a terceros* y el argumento *democrático*. La primera perspectiva, sostiene que el Derecho penal debe considerar delito los actos inmorales en términos de moral social, mientras que la segunda visión defendería que los delitos naturales -*mala in se*- deberían estar castigados penalmente y estarían definidos por el derecho Natural. Frente a esta visiones, el argumento liberal considera que deben ser delito aquellas acciones de los individuos que provoquen daño a terceros y que los delitos son exclusivamente lo que ha aprobado el órgano legal que tiene la competencia, el Parlamento elegido democráticamente, lo que se conoce como delitos *mala prohibitia*.

Por ejemplo, en el caso de un legislador democrático decidiera despenalizar el aborto durante las primeras semanas de gestación, el *argumento iusnaturalista* se opondría, el argumento del *moralismo jurídico* lo haría depender de la moral social, el *argumento liberal* lo aceptaría basándolo en la noción de autonomía individual y *argumento democrático* lo aceptaría como manifestación de la ley positiva, producto de la voluntad democráticamente elegida.

Es relevante estas diferencias en la justificación de las decisiones públicas ya que en la interpretación que hacen algunos jueces, en ocasiones, se utilizan argumentos que se basan en la idea de que el individuo conocía necesariamente que estaba cometiendo un delito porque era un delito natural y no podía no conocerlo. O era contrario a la moral social y eso lo haría depender del grado de asimilación del individuo a la cultura receptora.

La tarea, en el contexto de los *delitos culturalmente motivados*, sería delimitar específicamente qué fundamento está detrás de los bienes jurídicos que se protegen a través del Código Penal. Si ese fundamento está vinculado con los derechos humanos -*argumento de la universalidad*-, con una visión de la moralidad social -*argumento del moralismo jurídico*-, con los principios del Derecho Natural -*argumento de iusnaturalismo*- o como protección de la

autonomía individual –*argumento del liberalismo*-. La cuestión que se plantea, con estos delitos, es cuál es la legitimidad de juzgar acciones cuando el fundamento de los bienes jurídicos que se quiere proteger es localista y cultural.

Una posible respuesta es "*estos son nuestros valores y los queremos proteger*". Pero esto pone en evidencia que el Derecho tiene un impacto diferente para los miembros de las minorías y definitivamente no es un instrumento neutral. Como afirma Facchi, "es obviamente imposible crear normas que realmente traten todos los componentes culturales de la misma sociedad igualmente mientras la adopción de las categorías lingüísticas y conceptuales occidentales implica una pérdida de neutralidad."[47]

Una aproximación para abordar los elementos culturales en el Derecho Penal sería analizar la noción de *tabú*. Esta noción se define, según Freud, como "una serie de limitaciones a las que se someten los pueblos primitivos, ignorando sus razones y sin preocuparse siquiera de investigarlas, pero considerándolas cosas naturales y perfectamente convencidos de que su violación les traería los peores castigos."[48] Existe un componente cultural en el Derecho Penal y no es necesariamente neutral para los miembros de las minorías.

Viendo la lista de *delitos culturalmente motivados,* que ofrece Basile*,* es preciso recordar el *argumento de la igual dignidad* en el sentido de que las victimas habituales de estos delitos son mujeres y menores. La cultura no puede ser excusa para discriminar más a las mujeres o atenuar la pena a sus agresores. Este argumento debería ser también tenido en cuenta en la aplicación intercultural de los casos concretos del Derecho Penal.

3. *La aplicación del Derecho sensible a la diferencia cultural*

Un argumento a favor de los *delitos culturalmente motivados* es que la aplicación específica en los casos concretos puede suponer una vía intercultural de interpretación del Derecho. Esto supondría que la *defensa cultural* funcionaría desde las circunstancias particulares de los casos que abocan a los jueces a realizar un ejercicio interpretativo que incorpora un componente cultural. Son diversas las circunstancias y las formas en las que la cultura incluye en el Derecho, Levine sintetiza tres estrategias de la *defensa cultural*:

---

[47] A Facchi, 'Multiculltlicultural policies and female immigration in Europe' (1998) 11/4 Ratio Juris 352.

[48] S. Freud, *Totem y tabú* (Alianza Editorial, 1967) 33 citado por E. Lamo de Espinosa, 'El vicio y la ambivalencia normativa' (1988) 42 REIS 8.

a) La primera estrategia, *razón cultural:* Un demandado que es acusado de una específica intención de delito puede basarse en su cultura para proveer una explicación plausible –una razón cultural- para su conducta que puede refutar la inferencia que muchos de nosotros haríamos basados en nuestras tradiciones culturales americanas[49]; b) La segunda estrategia, *tolerancia cultural:* El acusado admite que cometió los actos alegados con la intención perjudicial requerida para hacer las acciones criminales en Estados Unidos. No obstante, sostiene que no debería ser considerado penalmente responsable en los Estados Unidos porque su cultura nativa tolera o condona el comportamiento ultrajante como una respuesta aceptable al comportamiento de la víctima;[50] c) La tercera estrategia, *requerimiento cultural:* La estrategia del requerimiento cultural se basa en un enfoque puramente interno de cultura. Cuando habla el acusado de que sus valores nativos le "compelen" a cometer un acto delictivo y sus acciones son "predeterminadas" por su trasfondo cultural, nosotros le quitamos de toda responsabilidad por su comportamiento.[51]

Dadas estas premisas y los anteriores argumentos expuestos, se analizarán a continuación algunos casos de *delitos culturalmente motivados,* con la intención de desarrollar el alcance de la interpretación del Derecho sensible a la diferencia cultural.

a. Caso Kimura

En 1985, en Santa Mónica, California. Fumiko Kimura se introdujo en el océano con sus dos hijos después de conocer que su marido le era infiel. Ella fue rescatada, pero sus dos hijos murieron. Kimura declaró que sus acciones constituían la práctica tradicional japonesa del *oya-ko shinju*, o suicidio de padres e hijos.[52]

La *defensa cultural* en este caso se basaría en una traición japonesa que consistiría en el suicidio conjunto de padres e hijos. La primera cuestión sería si se está ante un caso de tolerancia cultural o requerimiento cultural, donde parece que este último argumento no es aplicable ya que el hecho de una infidelidad marital no "compele" a un suicidio ritual del resto de la familia. Obviamente este punto podría ser debatido desde otras visiones en el trasfondo cultural japonés, pero una cosa es explicar comportamientos en términos culturales y otra es concebir que la cultura determina

[49] K.L. Levine, 'Negotiating the boundaries of crime and culture: A sociolegal perspective on cultural defense strategies' (2003) 28 Law an Social Inquiry 49.
[50] ibid. 56-57.
[51] ibid. 62.
[52] N.S. Kim, 'The cultural defense and the problem of cultural preemption: A framework for analysis', (1997) 27 *New Mexico Law Review* 101-102.

necesariamente el resultado del suicidio de la madre y los hijos. Por tanto, en este caso no funciona el *argumento del determinismo cultural*.

La defensa cultural en el caso Kimura sostendría que no puede hacerse penalmente responsable de la acción a la madre porque en su cultura existen otros valores que justifican estos rituales en determinados casos. Ella conoce que su acción está prohibida jurídicamente, pero no comprende que esto se aplique a su caso ya que sus valores culturales aceptan esta práctica bajo determinadas circunstancias. Como explica Monge Fernández, estos casos se denominan "error de comprensión" donde el sujeto conoce la norma prohibitiva, pero no se le puede exigir la comprensión de la misma, esto es, su proyección o interiorización como parte de su catálogo de valores. Es precisamente en esta categoría donde podrían ubicarse los casos de "socialización exótica", cuando el sujeto pertenece a una cultura o subcultura diferentes, y en su virtud ha recibido una educación distinta, incluyendo en su escala axiológica valores adversos e incompatibles con los del Ordenamiento jurídico.[53]

En este punto es relevante el grado de asimilación de la madre a la cultura occidental para delimitar si se está efectivamente ante un caso de socialización exótica o, en cambio, ella conoce y, en cierta medida comparte, los valores de la sociedad de acogida. Otro elemento relevante es, como señala Basile, la existencia en el país de origen de una norma penal del contenido análogo a la norma penal violada.[54] Es decir, ¿qué respuesta específica ofrece el ordenamiento jurídico japonés a un caso de suicidio tradicional de padres e hijos? La respuesta a esta cuestión es clave para comprender la conciencia de sancionabilidad que tenía la madre al cometer los hechos.

Un argumento relevante en este caso es de de la *igual dignidad,* aplicado a los menores. Es decir, considerar el punto de vista del más débil, de las víctimas. Otro elemento relevante, es el bien jurídico dañado, que es el de la vida. Precisamente la trasgresión de un bien jurídico tan fundamental hace difícil invocar la doctrina del *error de prohibición directo.* Como señala Felip Saborit, la Sentencia el Tribunal Suprema español de 11 octubre 1996 establece que "no es permisible su invocación en aquellas infracciones que sean de ilicitud notoriamente evidente, de tal modo que se manera natural o elemental se conozca y sepa la intrínseca ilicitud."[55]

---

[53] A. Monge Fernández, *El extranjero frente al Derecho penal. El error cultural y su incidencia en la culpabilidad* (Bosch Editor, 2008) 93.

[54] F. Basile, *Immigrazione e reati 'culturalmente motivati'.Il diritto penale nelle società multiculturali europee* (n 30) 311.

[55] D. Felip Saborit, *Error iuris : el conocimiento de la antijuricidad y el artículo 14 del Código Penal* (Atelier, 2001) 202.

La posibilidad de *defensa cultural* sería alegar un *error de prohibición indirecto* donde la acción estaría prohibida legalmente pero estaría excusada o justificada por algún motivo.
En este caso, sería un *error de comprensión* sobre la licitud de la práctica tradicional de suicidio padres e hijos. Descartado el argumento de *determinismo cultural*, en este caso la acción de la madre merece ser responsable penalmente por homicidio de sus hijos, pero no en un grado máximo. La *defensa cultural* justificaría en este caso una atenuación de la pena que debería ser ponderada según las circunstancias del caso, especialmente el grado de asimilación a la cultura occidental de la madre.

b. Caso Mouea

En 1985, Mouea, miembro de la tribu Hmong de Laos, secuestró a una mujer Hmong y tuvo relaciones sexuales con ella a pesar de sus protestas. En el juicio, el defensor argumentó que sus acciones eran coherentes con la práctica tribal Hmong de *zij poj niam*, o *matrimonio por captura*. El abogado de la defensa sostuvo que en el ritual del *matrimonio por captura*, un hombre secuestra a una mujer y la lleva a su casa familiar donde se consuma el matrimonio. Se espera que la mujer proteste ante las iniciativas sexuales como un testimonio de su virtud. El hombre muestra que es valioso para ser su marido continuando sus iniciativas sexuales a pesar de sus protestas. [56]

La *defensa cultural* sostendría que este caso sólo puede ser explicado realmente desde el trasfondo cultural de los participantes. Como sostiene Coleman "los defensores de esta posición sostienen que, a pesar de la establecida doctrina en contra, para los inmigrantes, la ignorancia del derecho *es* una excusa."[57]De esta forma, los proponentes de la defensa cultural argumentan que incluso cuando un inmigrante ha aprendido nuestras costumbres valores y leyes, sus prácticas culturales deben ser respetadas en nuestro sistema jurídico.[58]

En el caso Mouea, no sólo se da un *error de comprensión*, sino más bien un *error culturalmente condicionado.* En estas situaciones, explica Monge Fernández, se da la posibilidad de excluir la culpabilidad, si el autor hubiese tenido que esforzarse de tal modo que surja su *inexigibilidad jurídica*, negándose la *reprochabilidad*, sólo en el caso de que tal error de

[56] Kim., 'The cultural defense and the problem of cultural preemption: A framework for analysis', (n 53) 101-102.
[57] Coleman, 'Individualizing justice through multiculturalism: The liberals' dilemma' (n 20) 1101.
[58] ibid.1102.

comprensión sea, a su vez, un error de prohibición *invencible*. Singularmente concurre esta modalidad cuando la dificultad para la comprensión está condicionada culturalmente, es decir, el sujeto tiene conocimiento de la norma prohibitiva, aunque no interioriza el mandato de ésta por razones culturales. Por consiguiente, en este caso no se le puede reprochar la falta de *internalización* como "comprensión".[59]

El concepto de objeto de conocimiento de la antijuridicidad explica Felip Saborit se basa en tres principios: a) *Contrariedad a los principios éticos-morales y lesividad social;* b) *Infracción del ordenamiento valorativo material del Derecho*; c) *Conocimiento de la sancionabilidad o del carácter penal de la prohibición.*[60]

Es relevante que el primer punto no equivale al *argumento del moralismo jurídico* que defendía Devlin. Desde esta perspectiva Felip Saborit aclara que "los principios ético-sociales o morales imperantes en la sociedad *Sittenwidrigkeit* no es condición necesaria ni suficiente para formular una prohibición jurídica de cualquier clase."[61] Lo que parece clave en el caso Mouea es si el individuo conocía que la violación era un delito sancionable y que su conducta con aquella mujer formaba parte del hecho típico del delito. Según la Sentencia del Tribunal Supremo español de 21 noviembre 1995, "para que no haya error de prohibición basta que el sujeto conozca que lo que hace u omite es un comportamiento ilícito, es decir, contrario al ordenamiento jurídico, sin que sea preciso ningún otro conocimiento más concreto."[62]

En este caso juega un papel relevante el *argumento del determinismo cultural*, que se desarrollaría advirtiendo que entre los valores de la tribu Hmong de Laos estaría el ritual del *matrimonio por captura* y el hombre y la mujer del caso Mouea proceden de esa cultura. El argumento sostendría que el comportamiento sexual del hombre respecto a la mujer estaría determinado o condicionado por la cultura Hmong. Por tanto, se produciría un *error culturalmente condicionado* del individuo al no internalizar los valores del ordenamiento jurídico, ni de la moral dominante, de la sociedad receptora.

En este caso merecería reflexionar sobre la Sentencia del Tribunal Supremo español de 12 mayo 1994 respecto de un delito de corrupción de menores cuando afirmaba "el acusado no podía ignorar la gran inmoralidad

[59] Monge Fernández, *El extranjero frente al Derecho penal. El error cultural y su incidencia en la culpabilidad,* (n 54) 93-94.
[60] Felip Saborit, *Error iuris : el conocimiento de la antijuricidad y el artículo 14 del Código Penal* (n 56) 110-116.
[61] ibid. 110.
[62] ibid 178..

y consiguiente ilicitud de su conducta y las graves consecuencias que para su hija y los otros menores habría de tener."[63] El caso de la sentencia era de un individuo socializado en la cultura occidental. Lo que el caso Mouea pone en cuestión son los valores tras lo que se justifica una prohibición penal. Un itinerario buscaría justificaciones en los *delitos naturales* y los argumentos del *moralismo jurídico.* Según esta visión, la conciencia de la comisión de un delito debería ser "cuasi evidente" para todas las personas. El otro itinerario es el de los delito *mala prohibitia*, del *argumento democrático* y del *argumento de la universalidad*.

Analizando este ultimo argumento en el caso Mouea, si se aplica la 3 de la Carta de Derechos Humanos donde se establece: "todo individuo tiene derecho a la vida, a la libertad y a la seguridad de su persona." ¿Puede una interpretación intercultural de derecho universal a la libertad y la seguridad admitir como una práctica válida el *matrimonio por captura*? Parece que entra dentro del núcleo de certeza de la noción de libertad, la libertad sexual a la que se refiere el caso Mouea. Además, se debería añadir el argumento de *igual dignidad*, donde en este caso la víctima es una mujer. Aceptar esta práctica cultural supone agravar la subordinación de la mujer.

Por último, cabría una interpretación más sutil del caso Mouea que sostendría algo parecido a que lo reprensible del comportamiento del individuo depende sólo del contexto donde se produce, ya que en la tribu Hmong de Laos estas mismas acciones serían aceptadas. El problema con la práctica del *matrimonio por captura* es básicamente que se basa en la omisión de cualquier referencia a la verdadera voluntad de la mujer e, incluso, interpreta sus protestas como muestra de virtud. No es un matrimonio desde la libertad de las partes y con respeto a su igual dignidad. Lo terrible es que haga lo que haga la mujer, será *víctima de la captura.*

c. Caso Lu Chen

En 1987, en Brooklyn, Nueva York, Dong Lu Chen se enfrentó a su mujer sobre su relación sexual, cuando ella le explicó que tenía una relación extramarital, le pegó ocho veces en la cabeza con un martillo, matándola. En el juicio, un experto de la defensa testificó que en la cultura tradicional china, el adulterio de la esposa es una prueba del carácter débil de marido y que el divorcio está considerado una gran vergüenza para los ancestro de uno. [64]

Conviene recordar las palabras de Hart cuando mostraba algunas "verdades

[63] ibid 179.
[64] Kim, 'The cultural defense and the problem of cultural preemption: A framework for analysis' (n 53) 101-102.

obvias", que deberían incorporarse a todo sistema jurídico, que se refieren a la característica de la mutua vulnerabilidad humana donde la "prescripción más característica del derecho y la moral es *no matarás*."[65] Este caso, no se basaría en un *error de prohibición directo*, sino más bien un *error de prohibición indirecto*.

En estos casos, explica Monge Fernández, el sujeto *conoce* la tipicidad prohibitiva pero considera que su conducta está justificada, pudiendo reconducirse la figura a dos modalidades. En primer lugar, la falsa suposición de la existencia de una causa de justificación que la ley no reconoce (falsa creencia en la existencia de un precepto permisivo) y, en segundo lugar, la falsa suposición de circunstancias que conllevan la aplicación de una causa de justificación (falsa creencia en la existencia de una tipicidad permisiva objetiva –denominada erróneamente por la doctrina como *justificación putativa*). En síntesis, quien actúa bajo un error indirecto de prohibición conoce la significación antijurídica de la misma, sabiendo que es un hecho desvalorado por el Derecho, creyendo erróneamente que tal desvaloración queda desvirtuada por la concurrencia de una causa de justificación, por lo que podría recibir su solución como un "caso de creencia errónea."[66]

Un individuo –originariamente de cultura china - mata cruelmente a su mujer después de conocer que tiene una relación fuera del matrimonio. El componente cultural del caso viene de la diversa valoración del adulterio y del divorcio. Según la *defensa cultural,* el individuo, aunque conociera de la ilicitud del homicidio, se consideraría justificado por las graves consecuencias –en términos de su cultura- de la acción de su mujer. Parece claro que el argumento del determinismo cultural no sería aplicable ya que existen soluciones alternativas –que no comportan el homicidio de la mujer- para las circunstancias del caso. De hecho, se puede afirmar que matar a mujeres adúlteras no es una práctica cultural china *strictu sensu*.

Cabe plantearse qué solución establecería el ordenamiento jurídico chino para un caso similar, puesto que el homicidio también es delito. Este sería un caso de *razón cultural* que ofrecería una explicación plausible –como razón cultural- distinta de nuestras tradiciones culturales a la intención detrás de un delito. En este caso, la *razón cultural* buscaría explicar la comisión del homicidio después del conocimiento de un adulterio y el ensañamiento de los ocho martillazos.

---

[65] H.L.A. Hart, *El concepto de Derecho* (Genaro Carrió tr, Abeledo Perrot, 1998) 239-247.

[66] Monge Fernández, , *El extranjero frente al Derecho penal. El error cultural y su incidencia en la culpabilidad,* (n 54) 96.

El caso Lu Chen podría interpretarse como algunos casos análogos sin componente cultural donde se aplicaría el *estado de necesidad*. De esta forma, la *razón cultural* funcionaría como circunstancia atenuante de la pena en una causa de justificación. Pero es obvio que esta interpretación va contra el *argumento de la igual dignidad* o una posible interpretación intercultural del derecho universal a la vida.

d. Caso Pahn

En abril 1992, cinco amigos de un grupo de jóvenes budistas ayudaron a su amigo Binh Gia Pahn a llenarse de gasolina y prenderse fuego. El inmigrante de 43 años estaba protestando contra la legislación del gobierno vietnamita para suprimir el Budismo. Los amigos de Pham grabaron su muerte con cámaras de video, y entonces explicaron el incidente a la policía, no dándose cuenta que la ayuda al suicidio es un delito.[67]

Este caso parece estar dentro de las características de un *error de prohibición directo* donde los individuos no conocen que la acción está jurídicamente prohibida. La prueba de esta tesis es que graban el incidente y se ponen en contacto con la policía para explicarlo. En este punto, la clave es hasta qué punto el error de prohibición es vencible o no. Bajo esta perspectiva, se pueden concretar los motivos razonables que pueden provocar al autor a reflexionar sobre la antijuridicidad de su conducta, Roxin distingue tres grupos de casos: 1.-) En primer lugar, los casos de dudas, que provocaran un conocimiento eventual de la antijuridicidad, es decir, cuando el sujeto tenga razones para dudar de la ilicitud de su conducta, en el caso de que ésta infrinja las normas de la Ética Social realmente vigentes en una sociedad dada; 2.-) En segundo lugar, se admite la *vencibilidad* del error de prohibición en los casos en el autor desarrolla su actuación en el marco de una actividad reglada, cuya regulación jurídica el autor conoce o podía conocer, sin que hiciese nada por tener los conocimientos necesarios al respecto; 3.-) Finalmente se constata un error de prohibición vencible cuando el sujeto es consciente de la dañosidad social de su comportamiento, ya hacia otras personas, ya hacia la sociedad.[68]

La ayuda al suicidio, en el marco de una protesta pública, ¿va contra la moral social? ¿es contraria a una actividad reglada? ¿existe conciencia de su dañosidad social hacia otras personas o hacia la sociedad?

---

[67] Kim, 'The cultural defense and the problem of cultural preemption: A framework for analysis' (n 53) 101-102.

[68] Monge Fernández, , *El extranjero frente al Derecho penal. El error cultural y su incidencia en la culpabilidad,* (n 54) 88.

Respecto a la cuestión de la moral social, se debería contar con elementos de evaluación intercultural ya que la relación entre individuo y colectivo e, incluso, el valor de la vida tienen diversas concepciones en las diferentes culturales. Por ejemplo, hablando de manera general, Occidente y Oriente pueden tener visiones distintas de las relaciones entre individuo y colectivo. En este punto, en determinadas culturas, existen ocasiones donde individuos se inmolan por determinados ideales o valores colectivos. Esta visión es criticada desde Occidente.

Si se sitúa el caso Pahn en el contexto occidental donde jóvenes budistas de origen vietnamita ayudan a suicidarse a su compañero, cabe plantear que la moral social occidental condena generalmente la ayuda al suicidio, que no se trataría de una actividad reglada y que se produce daño a la sociedad. Este último argumento merece mayor desarrollo. El individuo que se suicida no debería ser sujeto de responsabilidad penal siguiendo el principio de Stuart Mill, que establece que la intervención estatal se justifica sólo cuando se produce daño a terceros. Sin embargo, las personas que ayudan al suicidio sí deberían ser responsables penalmente porque, en su acción, están produciendo un daño irreversible a un tercero -aunque éste no lo considere como tal- y a la sociedad.

La *defensa cultural* en este caso podría alegar que la punición de la ayuda al suicidio es una regla muy técnica del Derecho Penal que no es compartida en algunas culturas. La réplica es que se trata del bien jurídico "vida", que es fundamental y algunos consideran que irrenunciable por el propio sujeto. Existe la prevención, en diversos ordenamientos jurídicos, para no alentar comportamientos que ayuden al suicidio en la forma de un delito. La *defensa cultural* sostendría que lo que difiere es la jerarquía entre vida y libertad desde la evaluación en las diferentes culturas.

Este caso podrá interpretarse como un caso de error de prohibición directo donde existe una *razón cultural* que daría una explicación plausible del comportamiento. Sin embargo, dependiendo de algunos factores específicos del caso, podría interpretarse que el *error de prohición* es vencible y la *razón cultural* serviría como atenuante.

e. Caso cenizas hindú

Un hindú fue detenido mientras tiraba "basura" a un río en Holanda. De acuerdo con las regulaciones holandesas es una forma de contaminación. El acusado argumentó que la "basura" eran meramente restos quemados (madera y flores) de una ceremonia para recordar a una persona difunta y éstos, de acuerdo con una costumbre hindú, necesitaban ser lanzados a una

corriente.[69]

Este caso parece circunscribirse dentro del *error de prohibición* ya que el individuo no sabía que la acción era un delito. El juicio debería dilucidar si este error era evitable o no. Desde esta perspectiva, Monge Fernández se plantea "¿En qué casos, por tanto, podría afirmarse la *inevitabilidad* del error? Según Roxin, en aquellos supuestos en que el ciudadano desconoce la norma, a pesar de haber cumplido con la expectativas derivadas de un grado normal de fidelidad al Derecho. De un lado, en estado casos, la ausencia de pena no comporta en la sociedad ningún tipo de conmoción en los ciudadanos, es decir, desde el punto de vista de la *prevención general,* no existe necesidad de la pena. De otro lado, el sujeto no denota una posición contraria al ordenamiento jurídico que exigiera la imposición de una sanción: esto es, tampoco existe una necesidad del castigo desde el punto de vista de la *prevención especial."*[70]

El caso de la *cenizas hindú* tiene como uno de sus elementos clave el bien jurídico que se trata ya que aquí la cuestión reside no en la vida o la libertad sexual, como en casos anteriores, sino más bien en la contaminación que pueda producir en un río restos quemados de madera y flores. Esto supone que, desde la prevención general y especial, las consideraciones a realizar no provoquen especial conmoción en la sociedad. En esta línea, parece también más justificado considerar como inevitable el *error de prohibición.*
Como señala Hurtado Pozo, "este tipo de error es difícilmente admitido, sobre todo en el caso de las infracciones que forman el núcleo histórico del derecho penal. En efecto, el homicida difícilmente puede alegar un error de prohibición afirmando que no sabía que está prohibido matar a una persona. En caso de infracciones previstas en la legislación complementaria puede ser más fácil admitir el error de prohibición (por ejemplo, si concierne la prohibición de organizar juegos de azar, de violar las leyes sobre inmigración o la reglamentación de un actividad industrial o comercial)."[71] Una regulación sobre contaminación de los ríos parece un caso claro de legislación complementaria.

La conclusión de los argumentos esbozados parece concluir que en este caso sí funcionaría la argumentación de la *defensa cultural.*

---

[69] J. Van Broeck,'Cultural defence and culturally motivated crimes (cultural offences)' (n 4) 4.
[70] A. Monge Fernéndez, , *El extranjero frente al Derecho penal. El error cultural y su incidencia en la culpabilidad,* (n 54) 86-87.
[71] J. Hurtado Pozo, 'Derecho Penal y diferencias culturales: El caso peruano', (2008) 86-87 Derecho Penal y Criminología 84-85.

f. Caso Kargar

Mohammed Kargar, refugiado afgano junto a su familia, había contratado como *babysitter* una joven vecina. En presencia a de ésta, Kargar besó el pene de su hijo de un año y medio. La joven luego le comentó lo sucedido a su madre, quien recordó haber visto en el album de fotos de los Kargar un suceso similar, por lo que lo denunció a la policía. Kargar nunca negó los hechos, manifestando que besar el pene de su propio hijo es, dentro de su cultura, expresión de afecto paternal y que no tiene ninguna connotación sexual. A pesar de lo manifestado, se le imputó delito de abuso sexual, siendo condenado en una primera instancia, a pesar de tenerse claro que no había significado sexual en su acto y que era una expresión cultural, sobre la base de que la conducta se comprendía dentro del delito. Sin embargo, la Corte Suprema del Estado revocó el fallo resaltando el carácter cultural y no sexual del comportamiento de Kargar.[72]

Lo relevante de este caso es que no puede explicarse adecuadamente si no es en relación con el componente cultural. La cuestión reside en una diversa interpretación cultural de un hecho –un beso en el pene del hijo-. Según la visión occidental, podría tratarse de algún tipo de abuso sexual. Según la visión del padre afgano, es una muestra de afecto de carácter cultural y sin connotaciones sexuales. Como se ha dicho es una cuestión de interpretación de trasfondos culturales. Desde esta perspectiva, Monge Fernández señala que "en la argumentación de Roxin, es el juez el que debe configurar un *baremo de evitabilidad*, a partir de la formulación de tres cuestiones. La primera, consiste en responder a la pregunta sobre si la actuación del autor le ha ocasionado necesariamente una preocupación sobre la ilicitud de su conducta. En segundo término, si se constata la existencia de un *motivo* y el sujeto no ha actuado o no se ha esforzado por informarse o aquel es insuficiente, se constata la responsabilidad, atendiendo a razones preventivas. Finalmente, habrá de investigarse la opinión de un experto acerca de la antijuridicidad, que hubiese suscitado en el autor razones para inhibirse a actuar."[73]

¿Es ilícita la acción del caso Kargar? La cuestión a dilucidar debe partir del bien jurídico de que se trate –la libertad sexual del hijo- y de la gravedad de la posible trasgresión –las consecuencias para el hijo de recibir ese beso-. La segunda cuestión en este caso tendría que ver con si el delito de abuso sexual tiene un componente cultural, en la línea del argumento del moralismo jurídico o de los delitos naturales del iusnaturalismo. Si se parte

[72] R. Carnevali, 'El multiculturalismo: un desafío para el Derecho penal moderno' (2007) 3 Política Criminal 19.
[73] A. Monge Fernández, , *El extranjero frente al Derecho penal. El error cultural y su incidencia en la culpabilidad,* (n 54) 87-88.

de los delitos *mala prohibitia,* del argumento universalista y del argumento democrático se deben considerar exclusivamente los bienes jurídicos protegidos por decisión del legislador democrático. Lo relevante de este caso es el significado cultural diverso de la acción el que considera lícito o no el comportamiento.

La cuestión de si el padre podría evitar la acción habiéndose informado sobre la interpretación occidental del beso. Esta consideración es de difícil prueba y está vinculada con el grado de asimilación. Pero parece que el padre realizaba esas acciones como una muestra de afecto paternal genuino y desconocía otras implicaciones que pudieran darse. Obviamente los defensores del moralismo jurídico y de los delitos naturales podrían argumentar en contra de esta visión.

Respecto a la antijuridicidad y las razones para actuar, parece que en este caso la gravedad de la trasgresión del bien jurídico está en el elemento cultural o, mejor, en su interpretación. Dicho de otra forma, la gravedad de delito parece residir en el dolo del agresor y en las consecuencias para la victima. Sobre el dolo, parece probado en el caso que la voluntad de padre con el beso no tiene connotación sexual y, sobre las consecuencias, es difícil de calibrar, pero este no sería un caso de abuso grave.

En este caso, por tanto, la *defensa cultural* jugaría un papel relevante que, incluso, exculparía al padre de ser condenado de un delito por su acción.

g. Caso ablación

En 1983, la Corte Suprema de Francia tuvo que enfrentarse al caso de "Daniel Riecher" que hizo la ablación del clítoris y uno de los labios menores de su hija. La Corte sentenció el 20 de agosto de 1983, y sostuvo que el acto constituía una "mutilación" de acuerdo con la definición del art. 312.3 de Código Penal francés.[74]

Este es el tipo de *delito culturalmente motivado* que ha sido más expuesto en la opinión pública europea. Esta cuestión tiene una serie de dimensiones complejas que habitualmente no son tenidas en cuenta. Los argumentos que defienden esta práctica cultural están vinculados con el argumento del *determinismo cultural* y del *pluralismo jurídico de tipo subjetivo.* En este sentido, algunas posiciones sostienen que no es lo mismo la ablación de las niñas en su cultura que en las culturas occidentales.

[74] Foblets, 'Cultural delicts: the repercussion of cultural conflicts on delinquent behaviour. Reflections on the contribution of legal anthropology to contemporany debate' (n 8) 187.

Los argumentos en contra son el *argumento del universalismo* y el *argumento de la igual dignidad*. En concreto, si la interpretación intercultural de la libertad puede amparar que una madre decida la ablación genital de su hija. Un valor occidental es la libertad y la autonomía individual, pero hasta qué punto incluye ese poder padres e hijos. ¿Puede enfrentarse la hija cuando sea mayor y consciente a la decisión de su madre? El *argumento de la igual dignidad* defendería que la ablación es un ejemplo cultural de la sumisión de las mujeres a los hombres ya que tiene como finalidad limitar la libertad sexual de las mujeres.

Estos argumentos en contra parecen convincentes en este punto, aunque alguna versión del *argumento del determinismo cultural* debería considerarse para ponerse en el lugar de estas mujeres (madres e hijas) y analizar globalmente su problemática. Esta visión es la que defiende Facchi que realiza un agudo enfoque de esta práctica cultural en Occidente que busca "acercarse, comprender, hacerse cargo, explicar, adaptarse, mediar: todas las modalidades de acción opuestas a la represiva". [75] La visión que exclusivamente criminaliza la ablación, la convierte en la bandera de los peligros del multiculturalismo, pero se olvida de las condiciones legales, económicas y sociales de vida de las mujeres inmigrantes, no parece la adecuada.

Es clave, además, la educación en derechos humanos y el empoderamiento –*enpowerment*- de las mujeres. El *argumento de la igual dignidad* debería hacerse efectivo, no desde un paternalismo occidental, sino que surja desde las propias mujeres. Como sostiene Facchi, que "las mujeres adquieran una mayor autonomía dentro de su comunidad, de modo de poder llegar a la decisión de no infligir la escisión a sus hijas, sin que esto deba asumir un significado de ruptura con su cultura"[76]

La *defensa cultural* en este caso supone plantear la cuestión en términos más amplios que la técnica represiva de un delito. Aunque el *argumento de la universalidad* y el *de la igual dignidad* recomiendan que la ablación sea considerada delito en Occidente. El objetivo de erradicar esta práctica debería comprometer, necesariamente también, medidas promocionales, educativas y culturales para que las propias mujeres decidan que ese no es el mejor futuro para sus hijas.

h. Caso bebé albino

En Bélgica (Valonia), en la década de los ochenta, una joven madre de

[75]Facchi, *Los derechos en la Europa multicultural. Pluralismo normativo e inmigración* (n 27) 88, 63-88.
[76] ibid. 88.

origen africano, en un estado de desesperación porque había dado a luz a un bebé albino, lanzó el recién nacido al agua, causando que se ahogara.[77]

Una primera posibilidad, que deberá ser descartada, es considerar que este caso se trataría de una mujer inimputable. Según el artículo 20.1 del Codigo Penal español, la inimputabilidad es la "anomalía o alteración psíquica permanente que impide comprender la ilicitud del hecho o actuar conforme a esa comprensión." Monge Fernández distingue al respecto cuatro categorías: psicosis, oligofrenias, psicopatías y neurosis.[78] Es relevante que han habido algunos intentos en America Latina de equiparar a las personas indígenas como inimputables, visión que es contraria a los derechos humanos. En este sentido, Bartomei afirma que "son violatorios de los derechos humanos de los indígenas la aplicación rígida de leyes cuando éstas solamente no son comprendidas o son ignoradas, sino cuando, con frecuencia, no tienen significado alguno en el contexto de la cultura local; o bien cuando legislaciones penales consideran a los indígenas como "inimputables" o "incapaces" o sujetos a algún "régimen especial."[79]

Este caso se explica porque es una creencia cultural africana que el nacimiento de un bebé albino es símbolo de mala suerte y de los peores presagios. En este punto, es claro que el bien jurídico protegido es de suficiente relevancia, como el de la vida, y también funcionaria el *argumento de la igual dignidad*, referido a un menor. Por tanto, es lógico que el ordenamiento jurídico reaccione ante la trasgresión de un bien jurídico relevante protegido por el legislador democrático –*argumento democrático*- y manifestación de los derechos humanos –*argumento universalidad*-.

La *defensa cultural* funcionaría en este caso como alguna forma del *error de prohibición indirecto* ya que la mujer sabía que su acción estaba prohibida, pero podría creer que existía alguna causa que la justificara o la exculpara. El caso podría explicarse como un *estado de necesidad*, un poco *sui generis*, para conseguir salvarse de una maldición que crea una gran perturbación a la madre. Es relevante que el hecho de matar a un hijo es delito en las diferentes culturas y está penado en la cultura de origen. Una de las claves del caso es si es cierta, en términos culturales, la relación entre bebé albino, maldición y perturbación de la madre. Parece que se trate de un caso de tabú cultural africano. Pero de eso no se deriva que la solución deba ser la

[77] M.C. Foblets,'Cultural delicts: the repercussion of cultural conflicts on delinquent behaviour. Reflections on the contribution of legal anthropology to contemporany debate' (n 8) 188.

[78] A. Monge Fernández, *El extranjero frente al Derecho penal. El error cultural y su incidencia en la culpabilidad,* (n 54) 65-72.

[79] M.L. Bartolomei, 'Universalismo y diversidad cultural en América Latina' (1995) 20/7 El Otro Derecho 55.

muerte del hijo e, incluso, que ésta esté condonada en la cultura africana.

La *defensa cultural* podría funcionar en el caso de bebé albino, si fuera fehacientemente probada en juicio, como una forma de atenuante. Pero por motivos de prevención general y especial, por la relevancia del bien jurídico y del argumento de la igual dignidad, esta madre debería ser condenada por su acción.

IV. ALGUNAS CONCLUSIONES

La cuestión que subyace a estos casos de delitos culturalmente motivados es que la cultura es un elemento que, en ocasiones, debería ser tenido en cuenta para los jueces a la hora de motivar las sentencias. No siempre es un argumento definitivo, en ocasiones es una forma de mitigar la pena o un atenuante, pero lo relevante es que existe una dimensión a la hora de decidir casos sobre justicia penal que está vinculada al transfondo cultural.

Sin embargo, los argumentos en contra son fuertes y principalmente tienen que ver con el *argumento de la igual dignidad* ya que las víctimas suelen ser mujeres y menores. Si algunos delitos culturales son delitos de *honor,* la interpretación intercultural no debería ser una coartada de los agresores. Otro argumento es que existen unos límites de la interpretación intercultural de los derechos humanos donde *no todo vale.* El contenido esencial de los derechos humanos permite vislumbrar unas propiedades relevantes –de forma tenue–, que deberían ser interpretados –densamente- en las diferentes culturas.

La ponderación sobre la idoneidad de aplicar la *defensa cultural* debería tener en cuenta, la menos, estos elementos: a) *Bien jurídico:* ¿Se corresponde con un derecho humano? ¿Tiene que ver con la igual dignidad o la libertad sexual?; b) *Daño a terceros:* ¿Está claramente especificado el daño a terceros?; c) *Grado de asimilación:* ¿En qué medida el individuo conoce -o está integrado- en la sociedad receptora?; d) *Reciprocidad*: ¿La acción es delito en la sociedad de origen? De la respuesta a estas cuestiones, el argumento de la defensa cultural ganará o perderá fuerza, servirá para exculpar, para mitigar o simplemente no será adecuado. El caso de las *cenizas hindú* sería un caso claro de aplicación adecuada de la *defensa cultural*, mientras la aplicación de la *defensa cultural* en el caso *Mouea* produciría resultados contraproducentes.

Este artículo es un intento de fijar la atención en las coordenadas de Derecho, diversidad cultural y la aplicación judicial, desde los *delitos culturalmente motivados*. Es una perspectiva que debería desarrollarse más. En la tesitura de que la gestión del pluralismo requiere quizá modificar

algunos esquemas o refinar las justificaciones que se dan detrás de algunas decisiones. La *defensa cultural* quizá no sea un argumento definitivo para todos los casos, pero puede ser relevante para los jueces en el ejercicio de la ponderación, entre otros elementos, como un aspecto a tener en cuenta cuando personas de *otras* culturas deben responder por sus acciones.

ENFORCEMENT OF PENALTY CLAUSES IN CIVIL AND COMMON LAW: A PUZZLE TO BE SOLVED BY THE CONTRACTING PARTIES

Ignacio Marín García*

This paper claims the need of transnational rules to secure the enforcement of penalty clauses in international commercial contracts, since the contractual toolkit that parties may use seems to be insufficient to address both the clash between the civil and the common law traditions, and the existing disparities among civil laws in this area. The international community acknowledged this need a long time ago, but unfortunately the tremendous effort exerted in many different harmonization projects is unlikely to lead to the certainty that actors in international trade demand.

TABLE OF CONTENTS

* Attorney at law in Spain, CUATRECASAS. GONÇALVES PEREIRA. LL.M., Columbia Law School (2011), and PhD Candidate, Universitat Pompeu Fabra (Barcelona, Spain). The author is grateful to Professor Alejandro M. Garro for his guidance and helpful comments, and to *Fundación Rafael del Pino* for the Scholarship for Postgraduate Studies granted (2010-2011).

I. INTRODUCTION

In the field of penalty clauses, defined as any agreement for the payment of a fixed sum on breach of contract, one of the most distinctive features among civil and common law systems is the extent of the judicial review of the stipulated sum. While common law courts may declare unenforceable such agreement by virtue of the principle of just compensation,[1] civil law courts may only reduce a grossly excessive stipulated sum. The agreed sums exceeding the actual loss of the promisee are unlikely to be enforced by Anglo-American judges, [2] and those deemed extremely high by Continental European judges are also moderated. Therefore, broadly speaking, the principle of non-enforcement of contract penalties governs in common law, and the principle of enforcement of penalties subject to reduction controls in civil law.

The main difference between these two legal traditions lies on their different notions about contract liability: in common law systems, the payment of damages constitutes true fulfillment of the contractual promise. Whereas, in civil law systems, contract liability is an effect arising from the breach or a sanction.[3] Thus, for a civil lawyer, the amount

[1] In these jurisdictions, contract law does not aim to force the promisor to perform, but to compensate adequately the aggrieved promisee, E. Allan Farnsworth, *Contracts* (4th edn, Aspen 2004) 811. Regarding the principle of just compensation, the holdings of two cases, one American and the other English, are very illustrative of the fact that in common law systems freedom of contract encompasses such a wide autonomy for the parties to enter a contract, but a much more restrictive one to arrange remedies against its breach. First, in *Jaquith v Hudson* 5 Mich 123 (Mich 1858), the Supreme Court of Michigan stated that 'courts will not permit the parties by express stipulation, or any form of language, however clear the intent, to set it aside'. Second, in *Addis v Gramophone Co.* [1909] AC 488 (HL), the House of Lords insisted that 'damages for breach of contract [are] in the nature of compensation, not punishment'.

[2] Nevertheless, English law and American state laws present substantially different regimes governing liquidation of damages, with respect to the analysis of its validity and the legal consequences for a penalty clause.

[3] Judge Holmes noted that 'the duty to keep a contract at common law means a prediction that you must pay damages if you do not keep it,- nothing else' (Oliver W Holmes, Jr, 'The Path of the Law' (1897) 10 Harv L Rev 457, 477). See also Fernando Pantaleón Prieto, 'Las nuevas bases de la responsabilidad contractual' (1992) 46 Anuario de Derecho Civil 1719, 1737-40.

stipulated is always intended to be higher than the loss.[4]

In a comparative view, the major objection against the common law of penalties is that parties are placed in the worst of all possible scenarios, without the flexibility of enforcement of penalties subject to reduction (most civil law systems), and without the certainty of literal enforcement of penalties (Spain).[5] Indeed, from the economic analysis of law, the extreme rigidity of common law courts has been criticized on account of judges disregarding upon these provisions with disfavor,[6] since any judicial review resulting in the unenforcement of penalties threatens the function of this remedy against breach.[7] However, in international commercial contracts, the enforcement of those penalties constitutes an even major concern, since uncertainty is much higher due to the applicable law and the court decision when adjudicating the dispute or executing the judgment or the arbitration award.

Part I briefly presents the rules governing penalties in three different jurisdictions: a common law jurisdiction, the United States (Section II.1), and two civil law jurisdictions with fundamental distinctions, France and Spain (Section II.2); and the case law is explored to show how cases with the same facts lead to different outcomes depending on the applicable legal regime (Section II.3). Next, Part III denounces the lack of

[4] Ugo Mattei, 'The Comparative Law and Economics of Penalty Clauses in Contracts' (1995) 43 Am J Comp Law 427, 428.

[5] GH Treitel, *Remedies for Breach of Contract. A Comparative Account* (Clarendon 1988) 233.

[6] Aaron Edlin and Alan Schwartz, 'Optimal Penalties in Contracts' (2003) 78 Chi-Kent L Rev 33, 37. See also Steven Walt, 'Penalty Clauses and Liquidated Damages', *Encyclopedia of Law and Economics* (2d edn, 2011) vol 6, 178, defending that the wrong conviction that courts are capable of determining the value of contract performance for the promisee explains the judicial review of liquidated damages in common law systems.

[7] The selective enforcement of these provisions is controversial not only for economic efficiency reasons, but also for reasons of fairness. Phillip R Kaplan, 'A Critique of the Penalty Limitation on Liquidated Damages' (1978) 50 S Cal L Rev 1055, 1071-72; James Arthur Weisfield, '"Keep the Change!": A Critique of the No Actual Injury Defense to Liquidated Damages' (1990) 65 Wash L Rev 977, 993-95. In the United States, the unequal treatment of very similar cases bags the question about the real set of rules applied by courts. See also Elizabeth Warren, 'Formal and Operative Rules Under Common Law and Code' (1983) 30 UCLA L Rev 898 , dealing with loss above the stipulated sum; Ann Morales Olazábal, 'Formal and Operative Rules in Overliquidation Per Se Cases' (2004) 41 Am Bus LJ 503, examining cases of absence of loss. Moreover, under highly discretionary judicial review of contract penalties, parties would prefer to directly let the ascertainment of damages to courts instead of setting them in advance. Aída Kemelmajer de Carlucci, *La Cláusula Penal* (Depalma 1981) 109.

transnational rules to secure the enforcement of penalties in international commercial contracts (Section III.1). Furthermore, Part III explains why the will of the contracting parties may be at risk in an international litigation or arbitration in the absence of coordination instruments among the several jurisdictions (Section III.2). Finally, instead of transnational rules, the statutory recognition at national level of penalties in international commercial contracts is proposed in Part III as the most feasible solution to shield the enforcement of penalties in common law jurisdictions (Section III.3).

II. THE CIVIL-COMMON LAW COMPARISON OF RULES GOVERNING PENALTIES

1. *United States: the Principle of Non-Enforcement of Penalties*

American state laws stick to the common law rule of non-enforcement of penalties. Liquidation of damages is a permissible method of limiting the defaulting promisor's liability for compensatory damages: the parties agree at the time of contracting that damages for breach will be limited to a prescribed formula. Nonetheless, if the stipulated amount entails an undue oppression on the promisor, liquidated damages may be held to be a penalty and, therefore, unenforceable. This rule has been characterized as anomalous, particularly because contracting parties lack power to bargain over their remedial rights in a legal system in which freedom of contract is a deeply rooted principle.[8]

The most illustrative case on the American common law of penalties is Banta v. Stamford Motor Co. (1914),[9] opinion which firstly delineated the test to determine whether a provision for the payment of a stipulated sum in the event of a breach of contract will be regarded as one for liquidated damages. This test was formed by three conditions:

> These conditions . . . are (1) the damages to be anticipated as resulting from the breach must be uncertain in amount or difficult to prove; (2) there must have been an intent on the part of the parties to liquidate them in advance; and (3) the amount stipulated must be a reasonable one, that is

[8] Joseph M Perillo, *Calamari and Perillo on Contracts* (6th edn, West 2009) 531; Farnsworth (n 1) 811. See also Robert A. Hillman, 'The Limits of Behavioral in Legal Analysis. The Case of Liquidated Damages' (2000) 85 Cornell L Rev 717, 733-38, arguing that agreed damages provisions must be subject to judicial scrutiny but treated like any other contract term; Larry DiMatteo, 'A Theory of Efficient Penalty: Eliminating the Law of Liquidated Damages' (2000) 38 Am Bus LJ 633, 733, defending the same claim.

[9] *Banta v Stamford Motor Co.* 92 A 665 (Conn 1914), the Supreme Courts of Errors of Connecticut upheld as a valid liquidation of damages the agreed sum of $15 a day for delay in the delivery of a luxury yacht priced at $5,500.

to say, not greatly disproportionate to the presumable loss or injury.[10]

The subsequent case law further elaborated this test in such a way that the second condition, the intent of the parties, did not survive over time;[11] and the third condition has been relaxed in the sense that the reasonableness of the amount stipulated may also be ascertained in the light of both the anticipated or actual loss, instead of only the anticipated loss at the time of contracting (in *Banta*, the so-called 'presumable loss').[12]

In addition, the difficulty of proof of loss at the moment of contracting still continues as the other relevant factor for the assessment of the reasonability of the amount stipulated,[13] although the ease of proof alone should not be purported to deem the agreed sum as a penalty.[14]

Hence, American courts apply today one single test of reasonableness with

---

[10] ibid 667-68.

[11] Restatement (Second) of Contracts § 356 cmt c (1981): 'Neither the parties' actual intention as to its validity nor their characterization of the term as one for liquidated damages or a penalty is significant in determining whether the term is valid'. See also *Wassenaar v Panos* 331 N.W.2d 357 (Wis 1983), ruling the Supreme Court of Wisconsin that the 'subjective intent of the parties has little bearing on whether the clause is objectively reasonable'; Farnsworth (n 1) 817, explaining that the inquiry goes to whether the effect of upholding the stipulation improperly compels performance; Joseph M Perillo, *Corbin on Contracts*, vol 11 (11th edn, Lexis Nexis 2005) 427, stating that, even in those jurisdictions which formally keep intention as an independent factor, intention is derived from an objective test, so this prong of the test is redundant.

[12] This clash between the classical requirement that the sum must be a genuine pre-estimate of the harm (reasonableness ex ante) and the alternative that the sum must be reasonable at the time of breach when compared with the actual harm (reasonableness ex post) remains unsolved. In this vein, the Restatements have never opted for one of them, and the Uniform Commercial Code either. Restatement (First) of Contracts § 339(1) (1932), without referring to any of the two criteria; Restatement (Second) of Contracts § 356 cmt b (1981), explicitly admitting both criteria, albeit acknowledging that each one leads to different results; UCC § 2-718(1) (1977).

[13] Restatement (Second) of Contracts § 356 cmt b (1981): 'If the difficulty of proof of loss is great, considerable latitude is allowed in the approximation of anticipated or actual harm. If, on the other hand, the difficulty of proof of loss is slight, less latitude is allowed in that approximation'.

[14] Dan B Dobbs, *Law of Remedies*, vol 3 (2nd edn, West 1993) 251, claiming that this is the right interpretation of Restatement (Second) of Contracts § 356(1) (1981); William D Hawkland, *Uniform Commercial Code Series* § 2-718:03, vol 2 (Clark Boardman Callaghan 1994), with respect to the UCC § 2-718(1) (1977), advocating that, in contrast with other common law jurisdictions, the difficulty of proof of loss at the moment of contracting has never been a requirement for the validity of the agreed damages clause in American contract law.

two elements, namely the disproportion of the agreed sum and the difficulty of proof of loss, in order to determine whether a liquidation of damages is a penalty.[15]

2. *Civil Law: the Principle of Enforcement of Penalties Subject to Reduction (France) and the Principle of Literal Enforcement of Penalties (Spain)*

The literal enforcement of conventional penalties was a rule of classical Roman law that entitled the aggrieved party to recover the agreed sum without any restriction[16]. In the XIXth century, the codification brought back the principle of literal enforcement of penalty clauses to Continental European laws.[17] In this vein, the French Civil Code, as enacted in 1804, established the literal enforcement of conventional penalties in Article 1152: *'[l]orsque la convention porte que celui qui manquera de l'exécuter paiera une certaine somme à titre de dommages-intérêts, il ne peut être alloué à l'autre partie une somme plus forte ni moindre'*.[18] The Napoleonic Code was the model for the neighboring nations (Belgium, Italy, Portugal and Spain) and their laws copied this regulation. Nonetheless, the liberal Roman principle of literal enforcement of penalties was progressively abandoned, [19] and most

[15] In comparison with the above mentioned sections of both Restatements, the UCC § 2-718(1) (1977) added a new parameter, 'the inconvenience or nonfeasibility of otherwise obtaining an adequate remedy'. Although incorporated into state laws, courts rarely apply this additional factor. In fact, the American Law Institute has declared that the factors enumerated by the UCC do not operate as independent requirements for the validity of the clause, Motion Concerning Section 2-718(1) (May 11, 2001). See also Hawkland (n 14) § 2-718:04, arguing that the inclusion of this third factor is reiterative, since the difficulty of proof of loss already points to the availability of other adequate remedies; Ian R Macneil, 'Power of Contract and Agreed Remedies' (1962) 47 Cornell L Q 495, 528, asserting that historically court decisions had conferred great importance to this third additional factor when examining the validity of agreed remedies clauses.

[16] Paulus (D. 44, 7, 44, 6).

[17] The Justinian Code (C. 7, 47) limited the amount of damages claimable to the double of the value of what had been promised. *Ius commune* was also influenced by Canon law, which considered an unjustified gain those amounts that punished with severity the party in breach. Aristides N. Hatzis, 'Having the Cake and Eating It Too. Efficient Penalty Clauses in Common Law and Civil Contract Law' (2003) 22 Int'l Rev L & Econ 381, 399. See also Reinhard Zimmermann, *The Law of Obligations: Roman Foundations of the Civilian Tradition* (OUP 1996) 95-113.

[18] 'Where an agreement provides that the party who fails to perform it will pay a certain sum as damages, the other party may not be awarded a greater or lesser sum', French Civil Code <http://www.legifrance.gouv.fr> accessed 10 March 2012.

[19] The Italian Civil Code enacted in 1942 (Article 1384); the Portuguese Civil Code enacted in 1966 (Article 812); and in Belgium, without any statutory reform, after the Belgian *Cour de cassation* Judgment, 24 November 24, the case law considers that extravagant contract penalties are against the public order and, for this reason, void. In 1975, the French Civil Code was reformed too. Law No 75-

European legislations converged on allowing the judge to moderate those contract penalties which are grossly excessive. Thus, the judicial review of penalty clauses on the grounds of equity is the solution widely accepted by Continental European laws, since Germanic legal systems do also opt for it (Austria, Germany and Switzerland).[20]

In contrast with the majority of European civil law systems, Spanish law solely allows courts to reduce the penalty whether the breach of contract has less entity that the one anticipated by the contracting parties in the provision,[21] so the judicial review on the grounds of equity is excluded.[22]

---

597 of 9 July 1975, JO 10 July 1975 7076, added a second paragraph to Article 1152: '*Néanmoins, le juge peut modérer ou augmenter la peine qui avait été convenue, si elle est manifestement excessive ou dérisoire. Toute stipulation contraire sera réputée non écrite*' ('Nevertheless, the judge may moderate or increase the agreed penalty, where it is obviously excessive or ridiculously low. Any stipulation to the contrary shall be deemed unwritten', French Civil Code <http://www.legifrance.gouv.fr> accessed 10 March 2012). This article let the judge to increase or decrease a penalty found to be disproportionate. Moreover, Law No 75-597 reformed Article 1231, governing the reduction of the penalty in case of partial performance, stating that Article 1152 was also applicable: '*Lorsque l'engagement a été exécuté en partie, la peine convenue peut être diminuée par le juge à proportion de l'intérêt que l'exécution partielle a procuré au créancier, sans préjudice de l'application de l'article 1152. Toute stipulation contraire sera réputée non écrite*' ('Where an undertaking has been performed in part, the agreed penalty may be lessened by the judge in proportion to the interest which the part performance has procured for the creditor, without prejudice to the application of Article 1152. Any stipulation to the contrary shall be deemed not written', French Civil Code <http://www.legifrance.gouv.fr> accessed 10 March 2012). Therefore, the same penalty might be reviewed by a French judge on the grounds of partial performance and on the grounds of equity.

[20] German Civil Code (BGB § 343), although the German Commercial Code (HGB § 348) excludes contracts between professionals in the scope of their activity. Both Austrian law (§ 1336.2 Austrian Civil Code, ABGB) and Swiss law (Article 163-3 *Code des obligations*) admit the judicial review of disproportionate penalties too, but without a different regime for commercial contracts.

[21] Fernando Gómez Pomar, 'El Incumplimiento Contractual en Derecho Español' (2007) 3 InDret 29 <http://www.indret.com/pdf/466_es.pdf> accessed 10 March 2012. However, in lieu of the Spanish Civil Code, Navarrese civil law may apply, which is the only particular civil law of the Autonomous Communities with its own rules in the field of contract penalties. Actually, under Navarrese civil law, the coercive function of the penalty is especially protected, since the New Navarrese Code of Laws (Article 518) expressly provides that 'the agreed penalty should not be reduced by judicial discretion', so the penalty would not be adjusted on any ground, Navarra Superior Court Judgments, 27 January 2004 (RJ, No 2668), and 9 November 2005 (RJ, No 2006\377). See also José Ignacio Bonet Sánchez, 'La cláusula penal' in Ubaldo Nieto Carol and José Ignacio Bonet Sánchez (eds), *Tratado de Garantías en la Contratación Mercantil*, vol 1 (Civitas 1996) 887, 964-65. Recall that the Superior Courts of those Autonomous Communities with particular civil law have jurisdiction

The Spanish Civil Code (Article 1154) imposes on the judge the duty to moderate the penalty if, and only if, the undertaking has been partially or irregularly performed.[23] To moderate the penalty, the judge must assess the proportion between the actual performance and the performance that would have barred the claim of the penalty.[24]

Albeit the above mentioned differences concerning the grounds of the judicial review, European civil law systems share the same concept of penalty clause: a provision seeking to deter breach by requiring the payment of extra-compensatory damages.

Beyond the grounds of the judicial review, which serve to classify a civil law system as one of enforcement of penalties subject to reduction or one of literal enforcement of penalties, there exist other minor but significant differences among the several penalty clause regimes pertaining to the civil law tradition. Next, the French and the Spanish law of penalties are compared in order to point out the most basic traits of each of them.

---

to adjudicate cases in which arise an issue related to the corresponding particular civil law.

[22] The Spanish Supreme Court has constantly rejected the judicial review of penalty clauses on the grounds of equity, STS, 15 October 2008 (RJ, No 5692). However, among other relevant changes, a tentative draft bill aims to explicitly introduce the judicial review on the grounds of equity, *Comisión General de Codificación, Propuesta de Anteproyecto de Ley de Modernización del Derecho de Obligaciones y Contratos* (2009), Article 1150. Within Spanish legal scholars, the majority position has always defended the need of a law reform that allows the judicial review of penalty clauses on the grounds of equity, María Dolores Mas Badía, *La Revisión Judicial de las Cláusulas Penales* (Tirant Lo Blanch 1995) 216; Isabel Espín Alba, *Cláusula Penal. Especial Referencia a la Moderación de la Pena* (Marcial Pons 1997) 86. A minority position advocated that the current Article 1154 of the Spanish Civil Code (n 23) embraces judicial review on the grounds of equity, since the single requirement for the reduction is the disproportion between the penalty and the actual harm, Francisco Jordano Fraga, *La Resolución por Incumplimiento en la Compraventa Inmobiliaria. Estudio Jurisprudencial del Artículo 1504 del Código Civil* (Civitas 1992) 199-200; José Miguel Rodríguez Tapia, 'Sobre la Cláusula Penal en el Código Civil' (1993), 46 Anuario de Derecho Civil 511, 578-80.

[23] Article 1154: '*El Juez modificará equitativamente la pena cuando la obligación principal hubiera sido en parte o irregularmente cumplida por el deudor*' ('The Judge shall equitably modify the penalty where the principal obligation should have been performed partially or irregularly by the debtor', Spanish Civil Code <http://www.mjusticia.es/cs/Satellite/es/1215198252168/DetalleInformacion.html> accessed 10 March 2012.

[24] Manuel Albaladejo García, *Comentarios al Código Civil y a las Compilaciones Forales*, vol 15(2) (Edersa 1983) 486. In consequence, there would be no moderation if the penalty was agreed upon the partial performance actually occurred, STS, 14 September 2007 (RJ, No 5307).

a) Even though the judicial review under Spanish law is much more restricted, the judicial intervention of the penalty is still exceptional in French law, because the disproportion must be an abuse of the coercive function, being obviously excessive, and having no justification.[25]

b) While in Spanish law the judicial intervention of the penalty may consist only in the reduction of the sum stipulated,[26] in French law the judge may reduce the penalty if manifestly excessive, or increase it if ridiculously low.[27]

c) If applicable, Spanish courts must reduce the penalty,[28] although the question about the possibility of an ex officio judicial intervention is more debatable.[29] On the contrary, the French Civil Code (Articles 1152 and 1231) authorizes courts to exercise their judicial discretion when reviewing the '*clause pénale*', once it has been determined that the sum stipulated is manifestly excessive or pitiful and also in the event of partial performance. Furthermore, in French law, the adjustment of the sum stipulated on the judge's own motion is statutorily granted, [30] which reinforces the

[25] Geneviève Viney and Patrice Jourdain, *Traité de Droit Civil. Les Effets de la Responsabilité* 486-89 (2d edn, LGDJ 2001).
[26] Cristina Guilarte Martín-Calero, *La Moderación de la Culpa por los Tribunales (Estudio Doctrinal y Jurisprudencial)* (Lex Nova 1999) 139.
[27] Article 1152 of the French Civil Code (n 19). Actually, this judicial power to increase the agreed sum when ridiculously low constitutes a distinctive feature of French law in comparison with other European civil law systems. Unlike other regimes of contract penalties from the decade of the 70s (n 41) only the 1975 reform of the French Civil Code grants this faculty to the courts. Jean Thilmany, 'Fonctions et Révisibilité Des Clauses Pénales en Droit Comparé' (1980) 32 Revue Internationale de Droit Comparé 17, 40-1.
[28] Article 1154 of the Spanish Civil Code (n 23). The Spanish Supreme Court finally settled this historical controversy with consistent case law since mid 80s, STS, 7 February 2002 (RJ, No 2887).
[29] The Spanish Supreme Court has ruled so in some scattered decisions, being the last one STS, 12 December 1996 (RJ, No 8976). However, there is a tension with the rules of civil procedure, since an ex officio judicial intervention would imply a judicial action beyond the claims raised by the litigants, Luis Díez-Picazo y Ponce de León, *Fundamentos del Derecho Civil Patrimonial*, vol 2 (6th edn, Civitas 2008) 468. See also Charles Calleros, 'Punitive Damages, Liquidated Damages, and Clauses Pénales in Contract Actions: A Comparative Analysis of the American Common Law and the French Civil Code' (2006) 32 Brooklyn J Int'l L 67, 104-5, pointing out the same concern with respect to ex officio judicial review of penalties in French law, as mentioned below.
[30] Law No 85-1097 of 11 October 1985, JO 15 October 1985 11982, amended both Articles 1152 and 1231 of the French Civil Code, introducing the expression '*méme d'office*' ('even of his own motion').

discretionary judicial review of penalties.

In addition, in both legal systems, the question whether to adjust the sum stipulated and in which degree are reviewable by the appellate court but not by the highest court of ordinary jurisdiction, since each of these issues is considered a matter of fact instead of a matter of law. Therefore, the Spanish Supreme Court may decide these issues only on the basis of the prior finding that the lower court erred in qualifying promisor's performance.[31] In this regard, the French *Cour de cassation* balances the stronger discretionary judicial review of penalties with a demanding requirement of accountability, reversing those judgments which alter the sum stipulated without articulating the factual reasons why the amount set fits into the above mentioned category of 'manifestly excessive'.[32]

d) The French and the Spanish law of penalties have in common the application of an objective, retrospective test: despite not being entirely consistent,[33] French courts compare the sum stipulated with the actual damages,[34] and Spanish courts the breach anticipated in the provision with the actual breach.[35] Whereas, the American common law of penalties and the Uniform Commercial Code provide not only the use of the applicable test retrospectively (reasonableness ex post), but also prospectively (reasonableness ex ante).[36] Notwithstanding, in French law, the breaching party's bad faith in the performance is a relevant factor in the determination of whether a penalty is 'manifestly excessive',[37] unlike Spanish law, since this argument does not have any relevance.[38]

---

[31] STS, 20 December 2006 (RJ, No 2007/388), and STS, 20 September 2006 (RJ, No 8401).

[32] Cass 3e civ, 12 January 2011, Pourvoi No 09-70.262, <http://www.legifrance.gouv.fr/initRechJuriJudi.do> accessed 10 March 2012; Cass 3e civ, 13 July 2010, Pourvoi No 09-68.191 <http://www.legifrance.gouv.fr/initRechJuriJudi.do> accessed 10 March 2012; Cass 3e civ, 12 January 2010, Pourvoi No 09-11.856 <http://www.legifrance.gouv.fr/initRechJuriJudi.do> accessed 10 March 2012; Cass 1e civ, 28 November 2007, Pourvoi No 05-17.927 <http://www.legifrance.gouv.fr/initRechJuriJudi.do> accessed 10 March 2012.

[33] Calleros (n 29) 105.

[34] Denis Mazeaud, *La Notion de Clause Pénale* (LGDJ 1998) 57-58.

[35] Gómez Pomar (n 21).

[36] Restatement (First) of Contracts § 339(1) (1932); Restatement (Second) of Contracts § 356 cmt b (1981); UCC § 2-718(1) (1977) (n 12).

[37] Calleros (n 29) 106, specifying that the *Cour de cassation* rejects the behavior of the parties as the sole basis to find a penalty manifestly excessive, Cass com, 11 February 1997, Bull civ II, No 47.

[38] However, some scholars have defended the use of the argument of the bad faith

e) Lastly, another significant difference between the French and the Spanish law of penalties is that the former bans the cumulative penalty, i.e. the aggrieved party is not jointly entitled to the payment of penalty and the performance of the obligation,[39] while the latter allows the cumulative penalty, as long as this right has been clearly granted.[40] French law makes a single exception: the penalty for breach due to delay, which does not properly constitute a cumulative penalty, because the creditor will never obtain a timely performance of the already lately performed obligation. In the context of European civil law systems, the cumulative penalty is not a singularity of Spanish law.[41]

---

to expand the grounds on which the reduction of the penalty is permitted, Jaime Santos Briz, 'Comentario a los arts. 1152 a 1155 CC' in Ignacio Sierra Gil de la Cuesta (ed), *Comentario al Código Civil*, vol 6 (Bosch 2000) 289, 296-97. Against, María Corona Quesada González, 'Estudio de la Jurisprudencia del Tribunal Supremo sobre la Pena Convencional' (2003) 14 Aranzadi Civil 45, arguing that the claim of the sum stipulated may not be deemed against the good faith, since contract penalties are allowed in Spanish law.

[39] Article 1229 of the French Civil Code: '*Il [le créancier] ne peut demander en même temps le principal et la peine, à moins qu'elle n'ait été stipulée pour le simple retard*' ('He [the creditor] may not claim at the same time the principal and the penalty, unless it was stipulated for a mere delay', French Civil Code <http://www.legifrance.gouv.fr> accessed 10 March 2012).

[40] Article 1153 of the Spanish Civil Code: '*Tampoco el acreedor podrá exigir conjuntamente el cumplimiento de la obligación y la satisfacción de la pena, sin que esta facultad le haya sido claramente otorgada*' ('Neither may the creditor request jointly the performance of the obligation and the payment of the penalty, unless this power has been clearly granted', Spanish Civil Code <http://www.mjusticia.es/cs/Satellite/es/1215198252168/DetalleInformacion.html> accessed 10 March 2012).

[41] German Civil Code (BGB § 341(1)), allowing the claim of performance in addition to the payable penalty when the penalty was promised for improper performance. On the contrary, following the French solution, the Italian Civil Code (Article 1383), the Portuguese Civil Code (Article 811), and the Austrian Civil Code (§ 1336.1 ABGB), including this latter the non-compliance with the promised place of performance too. In accordance with French law, the mandatory prohibition of the cumulative penalty is the solution recommended by the Council of Europe, Committee of Ministers Resolution (78) 3 Relating to Penal Clauses in Civil Law (1978) [hereinafter Council of Europe Resolution (78) 3], Article 2: 'The promisee may not obtain concurrently performance of the principal obligation, as specified in the contract, and payment of the sum stipulated in the penal clause unless that sum was stipulated for delayed performance. Any stipulation to the contrary shall be void'. In fact, the cumulative penalty is not permitted in the tentative draft bill for the reform of the Spanish Civil Code (n 22) Article 1149; Isabel Arana de la Fuente, 'Algunas Precisiones sobre la Reforma de la Cláusula Penal en la Propuesta de Modernización del Código Civil en Materia de Obligaciones y Contratos' (2010) 4

On a comparative account limited to Western Europe, French law cannot be generalized, and deemed as the European civil law model of contract penalties, due to the judicial power of increasing an unreasonably small agreed sum, since usually the penalty may only be reduced. Nevertheless, French law features the other characteristics of the wide majority of European civil laws: (1) the validity of contract penalties, which may have the effect of coercing a party to perform her obligation; (2) the judicial review of penalties on the grounds of equity as a discretionary faculty, based on a retrospective test considering the actual harm, or on the grounds of partial performance; and (3) the promisee's entitlement either to the penalty or to specific performance, with the exception of delay, being deprived of claiming statutory damages.

Regarding this third common characteristic, German law is neither representative: not only the cumulative penalty is permitted,[42] but also the promisee is entitled to claim statutory damages, operating the penalty as

---

InDret 8-9 <http://www.indret.com/pdf/775_es.pdf> accessed 10 March 2012. Notwithstanding, shortly before the Council of Europe Resolution (78) 3, the Common Provisions Annexed to the Benelux Convention on Penalty Clauses (1973) Article 2(1)-(2), contained the exclusion of the cumulative penalty but as a default rule instead of mandatory, being excludable by the parties' agreement, Thilmany (n 27) 41. The exclusion of the cumulative penalty unless otherwise stipulated by the parties was also the solution adopted by the UN Commission on International Trade Law [UNCITRAL] in the Text of Draft Uniform Rules on Liquidated Damages and Penalty Clauses, together with a Commentary thereon (1981) UN Doc A/CN.9/218 [hereinafter UNCITRAL Draft] Article E: '(2) Where the agreed sum is to be recoverable or forfeited on non-performance, or defective performance other than delay, the obligee is entitled either to performance, or to recover or forfeit the agreed sum, unless the agreed sum cannot reasonably be regarded as a substitute for performance. (3) The rules set forth above shall not prejudice any contrary agreement made by the parties'. Despite acknowledging that the cumulation of the two remedies might unjustly enrich the obligee in some circumstances, the Revised Text of Draft Uniform Rules on Liquidated Damages and Penalty Clauses (1983) UN Doc A/CN.9/235 [hereinafter UNCITRAL Revised Draft], this revised draft of uniform rules does not follow the recommendation of the Council of Europe: Article E(3) was deleted, but Article E(2) was amended by including an exception under which the obligee is entitled to performance and the agreed sum when proving that the later cannot reasonably substitute the former, and Article X was added, providing that '[t]he parties may by agreement only derogate from or vary the effect of articles D, E and F of this (Convention)(law)'. See also the final endorsement of this solution, contrary to the recommendation of the Council of Europe, Uniform Rules on Contract Clauses for an Agreed Sum Due upon Failure of Performance (1983) UN Doc A/CN.9/243 [hereinafter UNCITRAL Uniform Rules] Annex I, Articles 6(2) and 9.

[42] German Civil Code (BGB § 341(1)).

the minimum amount of damages.[43]

In sum, in spite of the common traits already mentioned, there are not uniform rules governing contract penalties in Continental Europe, and historically there has not been a real political will of unifying contract law within the European Union,[44] even though some signs of change in 2010.[45]

---

[43] ibid BGB §§ 340(2) and 341(2), both referring to the obligee's assertion of additional damage in cases of non-performance and defective performance. Swiss law (Article 161-2 *Code des obligations*) also allows the recovery of the additional damage.

[44] The European Union lacks a general legislative competence in contract law, since its competence is limited to those areas related to consumer protection, which has been extensively exercised (the so-called consumer *acquis*). The enactment of a European Civil Code may be perceived as an expression of European identity, but this view is conflicting with the widespread opinion that national codes reflect their own national legal values and legal cultures, factor which explains the political opposition to move towards to the unification of private law, Simon Whittaker, '*The 'Draft Common Frame of Reference'. An Assessment*' (2008) 23-4 <http://www.justice.gov.uk/publications/eu-contract-law-common-frame-reference.htm> accessed 15 May 2011. The origin of the Europeanization of private law has scholarly roots, since the 1980s academics from different European countries formed research groups to embark on the harmonization of private law. Despite the shy institutional support that firstly arrived from the European Parliament, the series of Resolutions from 1986 to 2003, the Commission on European Contract Law, chaired by Professor Ole Lando, elaborated the Principles of European Contract Law [hereinafter PECL], meant to provide black letter rules of soft law using the drafting style of a restatement rather than a code in the civil law meaning of the term, Ole Lando and Huge Beale (eds), *Principles of European Contract Law, Parts I and II, Combined and Revised* (Kluwer Law International 2000); Ole Lando, Eric Clive, André Prüm and Reinhard Zimmermann (eds), *Principles of European Contract Law, Part III* (Kluwer Law International 2003). The second great achievement of this arduous process was the Draft Common Frame of Reference [hereinafter DCFR], commissioned by the European Commission, which combined rules from the PECL, rules from the existing European *acquis*, and rules from several teams of academics, Christian von Bar, Eric Clive and Hans Schulte-Nölke (eds), *Principles, Definitions and Model Rules of European Private Law: Draft Common Frame of Reference* (Sellier 2008). See also Luisa Antoniolli and Francesca Fiorentini (eds), *A Factual Assessment of the Draft Common Frame of Reference* (Sellier 2010) 7-10.

[45] These signs of change were the setting up of the Expert Group to review the DCFR for the European legislation harmonization in the matter of contract law, Commission Decision No 2010/233 [2010] OJ L105 119, and the launch of a public consultation, 'Commission Green Paper on Policy Options for Progress towards a European Contract Law for Consumers and Businesses' COM (2010) 348 final. See also Fernando Gómez Pomar and Marian Gili Saldaña, 'El Futuro Instrumento Opcional del Derecho Contractual Europeo: Una Breve Introducción a las Cuestiones de Formación, Interpretación, Contenido y Efectos' (2012) 1 InDret 4-13 <http://www.indret.com/pdf/872_es.pdf> accessed 16 March 2012.

These signs of change have led to a highly mature and innovative proposal of contract law harmonization, the Proposal for a Regulation on a Common European Sales Law,[46] the scope of which are those aspects which pose real problems in cross-border transactions without extending to aspects that are best addressed by national laws. Notwithstanding, this Common European Sales Law proposed by the Commission does not deal with contract penalties.

3. *Case Law: Same Facts Leading to Different Outcome Across the Jurisdictions*

An array of cases is presented in this Section in order to illustrate how much differ the three jurisdictions examined (United States, France and Spain) when adjudicating an issue involved in a dispute concerning an agreement for the payment of a fixed sum on breach of contract, regardless it is a liquidation of damages or a contract penalty. The issues discussed are the solutions to a disproportionate agreed sum, an unreasonably small agreed sum, and the promisee's entitlement to both the agreed sum and specific performance.

a. Disproportionate Agreed Sum

In American state laws, parties are left in a climate of uncertainty because courts may tackle differently the single test of reasonableness.[47] For instance, in Walter Implement, Inc. v. Focht,[48] the Washington Supreme Court held that a liquidated damages provision requiring the 20% of the outstanding rental payments in a lease of farm equipment was unenforceable, although liquidated damages amounted to $8,645.06 and actual damages were approximately $15,000.[49] In fact, the so-declared penalty, on the basis that the amount of liquidated damages was not reasonably related to the damages, and that the actual damages were easily ascertainable, showing a downward deviation of a 40%.

---

[46] Commission, 'Proposal for a Regulation of the European Parliament and of the Council on a Common European Sales Law' COM (2011) 635 final. This proposal made by the Commission coincides with the widespread thinking according to which the most likely is that an European Regulation adopts, totally or partially, a harmonized body of rules as an optional instrument which contracting parties may choose as the applicable law to their contract in order to opt out of their national laws (the so-called 'blue button'), Hans Schulte-Nölke, 'EC Law on the Formation of Contract—from the Common Frame of Reference to the "Blue Button"' (2007) 3 European Review of Contract Law 332, 348-49.

[47] The selective enforcement of these provisions raises not only efficiency concerns but also fairness concerns. See n 5 and n 7.

[48] *Walter Implement v Focht* 730 P.2d 1340 (Wash 1987).

[49] ibid 1345, calculation made by the Washington Supreme Court in the last paragraph of the decision.

On the contrary, in Bruce Builders, Inc. v. Goodwin,[50] the Court of Appeal of Florida upheld a liquidated damages provision under which the purchaser of eight lots of real estate for a total price of $173,800 forfeited the escrow deposit of $7,200, even though the seller made a net profit of approximately $2,500.[51] In *Bruce Builders*, the Court of Appeal of Florida argued that the amount of liquidated damages did not shock the court's conscience (the deposit was about the 4% of the total price), and that damages from breach were not ascertainable at the time of contracting.[52]

In French law, despite the Civil Code (Article 1152),[53] the judicial intervention is exceptional, provided that the penalty constitutes an abuse of the coercive function without justification.[54] The *Cour de cassation* is prone to reverse those judgments from the appellate courts in which a penalty is declared 'manifestly excessive' and, accordingly, moderated insofar as no factual reasons are articulated to support the application of Article 1152.[55] Nonetheless, Article 1152 may apply to reduce a disproportionate penalty on the grounds of equity:[56] for example, the *Cour de cassation* affirmed the appellate court decision to reduce from €30,000 to €22,900 the penalty stipulated in a contract for the sale of a building under the condition precedent of obtaining a loan.[57] Only €22,900 of the total amount of the earnest payment had to be forfeited, but the buyers claimed a larger reduction, which the *Cour de cassation* found to have been adequately denied due to their passive behavior,[58] since the buyers had not met the deadline even after a two-years extension, in spite of the quick sale

[50] *Bruce Builders, Inc. v Goodwin* 317 So. 2d 868 (Fla Dist Ct App 1975).
[51] ibid 870.
[52] Unlike English law, American state laws do not exclude all forfeiture clauses from the law of penalties. In English law the prohibition of penalties is deemed exceptional and, therefore, the penalty rule has to be applied restrictively, HG Beale and Joseph Chitty, *Chitty on Contracts*, vol 1 (30th edn, Sweet & Maxwell 2008) 1700, paras 136-138. See also Law Commission, *Penalty Clauses and Forfeiture of Monies Paid* (Law Com No 61, 1975), highlighting that the distinction of these figures leads to discrepancies.
[53] Article 1152 of the French Civil Code (n 19).
[54] Viney and Jourdain (n 25).
[55] See n 32.
[56] Article 1152 of the French Civil Code is often applied in conjunction with consumer protection rules, for instance, Cass 2e civ, 5 February 2009, Bull civ II, No 38 <http://www.legifrance.gouv.fr/initRechJuriJudi.do> accessed 10 March 2012. However, its wider scope embraces disputes in which the parties involved need not be consumers.
[57] Cass 3e civ, 30 January 2008, Bull civ III, No 15 <http://www.legifrance.gouv.fr/initRechJuriJudi.do> accessed 10 March 2012.
[58] Nevertheless, the behaviors of the parties can never be the sole basis to hold a penalty manifestly excessive. See n 37.

of the property at a good price.

On the other hand, the Spanish Supreme Court has even ruled that the fact that the sum stipulated is disproportionate or outrageous is irrelevant for a penalty to be reduced in the light of the Civil Code (Article 1154).[59] The Supreme Court is also reluctant to endorse other legal grounds for the reduction of excessive penalties, in spite of the serious scholarly attempts to find alternatives out of the scope of Article 1154.[60] In Spanish law, the earnest payments that operate bilaterally if any party breaches, i.e. the money is forfeitable by the recipient or the double amount is returnable to the depositor, are considered penalties and, in consequence, those sums

[59] Article 1154 of Spanish Civil Code (n 23). STS, 17 October 2007 (RJ, No 7307), the Spanish Supreme Court literally enforced the delay penalty included in a separation agreement, according to which the husband was entitled to €90.15 per day while his wife remains at the family home, leading to €72,211.60 due to 801 days of delay. See also STS, 29 November 1997 (RJ, No 8441), in a contract executed in 1988 and priced at 12,000,000 pesetas (€72,121.45), a defendant seller is bound to the penalty for delay in delivery of the property, 300,000 pesetas (€1,803.04) per day, which amounted to 19,800,000 pesetas (€119,000.40) due to 66 days of delay. Against, an isolated judgment dating back to the 50s, STS, 5 November 1956 (RJ, No 3805).

[60] Alternative legal grounds that scholar have suggested to reduce excessive penalties are the following: (1) Article 1103 of the Spanish Civil Code, courts may moderate the contract liability arising from negligence on a case-by-case basis, Javier Dávila González, *La Obligación con Cláusula Penal* (Montecorvo 1992) 473, favoring this solution; whereas, Mas Badía (n 22) 229-30 argues that the stipulation of the contract penalty excludes the application of the general contract liability rules; (2) Article 1258 of the Spanish Civil Code, parties should perform their obligations in accordance with good faith, see n 38; (3) Article 1275 of the Spanish civil Code, unjust enrichment, whenever there is an abuse of the coercive function or the penalty is not intended to coerce performance, Mas Badía (n 22) 232; (4) Article 7.2 of the Spanish Civil Code, abuse of rights, Mas Badía (n 22) 237; (5) *rebus sic stantibus* clause, the fulfillment of the contract becomes excessively burdensome due to unforeseen circumstances, Quesada González (n 38) 47, admitting the theoretical viability of this ground but stressing its highly unlikely application; Pablo Salvador Coderch, 'Alteración de Circunstancias en el Article 1213 de la Propuesta de Modernización del Código Civil en Materia de Obligaciones y Contratos' (2009) 4 InDret 8 <http://www.indret.com/pdf/687_es.pdf> accessed 10 March 2012, emphasizing that the application is still extremely restrictive under the tentative draft bill for the reform of the Spanish Civil Code (n 19) Article 1213. The Spanish Supreme Court has ruled about the application of some of the enumerated legal grounds to reduce excessive penalties, except good faith (Article 1258) and *rebus sic stantibus* clause, rendering inapplicable both unjust enrichment and abuse of rights, STS, 19 February 1985 (RJ, No 816), STS, 26 December 1990 (RJ, No 10374), and STS, 4 February 1991 (RJ, No 704). However, the Supreme Court exceptionally affirmed the reduction of an excessive penalty on account of Article 1103 in STS, 19 February 1990 (RJ, No 700).

are subject to judicial reduction under Article 1154.[61] Therefore, the Supreme Court held that Article 1154 was applicable to an earnest money agreement, but refused to moderate the 8,000,000 pesetas (€48.080,97) that the seller owed to the buyer because of the severity of the breach, the prior sale of the apartment to a third party.[62]

b. Unreasonably Small Agreed Sum

In American state laws, the general rule is that the sum stipulated in a valid liquidated damages clause limits the liability arising from promisor's breach,[63] having the aggrieved promisee no other remedy available for the recoverability of the portion of damages over the sum stipulated.[64] Nevertheless, with respect to unreasonably small agreed sums, an exception is made on the basis of the unconscionability doctrine:[65] '[a] term that fixes an unreasonably small amount as damages may be unenforceable as unconscionable'.[66] In this vein, in Roscoe-Gill v. Newman,[67] the Court of Appeals of Arizona reviews an unreasonably small liquidated amount in the light of the unconscionability doctrine, even

[61] Albeit the fundamental differences of the deposit, and the entitlement of the aggrieved party to claim the statutory damages exceeding the earnest payment, Silvia Díaz Alabart, 'Las arras (I)' (1996) 80 Revista de Derecho Privado 3, 37.

[62] STS, 10 October 2006 (RJ, No 8405). See Jordano Fraga (n 23) 187, defending that Article 1154 is applicable to this kind of earnest money agreement; María Corona Quesada González, 'Estudio de la Jurisprudencia del Tribunal Supremo sobre las Arras' (2003) 5 Aranzadi Civil 8, against the application of Article 1154 to these agreements.

[63] Perillo (n 11) 446. To illustrate the strict application of this general rule, see also *Wechsler v Hunt Health Sys.* 330 F Supp 2d 383, 426-27 (SDNY 2004), case in which the District Court declined to award early termination damages to the injured party in addition to the liquidated damages for the same concept.

[64] Perillo (n 8) 534, arguing that granting the aggrieved party any other remedy, even contractually conferred by the party in breach, implies that a valid liquidated damages clause may not constitute a reasonable forecast of the harm.

[65] A generally applicable doctrine in contract law, to which both Restatement (Second) of Contracts § 208 (1981) and UCC § 2-302 (1977) refer as a basis to hold unenforceable a contract term or all the contract. However, none of the provisions cited provides a definition of 'unconscionability', which has to be found in *Williams v Walker-Thomas Furniture Co.* 350 F.2d 445, 449 (DC Cir 1965), case in which the Court of Appeals stated that '[u]nconscionability has generally been recognized to include an absence of meaningful choice on the part of one of the parties together with contract terms which are unreasonably favorable to the other party'.

[66] Restatement (Second) of Contracts § 356 cmt a (1981). Also, UCC § 2-718 cmt 1 (1977) explicitly welcomes the unconscionability doctrine in the realm of liquidated damages. See Hillman (n 8) 738, n 128, in favor of a validity inquiry on the grounds of the generally applicable doctrines of contract law such as unconscionability.

[67] *Roscoe-Gill v Newman* 1997 Ariz App LEXIS 32.

though the Court concludes that the facts alleged by the seller failed to render the liquidated damages clause unconscionable: in a contract for the sale of a ranch at $380,000, the buyer forfeited the $5,000 paid as earnest money in escrow in the event of default, while the seller sought excess damages that amounted to $140,000.[68]

In French law, Article 1152 may also apply to increase a ridiculously low penalty ('*dérisoire peine*') on the grounds of equity,[69] as explained in Section I.2. In this regard, the *Cour d'appel* of Pau refuses to deem ridiculously low the penalty of €24,880 payable to the real estate agency for the breach of the exclusive right of sale, since this figure represents more than the 16% of the price at which the owners themselves sold the property (€152,400).[70] In its lawsuit, the real estate agency claim additional damages amounting to €48,000 to compensate its financial loss.

Unlike French law, the governing principle of Spanish law is the literal enforcement of the contract penalty (Article 1152), with the only exception of partial performance,[71] therefore courts are by no means allowed to increase an unreasonably small agreed sum. However, among Spanish legal scholars, the majority position has been to defend that the aggrieved party is entitled to be fully compensated whether the breach is willful instead of negligent, because of the prohibition to waive claims for damages arising from willful misconduct (Article 1102)[72] would render the contract penalty unenforceable under such circumstances.[73] Alternatively, a minority position of scholars has sustained another solution: regardless of Article 1102, the penalty is enforceable despite the willfulness of the breach, but

[68] Plaintiff seller claimed as damages the $120,000 difference between the original sale price of $380,000 and the actual sale price of $260,000, plus $20,000 in interest and lost discounts, legal fees, and payments for taking care of the ranch.
[69] Article 1152 (n 19).
[70] Cour d'appel Pau, 1e ch, 13 May 2008 <http://www.legifrance.gouv.fr/initRechJuriJudi.do> accessed 10 March 2012.
[71] Article 1154 of the Spanish Civil Code (n 23).
[72] Article 1102 of the Spanish Civil Code: '*La responsabilidad procedente del dolo es exigible en todas las obligaciones. La renuncia de la acción para hacerla efectiva es nula*' ('Liability arising from willful misconduct is enforceable for all obligations. Waiver of the action to enforce it shall be null and void', Spanish Civil Code <http://www.mjusticia.es/cs/Satellite/es/1215198252168/DetalleInformacion.html> accessed 10 March 2012.
[73] Ángel Carrasco Perera, 'Comentario al artículo 1.102 CC' in Manuel Albaladejo García (dir), *Comentarios al Código Civil y a las Compilaciones Forales*, vol 15(1) (Edersa 1989) 444, 468, defending that Article 1102 forbids any agreed sum below the statutory damages; Ferran Badosa Coll, *La Diligencia y la Culpa del Deudor en la Obligación Civil* (Publicaciones del Real Colegio de España 1987) 718, claiming that the liability arising from intentional breach should always be aggravated.

the promisee is entitled to recover the excess damages.[74] Anyhow, even if the breach is intentional, Spanish case law confirms the literal enforcement of the penalty, which bars the promisee's claim to recover excess damages.[75]

c. Promisee's Entitlement to Both the Agreed Sum and Specific Performance

In American state laws, the validity of the 'nonexclusive clauses' is under discussion. 'Nonexclusive clauses' are those contractual provisions that entitle the promisee to the liquidated amount and any other remedy, either specific performance or the general compensatory damages. Nevertheless, the former combination deserves a different treatment than the latter.

With respect to specific performance, the nonexclusive clause has no effect, since the parties may not alter the restrictive availability of this equitable remedy, which can be granted by courts anyway, except the parties intended the liquidated damages to be the exclusive remedy for breach.[76] In Stokes v. Moore, the Supreme Court of Alabama enforces a $500 liquidated damages provision for the violation by an employee of a covenant not to compete against his employers in the city of Mobile for one year after the contract termination, and the Court also grants temporary injunctive relief, because of the finding that parties never intended liquidated damages as the sole remedy.[77]

However, the aggrieved party will never be entitled to both the liquidated amount and the general compensatory damages. A stipulation with such content is held a penalty, because it is disproportionately beneficial for the promisee. For instance, in Schrenko v. Regnante, the Appeals Court of Massachusetts declared to be a penalty the clause that provided for

[74] Dávila González (n 59) 363, following the same understanding than José María Manresa, *Comentarios al Código Civil Español*, vol 7 (2nd edn, Imprenta de la Revista de Legislación 1907) 239. See also Rodríguez Tapia (n 22) 572-78, arguing that excess damages arising from any breach, intentional or negligent, should be recoverable.

[75] The ruling of the Spanish Supreme Court since mid 80s, STS, 7 July 1998 (RJ, No 5556), STS 20 February 1989 (RJ, No 1212), and STS, 23 May 1997 (RJ, No 4322).

[76] Dobbs (n 14) 189-201, explaining the narrow scope of this remedy in American state laws.

[77] *Stokes v. Moores* 77 So 2d 331, 335 (Ala 1955): 'the contract for liquidated damages will not operate to prevent an injunction . . . unless it appears from the contract that the provision for liquidated damages was intended to be the exclusive remedy for its breach'.

forfeiture of a $16,000 deposit in the event of a buyers' breach plus damages, in a failed contract for the sale of real estate in which the sellers received $25,000 more than the price the buyers would have paid.[78]

In French law, the prohibition of the cumulative penalty controls, as explained in Section I.2.e), with the single exception of penalties for delay (Article 1129).[79] Therefore, French courts will never grant to the aggrieved party both the penalty and the performance of the breached obligation. Logically, the *Cour de cassation* has ruled that this prohibition necessarily applies only with respect to the same obligation. In other words, if the promisor has breached two obligations, the injured promisse may claim the penalty arising from the breach of one obligation, and the performance of the other. For example, in a computer equipment lease contract, the *Cour de cassation* affirmed a judgment in which, in addition to the amount of unpaid rent, the lessee in breach was ordered to pay the agreed compensation in the event of termination.[80]

On the contrary, in Spanish law, cumulative penalties are permitted (Article 1553),[81] so the aggrieved party may be jointly entitled to the payment of the penalty and the performance of the obligation. Far from being the default rule, the high degree of coercion on the obligor and the wording of Article 1553 ('unless this power has been clearly granted') make that a cumulative penalty is never presumed. In this vein, if contract penalties, as an exception to the general rules of contract law, deserve a narrow interpretation, the interpretation of cumulative penalties should be even narrower. In accordance with this much stricter standard, the Spanish Supreme Court upheld as cumulative penalty a clause providing the additional payment of 15,000,000 pesetas (€90,151.82) in the event of delay or non-performance of the construction of two naves.[82]

III. HOW TO SECURE THE ENFORCEMENT OF PENALTIES IN INTERNATIONAL COMMERCIAL CONTRACTS

The General Assembly of the United Nations, when recommending the

[78] *Schrenko v Regnante* 27 Mass App Ct 282 (Mass App Ct.1989), buyers recovered the $16,000 deposit and sellers were not granted any compensation, since there was no loss at all, despite the sellers' damages claim for $18,831.62 ($10,581.62 out-of-pocket expenses attributable to the buyers' default, and the $8,250 difference in the commission paid to the broker).

[79] Article 1229 of the French Civil Code (n 39).

[80] Cass com, 9 May 1990, Pourvoi No 88-19.293 <http://www.legifrance.gouv.fr/initRechJuriJudi.do> accessed 10 March 2012.

[81] Article 1153 of the Spanish Civil Code (n 40).

[82] STS, 3 November 1999 (RJ, No 8859).

states to consider the adoption of the UNCITRAL Uniform Rules (1983),[83] summarized with brilliance the reasons for the harmonization of the conflicting common law and civil law rules governing penalties in the sphere of international commercial contracts:

Recognizing that a wide range of international trade contracts contain clauses obligating a party that fails to perform an obligation under contract to pay an agreed sum to the other party,

Noting that the effect and validity of such clauses are often uncertain owing to disparities in the treatment of such clauses in various legal systems,

Believing that these uncertainties constitute an obstacle to the flow of international trade,

Being of the opinion that it would be desirable for the legal rules applicable to such clauses to be harmonized so as to reduce or eliminate the uncertainties concerning such clauses and remove these uncertainties as a barrier to the flow of international trade,[84]

1. *Indetermination or Failure of the International Instruments of Coordination: Treaties and Soft Law*

Besides the UNCITRAL Uniform Rules, many other serious attempts have been made to broaden the enforceability of penalties in international trade, but nowadays there are no transnational rules that secure the enforcement of penalties in international commercial contracts. The lack of transnational rules in this area of law results from both the profound divergence between the civil and the common law traditions, and the relevant differences within the civil law countries.

The Benelux Convention on Penalty Clauses (1973)[85] was the earliest and perhaps the most courageous of these attempts, despite being addressed solely to three signatory states (Belgium, Netherlands and Luxemburg), with very similar national laws, and all members of the same regional trade organization.

Afterwards, the question was deliberately skipped in the Vienna

[83] See n 41.
[84] General Asseembly, Resolution 38/135 (1983) 270, UN Doc A/RES/38/135.
[85] Of course, the legality of contract penalties was not a controversial issue. This Convention deals with other questions such as the statute of limitations (Article 7). See n 41.

Convention (1980),[86] the most successful treaty offering uniform commercial law rules.[87] In my view, the CISG represented a lost chance to establish a path for the harmonization of contract penalties, given that its sphere of application is well-tailored (Article 1, 'contracts of sale of goods between parties whose places are in different States'), and parties to a contract may exclude or vary its application (Article 6).

Outside the domain of treaties, a wide variety of instruments have tackled this issue, however, none of them is legally binding for states, albeit potentially useful because parties may designate one of them as applicable law.

In the international arena, the UNCITRAL Uniform Rules (1983) were optimistically accompanied with a draft convention, mirroring the Vienna Convention,[88] even though the UNCITRAL Uniform Rules were never adopted.[89] The UNCITRAL Uniform Rules aimed to find a worldwide standard to balance the civil law enforceability, unless manifestly excessive, and the common law rule of unenforceability. The UNCITRAL Uniform Rules refer to 'contract clauses for an agreed sum due upon failure of performance' and non-sophisticated parties are excluded from its scope (Article 1),[90] providing that these clauses are presumptively valid, so the judicial intervention may consist only in the reduction of the agreed sum if 'substantially disproportionate' with respect to the actual harm (Article 8).[91] Nevertheless, the civil approach turned out to be predominant,[92] as

[86] United Nations Convention on Contracts for the International Sale of Goods (1980) 1489 UNTS 3 [hereinafter CISG]. Farnsworth (n 1) 812, n 5: 'Because of the wide gulf between common law systems and other legal systems, the Vienna Convention contains no provision on the important subject of stipulated damages'.

[87] Bruno Zeller, *CISG and the Unification of International Trade Law* (Routledge 2007) 94, in spite of relevant absences like Brazil, India, and United Kingdom.

[88] Draft United Nations Convention on Contract Clauses for an Agreed Sum Due upon Failure of Performance (1983) UN Doc A/CN.9/243, Annex II.

[89] Jonathan S Solórzano, 'An Uncertain Penalty: A Look at the International Community's Inability to Harmonize the Law of Liquidated Damages and Penalty Clauses' (2009) 15 Law & Bus Rev Am 779, 813: 'What is clear, however, is that somehow the proposal died. Model law or convention was ever adopted or entered into . . . The question we are left is *why*?'.

[90] Article 1 of UNCITRAL Uniform Rules: 'These Rules apply to international contracts in which the parties have agreed that, upon a failure of performance by one party (the obligor), the other party (the obligee) is entitled to an agreed sum from the obligor, *whether as a penalty or as compensation*' (emphasis added).

[91] Article 8 of UNCITRAL Uniform Rules: 'The agreed sum shall not be reduced by a court or arbitral tribunal unless the agreed sum is substantially disproportionate in relation to the loss that has been suffered by the obligee'.

[92] Against, Larry A DiMatteo, 'Enforcement of Penalty Clauses: A Civil-Common

evidenced by the non-trivial dropping of the 'genuine pre-estimate' between the revised draft (Article G) and the definitive version (Article 8).[93] For common law countries, the public policy concern against inequitable bargains together with the application by courts of two standards of justice, one for domestic and another for international transactions, or just the lack of interest may explain the failure of the UNCITRAL Uniform Rules.[94]

In the international arena too, the UNIDROIT Principles (Article 7.4.13),[95] the major instrument of soft law in the field of international commercial contracts, have also resolved the question following the civil law principle of enforcement of penalties subject to reduction:[96] after giving an broad definition intended to include both liquidated damages and penalties,[97]

Law Comparison' (2010) 5 Internationales Handelsrecht 193, 199, for whom the UNCITRAL Uniform Rules had a 'middle ground approach', arguing that '[b]y using the word "disproportionate" the Rules adopt the disproportionate standard found in American law and provides a wider scope to the voiding or reforming of penalty clauses in the civil law', statement which ignores the relatively higher familiarity with the term "disproportionate" or equivalent ones in civil law.

[93] Solórzano (n 89) 811-12. Article G of UNCITRAL Revised Draft: '(1) The agreed sum shall not be reduced by a court or arbitral tribunal. (2) However, the agreed sum may be reduced if it is shown to be grossly disproportionate in relation to the loss that has been suffered by the obligee, and if the agreed sum cannot reasonably be regarded as a genuine pre-estimate by the parties of the loss likely to be suffered by the obligee'. Nonetheless, already in the UNCITRAL Revised Draft, the prevailing view was that this element was not required for reduction, see UNCITRAL Revised Draft (n 41) 13, n 29.

[94] Solórzano (n 89) 804 and 813-14. The general lack of interest is a highly plausible explanation, especially regarding the common law countries, since only eighteen countries responded when the UNCITRAL Draft was circulated, and only one of them was a true common law country (Canada).

[95] International Institute for the Unification of Private Law, UNIDROIT Principles of International Commercial Contracts (1994) [hereinafter UNIDROIT Principles]. The Article dealing with contract penalties (7.4.13) has the same content in the 2004 version of the UNIDROIT Principles. Article 7.4.13 UNIDROIT Principles: '(1) Where the contract provides that a party who does not perform is to pay a specified sum to the aggrieved party for such non- performance, the aggrieved party is entitled to that sum irrespective of its actual harm. (2) However, notwithstanding any agreement to the contrary the specified sum may be reduced to a reasonable amount where it is grossly excessive in relation to the harm resulting from the non-performance and to the other circumstances'.

[96] DiMatteo (n 92) 199.

[97] Ewan McKendrick, 'Article 74.13' in Stefan Vogenauer and Jan Kleinheisterkamp (eds), *Commentary on the Unidroit Principles of International Commercial Contracts (PICC)* (OUP 2009) 919, 923. See also Michael Joachim Bonell (ed), *The Unidroit Principles in Practice. Caselaw and Bibliography on the Unidroit Principles of International Commercial Contracts* (2nd edn, Transnational

'agreed payment for non-performance', the general rule is the recoverability of stipulated damages regardless of the actual harm (Article 7.4.13(1)), but the court may reduce those 'grossly excessive amounts' (Article 7.4.13(2)).

Within the European context, the scholar-made soft law rules of both the Principles of European Contract Law (Article 9:509),[98] and the Draft Common Frame of Reference (Article III-3:712)[99] stuck to the pattern set by the UNIDROIT Principles: stipulated damages are named again 'agreed payment for non-performance' in the PECL, or 'stipulated payment for non-performance' in the DCRF, and in the both texts the governing norm is the recoverability of the sum irrespective of the actual harm, unless the court finds it to be 'grossly excessive', case in which the sum will be reduced. The antecedent of them was the Council of Europe Resolution (78) 3,[100] a set of eight non-binding rules that the member states were recommended to adopt in order to harmonize the civil law regimes.

The Council of Europe Resolution (78) 3, considered as a whole, contains much more detailed and elaborated rules than the soft law instruments examined until now (UNIDROIT Principles, PECL, and DCFR). Not only for using an inclusive definition of penalty (Article 1),[101] and turning to the principle of enforcement of penalties subject to reduction (Article 7), but also for dealing with the prohibition of cumulative penalties (Article 2), and the compatibility of the penalty with claims for specific performance, statutory damages, and additional damages (Articles 3, 5 and 6). The impact on the current civil law codes was minimal, since most reforms of the national laws towards the aforementioned principle occurred years before,[102] as described in Section I.2. Nonetheless, the Council of Europe

---

Publishers 2006) 342.

[98] See n 44. Article 9:509: '(1) Where the contract provides that a party who fails to perform is to pay a specified sum to the aggrieved party for such non-performance, the aggrieved party shall be awarded that sum irrespective of its actual loss. (2) However, despite any agreement to the contrary the specified sum may be reduced to a reasonable amount where it is grossly excessive in relation to the loss resulting from the non-performance and the other circumstances'.

[99] See n 44. Article III-3:712: '(1) Where the terms regulating an obligation provide that a debtor who fails to perform the obligation is to pay a specified sum to the creditor for such non-performance, the creditor is entitled to that sum irrespective of the actual loss. (2) However, despite any provision to the contrary, the sum so specified in a contract or other juridical act may be reduced to a reasonable amount where it is grossly excessive in relation to the loss resulting from the non-performance and the other circumstances'.

[100] See n 41.

[101] Arana de la Fuente (n 41) 6.

[102] Against, DiMatteo (n 92) 199, defending the influence of the Resolution in the

Resolution (78) 3 might be viewed as the European civil law model of contract penalties, given that the main characteristics of European civil laws are captured: (1) the validity of contract penalties, which may have the effect of coercing a party to perform her obligation; (2) the judicial review of penalties on the grounds of equity as a discretionary faculty, based on a retrospective test considering the actual harm, or on the grounds of partial performance; and (3) the promisee's entitlement either to the penalty or to specific performance, with the exception of delay.

2. *Fighting Uncertainty: Contractual Arrangements for the Enforceability of Penalties and their Effectiveness*

The lack of transnational rules that control the enforceability of penalty clauses in international commercial contracts, as shown in Section II.1, puts at risk the will of the contracting parties. In addition to this lack of transnational rules, the absence of coordination instruments among the several jurisdictions at a national level[103] entails that parties are unable to secure the enforceability of contract penalties by resorting to the available contractual devices, such as choice of law, forum selection, and arbitration clauses.[104] These contractual arrangements might turn out to be ineffective for several reasons, in particular whether the enforcement of the penalty is sought in common law courts, either adjudicating the dispute or executing the judgment or the arbitration award, since the mandatory rules against penalties might never be displaced.[105]

Obviously, the effectiveness of these contractual arrangements is likely to be higher when parties have chosen a civil law, and the court involved in adjudication or execution is also a civil law one, since general policy considerations that may render the penalty void will not arise so long as *lex*

later legislation regarding the generalization of the 'manifestly excessive' standard and the preference for reformation or reduction of the stipulated damages.

[103] Besides transnational rules, coordination instruments might also be unilaterally provided by purely national rules, for instance, by granting the application of the foreign penalty law designated by the parties, or by granting the execution of a foreign judgment or arbitral award.

[104] Pure drafting techniques intended to increase the chances of enforceability of penalty clause if a common law regime is applicable are not considered here, because these techniques are not capable to provide a minimum level of certainty under the case-by-case approach and the selective enforcement of stipulated damages. See n 7. See also DiMatteo (n 91) 200-01, making useful suggestions for drafting a penalty clause under American state laws.

[105] Farnsworth (n 1) 812, n 5, fearing that soft law may not derogate from this common law prohibition, albeit designated as applicable law by the parties: 'Whether this provision [Article 7.4.13 UNIDROIT Principles] can have any effect on a mandatory rule such as the common law rule prohibiting penalties is an open question'.

contractus and *lex fori* belong to the same legal tradition. For instance, if parties designate Spanish law as applicable, and the selected forum is Chile.[106] Within the European Union, the effectiveness of these clauses is even higher, due to the general rules of international private law in the area of contracts,[107] which even allow that parties derogate from certain mandatory rules. As an illustration, if parties decide to severely limit contract liability by using a penalty clause with an unreasonable small agreed sum, Italian law may be the applicable law designated to trump the French Civil Code (Article 1152) when the selected forum is France in order to prevent the judge from increasing a ridiculously low penalty (*'dérisoire peine'*) on the grounds of equity. The Rome I Regulation (Articles 3.3 and 9.1)[108] grants this possibility — only the 'overriding mandatory provisions' of the law of forum will resist the law chosen by the parties. In this regard, this French rule is mandatory,[109] but it can hardly be deemed an 'overriding mandatory provision' in accordance with the law of the European Union.

Conversely, the effectiveness of these contractual arrangements is uncertain when the parties intended to avoid the common law prohibition

[106] In the example, the parties of an international commercial contract intended the literal enforcement of the agreed penalty, avoiding the objective limit for pecuniary obligations imposed by the Chilean Civil Code (Article 1544), i.e. the penalty may not exceed the double of the value of the undertaking not performed.

[107] In the realm of contract law, Council Regulation (EC) 44/2001 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters (Brussels I) [2001] OJ L12; European Parliament and Council Regulation 593/2008 on the law applicable to contractual obligations (Rome I) [2008] OJ L177.

[108] Article 3.3 of Rome I Regulation: 'A contract shall be governed by the law chosen by the parties. The choice shall be made expressly or clearly demonstrated by the terms of the contract or the circumstances of the case. By their choice the parties can select the law applicable to the whole or to part only of the contract'. Article 9.1 of Rome I Regulation: 'Overriding mandatory provisions are provisions the respect for which is regarded as crucial by a country for safeguarding its public interests, such as its political, social or economic organisation, to such an extent that they are applicable to any situation falling within their scope, irrespective of the law otherwise applicable to the contract under this Regulation'. See Ana Quiñones Escámez, 'Ley Aplicable a los Contratos Internacionales en la Propuesta de Reglamento "Roma I" de 15.12.2005' (2006) 3 InDret 16-7 <http://www.indret.com/pdf/367_es.pdf> accessed 15 March 2012, explaining the origin and evolution of the concept of overriding mandatory provisions (*leyes de policía*) in the case law of the Court of Justice of the European Communities.

[109] Article 1152 of the French Civil Code (n 19), providing that '[a]ny stipulation to the contrary shall be deemed unwritten'.

of penalties, and the court deciding the case or executing the foreign judgment or arbitral award is a common law court. The likelihood of success increases in the next order: (1) only a pro-penalty choice of law, (2) a pro-penalty choice of law in conjunction with the selection of a civil law forum, and (3) a pro-penalty choice of law and arbitration in a civil law country, which is the safest way to secure the enforcement of penalties in international commercial contracts.[110]

However, considering the third solution to secure the enforcement of penalties in common law jurisdictions, one may doubt whether the enforcing court would refuse the recognition and enforcement of the arbitral award, since even the New York Arbitration Convention[111] grants this refusal if the recognition or enforcement would be contrary to the public policy of that country (Article V(2)(b)).[112] This ground under the New York Arbitration Convention casts doubt on the enforcement of an arbitral award in common law countries, in particular, in the United States, one of the jurisdictions analyzed here. Absent any decision from an American court, there is not a definitive answer to this question yet. Nevertheless, in accordance with the case law from another common law jurisdiction, the award would be enforced.[113]

An additional precaution that might be taken, as DiMatteo cleverly suggests,[114] is the prepayment of the penalty by means of an escrow account within the jurisdiction of the selected civil law forum, or within the same civil law country agreed for the arbitration. Notwithstanding, the contract may provide several penalties or penalties of a considerable amount, then none of the potential breaching parties will be prone to

---

[110] DiMatteo (n 92) 200, sharing the same view in this regard, despite a more pessimistic opinion about the execution of the award in the United States.

[111] Convention on the Recognition and Enforcement of Foreign Arbitral Awards (1958) 21 UNST 2518 [hereinafter New York Arbitration Convention].

[112] Article V(2)(b) of the New York Arbitration Convention: 'Recognition and enforcement of an arbitral award may also be refused if the competent authority in the country where recognition and enforcement is sought finds that . . . (b) The recognition or enforcement of the award would be contrary to the public policy of that country'.

[113] Dirk Otto and Omaia Elwan, 'Article V(2)' in Herbert Kronke et alii (eds), *Recognition and Enforcement of Foreign Arbitral Awards. A Global Commentary on the New York Convention* (Kluwer Law Interantional 2010) 345, 401 n 268, referring to a very recent Hong Kong court decision ruling that the Danish arbitration award providing for overcompensatory liquidated damages does not violate public policy, *A v R* [2009] HKCFI 342 (Court of First Instance of the High Court, Hong Kong). Against, DiMatteo (n 92) 200, sustaining that the award is likely to be questioned by American courts.

[114] ibid 202.

deposit the full amount of the penalties stipulated in the contract. Therefore, albeit the payment of the potentially due penalty is not completely secured, the deposit in the escrow account secures at least the partial payment, acting as well as a powerful incentive to ensure performance.

3. *A Quick and Safe Solution: To Shield the Enforcement of Penalties in International Commercial Contracts in Common Law Jurisdictions*

After having examined prior attempts for the harmonization of the conflicting common law and civil law rules governing penalties, the main reason of the failure of all these harmonization projects (treaties or bodies of soft law) has always been that the root principles of each legal tradition are not compatible, therefore the adoption of one root principle necessarily supersedes the other. In this regard, treaties and bodies of soft law have usually opted for the principle of enforcement of penalties subject to reduction, the civil law principle, a choice that has involved the understandable rejection of common law countries.

Under this dilemma, the demand of legal certainty in the field of enforcement of penalty clauses by the actors in international trade points to a relatively easy response: shielding the enforcement of penalties in international commercial contracts in common law jurisdictions by means of their statutory recognition at national level. Statutory recognition at national level that should be narrowly tailored to penalties expressly agreed by the parties in contracts in which at least one party is non-national, and the choice of law designates a foreign law according to which penalties are permissible.

In my opinion, the proposed solution is the most feasible for the enforceability and effectiveness of penalties in international commercial contracts because (i) it would be unilaterally adopted by each single common law jurisdiction, which implies that it has no coordination costs and that its success does not depend on an agreement among a high number of states; and (ii) it would be legally binding, which of course means a stronger effect than an optional regime designed in a body of soft law.

Nevertheless, there exists the well-founded fear of the rejection of the proposed solution by the legislatives of the common law countries, since it would lead to the application of two standards of justice, one for domestic and another of international transactions. This reasoning was one of the

grounds to turn down the UNCITRAL Uniform Rules.[115]

IV. Conclusion

After having explored the clash between the civil and the common law traditions, and the existing disparities among civil laws in the field of penalty clauses, this paper urges the adoption of transnational rules to secure the enforcement of penalty clauses in international commercial contracts in order to provide the actors in international trade with the certainty that they demand.

The international community acknowledged this need three decades ago, when the first UNCITRAL Draft was submitted in 1981, but the final UNCITRAL Uniform Rules and other harmonization projects have failed in that respect. Basically, the reasons that might explain this failure are two: on the one hand, all these projects have always aligned with the civil law legal tradition — in particular, the UNCITRAL Uniform Rules —; and, on the other hand, common law countries are unwilling to give up the prohibition of penalties, and tend to be prejudiced against the enforcement of penalties, even when the parties to a contract are merchants.

In the absence of coordination instruments among the several jurisdictions, the will of the contracting parties is at risk. Nevertheless, this lack of transnational rules is much more detrimental for the parties if a common law jurisdiction is involved. Not only civil law jurisdictions usually do not present sharp differences, but also the effectiveness of the contractual arrangements (choice of law, forum selection and arbitration clauses) is generally higher, especially within the European Union. On the contrary, whether a common law jurisdiction is involved, parties can only fight uncertainty by incurring substantial transaction costs to secure the enforcement of contract penalties, since all the available contractual devices have to be employed to diminish the likelihood of unenforcement. In this second scenario, the contractual toolkit may turn out to be insufficient, and therefore the need of transnational rules to bridge the gaps between civil and common law systems becomes critical.

Nevertheless, from a practical point of view, given the failure of all the attempts of the international community in the field of the enforcement of penalties, a quick and safe solution is the shielding of the enforcement of penalties in international commercial contracts in common law jurisdictions by means of their statutory recognition at national level. This

[115] Solórzano (n 89) 804 and 813-4.

recognition in each state would be restricted to penalties expressly agreed by the parties in contracts in which at least one party is non-national, and the choice of law designates a foreign law according to which penalties are permissible. Paradoxically, despite the need of transnational rules in this realm, the most feasible solution consists of the approval of national rules of limited scope but amazingly positive effects in international trade.

EUROPEAN UNION LAW AS INTERNATIONAL LAW

Dr Timothy Moorhead*

International law principles enable a rationalisation of the values to which the Union order aspires as a collective political and legal commitment amongst the Member States. The doctrine of Union law supremacy, which parallels that of international law supremacy, emphasises the overriding character of Union legal demands as a set of values and objectives over those of purely domestic origin. A common view that the Union legal order is sui generis or municipal in character fails to explain the directing character of the values underlying the Union project including its legal order.

In this article I therefore explore and defend the view that the Union legal order is essentially one of international law. A central contention in this regard is that the supremacy of Union law obligations within the Member States is based on the principle of the supremacy of international law obligations over those originating in the domestic arena. The intensive rationalisation of this principle by the Court of Justice within its case law manages the intrusive domestic legal effects of the values and ideals found in the Union Treaties and illustrates the evolutionary character of the Union project.

TABLE OF CONTENTS

* University of Kent

I. INTRODUCTION

The potential conflicts and collisions of systems that can in principle occur as between Community and Member States do not occur in a legal vacuum, but in a space to which international law is also relevant.[1]

I will explore two strands of inquiry in this article. The first concerns an interrogation of the proposition that the activities of the Court of Justice form an expression of the doctrine of the supremacy of international over domestic legal rules. This doctrine has traditionally been subject to a number of qualifications peculiar to individual States, relating for example to the ability of a later domestic legislative provision to set aside an earlier conflicting rule of international law.[2] In the Union setting however, qualifications associated with the doctrine have gradually diminished within Member States. These developments have been paralleled by an increase in the reach and influence of the supremacy doctrine where Union as opposed to 'traditional' international law demands are concerned. The 'strength in depth' of the [Union supremacy] doctrine is evident in relation to the variety of legal effects that have emerged within the Court's judgments in relation to Union laws, the broadening of conditions relating to the 'justiciability' of Union legal rules and finally concerning the location of jurisdictional authority to decide upon the scope of the doctrine which arguably now lies de jure and de facto with the Court of Justice.

A second strand of inquiry concerns a counter-intuitive feature of the defended international law character of Union legal ordering. Notwithstanding the Court's apparently complex and expansive development of the supremacy doctrine giving effect to the legal demands of the Treaties, the Court of Justice has in fact adopted a restrictive approach to their articulation. The reason for this is that the values or objectives of the Union are open-ended. They are therefore subject to a medium or long-term process of realisation that takes place within a shifting and often charged political environment. As a result, the Court has consistently qualified and 'managed' the legal potential of the objectives set out in the Treaties so as to ensure the continued viability of Union legal demands within the Member States and hence that of the

[1] MacCormick N, *Questioning Sovereignty* (Oxford University Press, Oxford 1999) at 120.

[2] The lex posteriori derogat lex priori principle, whose application (and rejection by the Court of Justice) in the context of a Union law demand was considered in Case 106/77 *Italian Finance Administration v Simmenthal* [1978] ECR 629.

entire Union project.[3]

The article will be structured as follows. I will consider first arguments for and against the international law character of the Union order, affirming the former position. Next, I will attempt to show that Union law, while rooted in the principles and practices associated with international law, represents a significant evolution of these principles. This is evident both in relation to the removal of qualifications previously applicable to international law effects in the domestic setting as well as in the location of the jurisdictional authority to determine these effects. Finally, I will consider how and why, notwithstanding these developments, the Court of Justice – and hence domestic courts when applying Union legal demands – have in fact adopted a markedly restrained approach towards their articulation given the profound political implications of a 'fully integrated' Europe mandated under the Treaties as a matter of international law.

II. THE UNION: INTERNATIONAL, SUI GENERIS OR MUNICIPAL TYPE OF LEGAL ORDER?

The prevailing view concerning the nature of the European Union legal order is that it is sui generis in character given the institutional characteristics it shares with both international and municipal orders.[4] According to this view, it is said that the extensive regulatory scope, legislative and adjudicatory independence and rule of law character of the Union order all indicate its unique character which represents a departure from ordinary principles of international law. It is generally agreed that the *origins* of the European Union (originally the European Communities) are to be found in public international law.[5] However, the international law basis of the Union as a dynamic, evolutionary body of 'governmental' institutional practices has been called into question by both the Court of Justice and academic commentators. In its *Van Gend en Loos* judgment the Court refers to the Union as a *new* type of international legal order.[6]

[3] For example by limiting the circumstances in which direct effects may arise.

[4] See for example Weiler J H H and Haltern, U R, 'The Autonomy of the Community Legal Order – Through the Looking Glass' (1996) 37 Harv.Int'l L.J. 411, hereafter Weiler and Haltern, 'Through the Looking Glass'.

[5] '[t]he origins, powers and objectives of the three Communities are all to be found in international treaties.' McMahon J F, 'The Court of the European Communities: Judicial Interpretation and International Organisation' (1961) 37 Brit.YB Int'l L 320 at 329. Weiler and Haltern have noted that, '[t]here is no doubt that the European legal order started its life as an international organisation in the traditional sense, even if it had some unique features from its inception.' Weiler and Haltern, 'Through the Looking Glass' at 419.

[6] The Court stated in its *Van Gend en Loos* judgment that 'The Community constitutes a new legal order of international law for the benefit of which the

Similar positions have been taken by Member State courts.[7] Academics have also sought to distinguish Union from international law, focussing upon the alleged constitutionalisation of Union legal demands. As Weiler notes,

> '[m]ost commentators focus on the legal doctrines of supremacy of European law, the direct effect of European law, implied powers and pre-emption, and on the evolution of the protection of fundamental human rights as hallmarks of this "constitutionalisation".'[8]

Weiler himself questions whether the distinction between international and constitutional legal ordering is a relevant one, suggesting instead that:

[a]ssuming the distinction between an international and a constitutional order makes any sense at all . . . we would prefer to focus on the following features that distinguish the European legal order from public international law: the different hermeneutics of the European order, its system of compliance, which renders European law in effect a transnational form of "higher law" supported by enforceable judicial review, as well as the removal of traditional forms of state responsibility from the system.[9]

A further consideration militating against the international law character of the Union order is its 'governmental' institutional framework and associated conferred powers. In this regard, the Treaties provide for a supranational institutional framework empowered to both legislate and provide legally binding adjudication[10] in relation to the various economic and social objectives they contain. The institutional practices of the Union moreover directly impact on individuals' rights and obligations within the domestic arena. Accordingly, Mark Jones notes that,

> . . . there are two relevant, fundamental distinctions between the obj ectives of Community Law and those of more traditional internation al law. First, the legal position of individuals is modified not just by the Treaties themselves, but also by the exercise of *governmental pow*

states have limited their sovereign rights, albeit within limited fields, and the subjects of which comprise not only Member States but also their nationals.'

[7] For domestic cases that have affirmed the *sui generis* character of the Union legal order, see Weiler and Haltern, 'Through the Looking Glass' at 421, fn 44.

[8] Ibid at 420 citing G F Mancini, 'The Making of a Constitution for Europe' (1989) 26 CMLR 595 and E Stein, 'Lawyers, Judges and the Making of a Transnational Constitution' (1981) 73 AJIL 1.

[9] Ibid at 420-421.

[10] Arts 267 and 258 TFEU.

> *ers* conferred upon the Community institutions by the Treaties. Seco nd, the Treaties and the powers they confer are concerned with mod ifying the legal position of individuals over an extremely wide range of economic and social activities.[11]

Weiler also highlights the Union's institutional structure as being characterised not by ' . . . general principles of public international law, but by a specified interstate governmental structure defined by a constitutional charter (the Treaties) and constitutional principles.'[12] Do then these features of the Union legal order noted by Weiler and Jones undermine the claim that the jurisdiction conferred under the Treaties can be understood in terms of international law principles? In short, does the directly applicable and supreme character of Union laws, its governmental institutional framework, wide ranging jurisdiction, and claimed rule of law basis underline its *sui generis* or even municipal law character? I will contend that these features do not n fact compromise the Union orders' essentially international law basis.

III. THE UNION AS INTERNATIONAL LEGAL ORDER

The claim defended is that the Union order is properly understood as a species of international law notwithstanding its developed institutional structure and extensive jurisdictional scope. The intrusive nature of the Union legal order within Member State jurisdictions does not require any departure from established international law principles. The supremacy of Member States' legal obligations arising as a matter of international law, can fully explain the domestic legal effects of Union laws and hence the relationship between Union and domestic legal orders. To make out these claims requires that I counter the arguments raised above – concerning the supremacy and direct applicability of Union laws on the one hand and the sophisticated governmental institutional framework underlying Union governance on the other – and offer a credible affirmation of its international law quality. Dealing with each of these points in turn.

First, the *sui generis* character of the Union legal order is attributed to the supremacy and direct effect doctrines which relate in turn to the overriding character of Union over domestic legal demands and the ability of individuals to rely upon or invoke Union legal demands before domestic

---

[11] M L Jones, 'The Legal Nature Of The European Community: A Jurisprudential Analysis Using H L A Hart's Model of Law and A Legal System' (1984) 17 Cornell Int'l L.J. 1 at 28 (emphasis in original). Hereafter Jones, 'The Legal Nature of the European Community'.

[12] J H H Weiler,'The Transformation of Europe' (1991) 100 Yale L.J. 2403 at 2407. Hereafter Weiler, 'The Transformation of Europe'.

courts. Do these features however actually represent a departure from established international law principles? For Spiermann, 'There is however no doubt that under international law, a national court, being an organ of the State, is obliged to reach decisions that are in accordance with the international obligations of the State . . .'[13] furthermore, ' . . . in modern international law, interests in the subject matter governed by a rule normally breed rights (to lay claims and bring actions) on the basis of the rule, also for individuals.[14]

The fact that domestic courts are bound in effect to uphold Union laws and that these laws are capable of modifying the legal rights or obligations of individuals before domestic courts does not of itself represent a principled departure from international law principles. It may be replied to this that while the 'defining' legal doctrines, of Union law supremacy and direct effect, are entirely familiar to international law, their expansive character within the Union setting does in fact represent a significant advance in jurisdictional authority associated with international law regimes. This is correct; however, the expansive character of the Union jurisdiction does not however of itself compromise the Union's international law pedigree unless we can find a principled justification of the distinction between municipal and international jurisdictions based upon the extent of jurisdictional authority alone. Given that there is no inherent limitation on either the scope or subject matter of international law agreements, such a justification for maintaining a principled distinction between the Union order and international law is likely to fail.

Does then the Union's governmental institutional character suggest the *sui generis* character of its legal order? In terms of the ability of domestic institutions to control or manage Union institutional demands, these demands are far-reaching by comparison with those arising under international law treaties generally. In addition, it is correct to say that the scope and claimed rule of law basis of these demands does represent a serious challenge to the jurisdictional claims of domestic orders. Finally, the ability to authoritatively interpret the substantive meanings of the Treaties being vested in a supranational judicial institution[15], the Court of Justice, is arguably a departure from the historically accepted prerogative

[13] O Spiermann, 'The Other Side of the Story: An Unpopular Essay on the Making of the European Community Legal Order' (1999) 10 EJIL 763 at 770. Hereafter Spiermann, 'The Other Side of the Story'.
[14] Ibid.
[15] Although recourse to international tribunals is now more common in relation to individual Treaty agreements. See for example the remit of the WTO tribunals to provide authoritative rulings on the GATT agreements.

of domestic executive or judicial branches of signatory states to do so.[16] These features once again do not in my view undermine the Union's international law pedigree, for the following reasons.

Any claim as to the 'distinctiveness' of the Union order as a result of its 'governmental' institutional structure including legislative, executive[17] and adjudicatory branches does not mean that its international law character is somehow altered. This would require acceptance of the proposition that, at a certain degree of institutional complexity, an international legal order loses its character as such. This proposition however confuses the core characteristics of international legal order with those of legal orders generally. The distinguishing features of international legal orders are a focus on the achievement of specific objectives or a commitment to the realisation of more broadly drawn social welfare outcomes within a limited jurisdictional sphere or a combination of both possibilities, by a number of cooperating States by way of Treaty agreements.

Legal orders as a generic category are evidenced by the existence of an institutional governmental framework operating in legislative, executive and judicial capacities and governed by values associated with the rule of law. To the extent that an institutional order possesses these features, it will be regarded as a 'developed' legal order. The fact that an international legal order which by definition pursues specified and jurisdictionally contained objectives, does so by means of an institutional framework which shares features with developed, rule of law legal orders does not mean that it has departed from its international law basis. The Union order involves institutional practices collectively directed to the achievement of the objectives contained in the Union Treaties. These are the establishment of the common market and the gradual harmonisation of the social and economic policies of the Member States under the overall

[16] So for example, in relation to practices of US courts, Morgenstern has noted that, ' . . . a treaty, as part of the law of the land can be interpreted by the courts, but great weight will be given to the view of the executive.' F Morgenstern, 'Judicial Practice and the Supremacy of International Law' (1950) 27 Brit.YB Int'l L. 42 at 79. Hereafter Morgenstern, 'Judicial Practice and the Supremacy of International Law'. The view that within the domestic setting, it is the executive as opposed to the judicial branch that may provide authoritative rulings on a Treaty's meaning has however been called into question by the European Court of Human Rights judgment in *Chevrol v France*, 2003-III Eur. Ct. H.R. 159 where it was held that the determination of rights arising under public international law must, in order to comply with fundamental procedural guarantees contained in Article 6 of the European Convention on Human Rights, be made by the judicial as opposed to executive branches of States.

[17] The executive activities of the Union are shared by the Commission, which is the dominant institution in this regard, and the Council of Ministers.

rubric of closer European integration. The fact that these objectives are supported by complex institutional structures governed by rule of law principles certainly gives an impression of a municipal-type order. However this can be seen as evidence of the extent to which the Member States have been willing to obligate themselves *as a matter of international law* to the realisation of these objectives.

Next, it is fully in accordance with international law principles that a supranational institution in the Court of Justice and not domestic constitutional courts should possess authority to adjudicate on the substance, status and scope of Union law obligations. Member States may not agree with the results where this extends Union legal demands beyond what they see as the competences conferred by the Treaties. However, there is no doubt that as a matter of international law, this is precisely the institutional role conferred on the Court by the Treaties. As Weiler notes, ' . . . the European Court, in adopting its position on judicial *Kompetenz Kompetenz*, was not following any constitutional foundation but rather an orthodox international law rationale."[18] Moreover, domestic courts have on the whole recognised the final authority of the Court of Justice to rule on the legal demands arising under the Treaties, thereby accepting the supreme character of the body of Union obligations and rights in accordance with the principle of international law supremacy.[19]

Finally, the fact that the Union possesses legislative competence in relation to the matters set out in the Treaties does not support an argument that the Union is closer to a municipal or *sui generis* order than one of international law *simpliciter*. The fact that the Union may enact directly applicable laws that (as such) take automatic effect within the Member States undoubtedly represents a significant advance upon the ability of international law norms to take domestic legal effects. In this regard, the Treaties transfer legislative authority to the institutions created under the Treaties themselves in order to achieve the aims they contain. We may accordingly assume that the Member States are empowered as a matter of international law to confer this authority. It is counter-intuitive to suppose that the resulting obligations arising as a matter of Union

[18] Weiler and Haltern, 'Through the Looking Glass' at 415.

[19] In Hartian terms, the recognition of the various international law features of the Union order – Union law supremacy and direct effects as well as the authority of the Court of Justice as the final arbiter of the scope, meaning and legal effects of Union legal norms – has emerged as a Rule of Recognition within Member State legal orders. See Jones, 'The Legal Nature of the European Community'. Interpretations of the effects of Union laws by the Court of Justice will apply as binding legal authority across all the Member States, see the Advocate General's opinion in Cases C-10/97 *Ministero delle Finanze v IN.CO.GE.'90 Srl* [1998] ECR-I6307.

legislation, and equivalent in status to the Treaty articles as far as domestic orders are concerned, are not themselves norms of international law.[20] The legislative competences of the Union rather evidence a supranational competence to create legal obligations that possess the character of international law. That is, they are supreme over all aspects of domestic legal ordering and operate in the service of the ideals or objectives found in the Treaties.

To sum up, the allegedly municipal features of the Union legal order identified by Weiler and Jones *inter alia* do not represent a principled departure from the international law character of the Union. Any principled distinction between international and municipal law does not rely on the presence or otherwise of developed institutional structures. Nor does it depend on whether institutions created under international agreements possess sovereign powers transferred from domestic jurisdictions beyond some (un)defined level. Nor finally does the distinction depend on the scope or status of legal effects promulgated and adjudicated upon by supranational institutions. The key elements of an international legal order are that all legal demands arising thereunder prevail over all domestic regulation and directed to the achievement of defined objectives set out in Treaty agreements among a collectivity of States, including the possibility of the independent institutional promotion of these objectives through the exercise of conferred powers. At this point, the question remains however as to whether the expansive character of the Union's legal jurisdiction adds something to our understanding of the operation of international law and specifically of the role of supranational courts charged with determining the legal effects of a highly intrusive international law jurisdiction.

IV. THE EXPANSIVE CHARACTER OF THE UNION'S INTERNATIONAL LAW JURISDICTION

[20] In this respect, Weiler's contention in relation to the direct effects of Union laws that these ' . . . reversed the normal presumption of public international law [that] . . . if the state fails to bestow the rights [conferred by Treaty], the individual cannot invoke the international obligation, unless internal constitutional or statutory law, to which public international law is indifferent, provides for such a remedy' is contradictory. If internal law were entirely independent of international law then it could not possibly allow the invocation of international law obligations before domestic courts unless these are said to operate as two wholly independent and distinct jurisdictions. This is plainly not the case. Indeed it is the very incorporation of Union law obligations within domestic orders that underlies the viability of the Union's legal jurisdiction. Weiler J H H, 'The Transformation of Europe' (1991) 100 *The Yale Law Journal* 2403 at 2413-2414.

The Court of Justice possesses competence under the Treaties and hence as a matter of international law to ensure that ' . . . in the interpretation and application of the Treaties the law is observed'[21] The Court therefore determines the status and scope of the legal obligations that flow from the application of Union laws and hence the legal effects deriving from the doctrine of Union law supremacy. The Court's institutional role conferred under the Treaties represents a significant transfer of jurisdictional authority, previously associated with domestic courts, to determine the legal effects of international law norms. This jurisdictional transfer has been 'reflected back' to domestic courts, with the Court counter-intuitively *increasing* the authority of domestic courts by enabling them to act as de facto 'Union courts' thereby exercising a constitutional power of review over all norms of domestic origin in light of all Union law requirements.[22]

In seeking to promote the substantive aims and objectives found in the Treaties, and in according these a supreme status over domestic law, the Court self-evidently acts as an international tribunal. In this regard, a feature of the Treaties that has allowed the development of a nuanced and hence viable portrayal of the supremacy doctrine by the Court of Justice is that the Treaties contain a combination of aspirational objectives relating to European integration and precise or 'hard' legal rights and obligations designed to put these aspirations into effect. This has allowed the affirmation of a regulatory framework within the Member States that factually supports the Union's underlying integrationist values while at the same time permitting these values to possess a direct legal influence as interpretive authority over all areas of legal regulation that possess a connection with the Treaty objectives. This brings us to closer consideration of the legal effects taken by Union legal rules.

1. *Justicialblity Factors I: The Legal Effects of Union Norms*

The legal effects taken by Union norms within the Member States depend on a number of factors, some of which, for example in relation to issues of justiciability also apply in the determination of the legal effects of laws of domestic origin. These factors may be summarised as follows. The first is whether a legal standard concerns subject matter that is suitable for judicial determination, a test of institutional suitability. This includes the questions of the legislator's intent (the Member States where the Union Treaties are concerned), the subject matter under consideration (which subjects are properly the subject of judicial as opposed to executive or

[21] Art 19 TEU.
[22] R v Secretary of State for Transport ex p Factortame [1989] 2 CMLR 353.

legislative determination), as well as broader issues relating to the role of the judiciary in a rule of law system of governance.[23] Second, whether a measure provides sufficient linguistic certainty (in identifying legal rights or obligations), the test of linguistic suitability. Third, whether further implementing measures are needed in order to put a measure into effect, the test of (un)conditionality.[24] In each of these areas, a single, definitive test is impossible to identify. What may be observed is that justiciability requirements may vary according to the type of legal effects taken (by Union laws). As a result, the Court's portrayal of the justiciability of Union norms and their consequent legal effects has been central to its development of the supremacy doctrine.[25]

In this regard, the Court has maximised the impact of Union law over domestic orders, by developing a variety of possible legal effects attributable to binding Union laws. These include direct effect in the

[23] Bickel A, *The Least Dangerous Branch: The Supreme Court at the Bar of Politics* (The Bobbs-Merrill Company. Inc. Indianapolis, New York 1962).

[24] These factors are equally applicable in the domestic setting regarding the suitability of domestic norms for judicial application. Carlos Vazquez has noted therefore that '[t]hese questions are not unique to treaties. The lack of "judicially discoverable and manageable standards" is often cited as bearing on whether statutory or constitutional provisions are judicially enforceable.' C Vazquez, 'The Four Doctrines of Self-Executing Treaties' (1995) 89 AJIL 695 at 714 (footnote omitted). Hereafter Vazquez, 'The Four Doctrines of Self-Executing Treaties'. In this regard, Vazquez, in considering factors that US courts have looked to in deciding whether Treaty provisions are 'self-executing' and hence give rise to enforceable individual rights, has noted that ' . . . courts have examined under the "self-execution" rubric various concepts that are not unique to treaties. These include matters such as whether the claim is justiciable, whether the litigant has standing, and whether the litigant has a right of action'. Ibid at 711.

[25] In his analysis of the self-executing / non self-executing distinction set out in US court judgments relating to the internal applicability of Treaty norms, Carlos Vazquez notes that, ' . . . the self-execution "doctrine" addresses at least four distinct types of reasons why a treaty might be judicially unenforceable. First, a treaty might be judicially unenforceable because the parties . . . made it judicially unenforceable. This is primarily a matter of intent. Second, a treaty might be unenforceable because the obligation it imposes is of a type that, under our system of separated powers, cannot be enforced directly by the courts. This branch of the doctrine calls for a judgment concerning the allocation of treaty-enforcement power a between the courts and the legislature. Third, a treaty might be judicially unenforceable because the treaty makers lack the constitutional power to accomplish by treaty what they purported to accomplish. This branch of the doctrine calls for a judgement about the allocation of legislative power between the treaty makers and the lawmakers. Finally, a treaty provision might be judicially unenforceable because it does not establish a private right of action and there is no other legal basis for the remedy being sought by the party relying on the treaty'. Ibid at 722-723.

'narrow'[26] sense of freestanding, individually enforceable rights, legality review effects and finally indirect or interpretive effects. Cumulatively these legal possibilities have been allied to an evolutionary and complex expression of the supremacy doctrine by the Court and hence the fullest possible expression of the underlying values of the Treaties by Union and domestic courts. In relation to 'direct effects'[27], the most intrusive of Union law effects, the Court stated in Becker that Union provisions must be 'sufficiently precise and unconditional' to give rise to enforceable individual rights.[28] For the legality review effects of Union norms to arise, these conditions will generally not need to be met[29] providing that an 'identifiable conflict' between Union and domestic provision(s) is present.[30] Finally, in relation to the interpretive effects of Union laws, we find that criteria of linguistic certainty do not play any part in determining whether these effects arise.[31] This enables the broad concerns of European integration, set out in the opening Treaty articles tol create legal effects.

2. *Justiciability Factors II: Linguistic Certainty, Political Questions and the Intent of the Member States*

Within the context of the determination of the legal effects of Treaty norms by domestic courts generally, the linguistic clarity of a provision of international law will often be linked to the question of whether the substantive topic raised is one suitable for determination by courts as opposed to other institutional branches. For example, Vazquez has suggested that in relation to the treatment of Treaty rules by US courts, 'precatory', 'hortatory' or aspirational provisions will not be self-executing as they evidence a commitment to achieve certain objectives in the

[26] A broad understanding of direct effectiveness refers to the invocability of Union norms in a domestic context so as to create legal effects that fall short of the direct conferral of rights. On the difference between broad ad narrow definitions of direct effect, see Craig and de Burca, *EU Law* at 178 ff.

[27] In the 'narrow' sense of the recognition of freestanding Union rights, enforceable by individuals before domestic courts.

[28] Case 8/81 *Becker v Finanzamt Munster-Innenstadt* [1982] ECR 53 at para 25 of the judgment.

[29] See inter alia, Case C-129/94 *Ruiz Bernaldez* [1996] ECR I-1829. For contrary judicial dicta, see the Court's judgment in *CIA*, which suggested that in order for the legality review (exclusionary) effects of a directive to arise in a horizontal situation, the conditions for direct effect must be met. Case C-194/94 *CIA Security International SA v Signalson SA* [1996] ECR I-220.

[30] Lenaerts and Corthauts, 'Of Birds and Hedges'.

[31] Case 6/72 *Europemballage and Continental Can v Commission* [1973] ECR 215, particularly paras 24 and 25 in which the Court held that the Treaty competition rules prohibiting abuse of a dominant position in the market had to be interpreted in light of the overall objectives of the Treaty contained in the opening articles which affirmed the promotion of the common market amongst the Member States.

political arena as opposed to an intention to confer legally enforceable rights:

> "Precatory" treaty provisions are deemed judicially unenforceable not because of the parties' (or anyone's) intent, but because what the parties agreed to do is considered in our system of separated powers, a "political" task not for the courts to perform.[32]

For the Court of Justice, the determination of whether a norm reveals a 'political' question that is as such not suitable for judicial enforcement, requires a broader judgment, one that addresses its role in relation to both the other Union institutions as well as the Member States. In addition, for the Court, so far as the legal effects of the aspirational provisions of the Union Treaties are concerned, the distinction between vaguely worded aspirations and those that are 'sufficiently precise', relates as noted above, to the type of legal effect produced as opposed to being determinative of the question of whether legal effects may arise. The proposition that all objectives of the Union project are capable of creating legal effects is supported by the first articles of the Treaties. These objectives are worded in mandatory terms. Article 3 (3) TEU for example states that:

The Union shall establish an internal market. It shall work for the sustainable development of Europe based on balanced economic growth and price stability, a highly competitive social market economy, aiming at full employment and social progress, and a high level of protection and improvement of the quality of the environment. It shall promote scientific and technological advance. It shall combat social exclusion and discrimination, and shall promote social justice and protection, equality between women and men; solidarity between generations and protection of the rights of the child. It shall promote economic, social and territorial cohesion, and solidarity among Member States. It shall respect its rich cultural and linguistic diversity and shall ensure that Europe's cultural heritage is safeguarded and enhanced.[33]

This wording evidences an intention on the part of the Member States, that the overall objectives of the Union – in addition to the precise regulatory demands arising under Treaties and associated secondary legislation, should be able to create legal effects notwithstanding that these objectives undeniably fall under the heading of 'political questions'. Moreover, from the perspective of the Court of Justice, the political objectives found in the Treaties are precisely those matters to which the

---

[32] Vazquez, 'The Four Doctrines of Self-Executing Treaties' at 712.
[33] Art 3 TEU (emphases added).

'hard' or precise legal provisions of the Treaties are directed to achieving. All legal regulation arising under the Treaties is directed in some way to the achievement of the social, economic and political ideals, of European integration. For the Court of Justice therefore, as well as domestic courts, the institutional considerations relevant to whether a Union law is justiciable require a reformulation of the political tasks doctrine noted by Vazquez above and based on received notions of the separations of powers between the courts and other domestic governmental branches.

Union law effects within the domestic setting entail a modified understanding of the separation of powers doctrine, one that seeks the constitutionalisation of the body of Union laws through the activities of domestic courts. Such practices redefine the domestic judicial role in relation to the other governmental branches and are in contrast to the manner in which domestic courts have determined the internal legal effects of international treaties outside the Union context. Here, domestic courts seized with questions of international law will focus exclusively on the separation of powers doctrine embodied within the constitutional settlement of the State concerned without considering how international law measures may themselves qualify or alter that settlement.[34] Finally, in relation to questions of intent, the Court of Justice does not consider the intentions of the Member States in assessing the legal effects of the Union provisions it is called on to interpret and apply. Instead it will assess the language of a provision in light of the overall purposes of the Treaties. In this regard, a 'general' interpretive assumption is that all individual Union measures are intended by the Member States collectively to fit within the overall scheme of the Union legal order, which is based on the achievement of the Treaty objectives.

In sum, the suggestion that: 'Where the line is drawn between "precatory" [and hence judicially unenforceable] and "obligatory" [and hence judicially enforceable] treaty provisions is a matter of domestic constitutional law'[35] is revised in the Union setting. This revision concerns on the one hand the

[34] For these reasons, Marc Amstutz, speaking of the Court's practices in relation to the interpretive effects taken by Directives notes that 't]he thoroughly courageous decision to intervene at the level of the rules governing legal reasoning (where the link between text setting and text interpretation is made) and – propter unitatem juris – extend the law-making powers of the national judiciaries beyond the contra legem boundary drawn by long-established legal theory (and thus into the legislative sphere defined mirror wise by the same legal tradition) is . . . a socially adequate (albeit also highly risky) alternative strategy.' Amstutz M, 'Marleasing and the Emergence of Interlegality in Legal Reasoning' (2005) 11 ELJ 766, *at 777-778,* Amstutz M, 'Marleasing and the Emergence of Interlegality in Legal Reasoning'.

[35] Vazquez, 'The Four Doctrines of Self-Executing Treaties' at 713.

ability of the Court to authoritatively determine the meaning of Union Treaty provisions and secondary Union laws, and on the other its active constitutionalisation of Union norms within the domestic arena.[36] This evolution of the doctrine of international law supremacy by the Court of Justice in upholding the supremacy of Union over domestic legal requirements does not represent a decisive break from the Union's international law basis but rather expresses illustrates an organic expression of principles of international law.

To sum up these points, we may say that all the Member States, whether monist or dualist in their approaches to international law have duly incorporated, according to the terms of their respective constitutions, the Union legal order as a directly applicable system of supreme legal rules that exists alongside laws of domestic origin. The fact of incorporation does not however *conclude* the legal effects of Union norms, a fact illustrated by the gradual evolution of the doctrine of Union law supremacy by the Court of Justice and its corresponding acceptance within the Member States. The evolving quality of Union law effects within the Member States highlights the potential inherent in the Union jurisdiction within Member States. In practice, this potential has been far from realised as a result of an attitude of restraint by the Court of Justice towards the legal demands arising under the Treaties. Reasons for this restraint are explored in the following section.

3. *The Restrictive Expression of Union law demands by the Court of Justice*

The ability of the Court of Justice to promote the Treaty aims depends on the extent to which its conclusions regarding the supremacy of Union over domestic law are accepted by domestic courts. In this regard, Union law supremacy represents a 'value' whose promotion and articulation by the Court of Justice and recognition by domestic courts is crucial to realisation of the Union project.[37] In common with the Court's challenge to the incumbent role of domestic courts under pre-existing separation-of-powers arrangements, the Court's presentation of the supremacy of Union law represents an invitation or challenge to domestic judicial practices regarding their treatment of international legal norms. This does not

[36] In this regard, Jones has noted that, ' . . . the Court consistently has resisted arguments by the national governments that, in accordance with the practice of international law, the question of penetration [of Union laws within the domestic setting] is to be determined by national constitutional law. The practices of Member States vary considerably and, therefore, any solution based on the provisions of traditional national constitutional law will not ensure the full and uniform application of Community Law in all the Member States'. Jones, 'The Legal Nature of the European Community' at 45 (footnotes omitted).

[37] Famously in its *Van Gend* and *Costa* judgments.

mean that the international law character of Union legal obligations is somehow qualified. However, the need to ensure uniform Union law effects while developing the Treaty aims has required the careful 'management' of the (supremacy) doctrine by the Court of Justice

Given the character of the legal obligations set out in the Treaties, backed by an independent and sophisticated institutional framework, a significant untapped potential exists in relation to the domestic effects of Union laws as a matter of international law. The legal obligations placed on the Member States under the Treaties are in fact *more* extensive than the portrayal of these obligations by the Court of Justice. The Court has in fact offered a moderate and limited expression of Union legal demands. In relation to the *pace* and *extent* of European integration, the Court has vouchsafed the continued development of the Union by remaining responsive to a political imperative that concerns the *acceptability* of its legal demands within the Member States.

The Court of Justice has been faced not only with ensuring that ' . . . in the interpretation and application of the Treaties the law is observed'[38] but also with the need to maintain the viability and indeed survival of the Union's intrusive international law jurisdiction. Recognising that the full import of the Treaties' legal demands would be unacceptable to the Member States and hence practically unenforceable, the Court has therefore developed the doctrine of Union law supremacy by reference to what it deems institutionally *possible*. While this has entailed an expansive engagement by domestic courts by comparison with other legal demands of international law origin, the Court's articulation of the Treaty demands has at the same time been consistently restrained. It has advanced the doctrine of Union law supremacy in incremental steps, consistently maintaining a limited view of the legal obligations found in the Treaties.

The Court of Justice's allegedly dynamic and extensive approach to Treaty interpretation then obscures the fact that the legal possibilities represented by the Treaties as a matter of international law are in fact diminished by political, legal and jurisdictional concerns that prevent their full expression. The Court's case judgments reveal a 'thin' interpretive approach to the principle of international law supremacy where the legal effects of the Treaties are concerned. The articulation of Union law demands by the Court therefore represents as noted above, a *managed*, challenge to the jurisdictional expectations of domestic courts. In addition, as pointed out by Spiermann, given that the Member States have signed up to a project of international cooperation in the Union Treaties

[38] Art 19 TFEU.

that has clear potential to challenge incumbent notions of State sovereignty, the Court has highlighted the relevance of *State* sovereignty in developing the Treaties' legal effects, an approach he traces to the Court's judgments in *Costa* and *Wilhelm*[39]:

> . . . as in *Costa v ENEL*, in *Wilhelm*, state sovereignty was treated as a key ingredient of treaty interpretation, essentially because the Court by then had recognised such a strong position of national law in regulating market structures that the Treaty was binding only within the context of national law, thus making precedence an appropriate synonym of [the international law principle of] *pacta sunt servanda.*[40]

The recognition of domestic legal regulation and hence the *potential* for conflict between Union and domestic law by an international tribunal is, as pointed out by Spiermann, unusual from an international law perspective given that, '[f]rom the point of view of international law, there can be no conflict between a treaty rule and a national law rule, for the rules do not belong to the same system.'[41] The Court has affirmed, in Spiermann's view, a 'national lawyers' perspective on the relationship between international (Union) and domestic law, a perspective that prioritises what he terms the international law of co-existence. [42] This approach emphasises (the role of) domestic sovereignty in defining a States' international law obligations without recognising the possibility of an *unconditional* joint limitation of sovereignty, a possibility which is an established feature of what Spiermann terms the international law of co-operation.[43] Under the former approach, the emergence of the Union doctrines of supremacy and direct effect indeed evidence a 'new' kind of legal order but only by

---

[39] Case 14/68 *Walt Wilhelm and Others v Bundeskartellamt* [1969] ECR-I. For consideration of the judgment in this case, see Bengoetxea, *The Legal Reasoning of the European Court of Justice* at 265-266.

[40] Spiermann O, 'The Other Side of the Story: An Unpopular Essay on the Making of the European Community Legal Order' (1999) 10 *European Journal of International Law* 763, hereafter, Spiermann, 'The Other Side of the Story' at 785.

[41] Spiermann, 'The Other Side of the Story' at 773 citing Article 27 of the Vienna Convention on the Law of Treaties according to which: '[m]unicipal law may not be invoked as a justification for failure to perform international obligations'. Vienna Convention on the Law of Treaties (Adopted 23rd May 1969, entered into force 27th January 1980) 1155 UNTS 331.

[42] '[t]he ahistorical idea of international law embraced by the Court of Justice in *Van Gend en Loos* and *Costa* was the international law of coexistence.' Spiermann, 'The Other Side of the Story' at 779.

[43] ' . . . as an international lawyer will know, when compared to other parts of the international law of cooperation, there is nothing new about direct effect and nothing innovative about precedence.' Ibid at 787.

reference to what Spiermann terms a 'narrow' understanding of international law.[44] This however neglects possibilities found within the international law of cooperation which represents a more credible understanding of international law principles and one which can account fully for the features of the Union legal order explored above.

The international law obligations found in the Treaties then have been presented by the Court, not as a limitation on and corresponding transfer of domestic sovereignty that automatically takes precedence over domestic law, but instead as effects that occur strictly within the context of (competing) domestic legal demands. It is this engagement with domestic legal concerns by the Court of Justice that is remarkable, in Spiermann's view, from a (true) international law perspective. The Treaties have never been used as a platform to challenge failures by the Member States to meet the obligations arising as a matter of international law to *positively* promote the aims that they contain.[45] The Court's approach has rather secured the co-operation of domestic courts in expressing these demands. This in turn has promoted the acceptance of Union requirements by domestic institutional actors generally and served to manage conflicts between Union and domestic governmental activities as well as 'internal' constitutional conflicts that may arise from an 'over empowerment' of domestic courts.[46]

---

[44] Thus, Spiermann notes in relation to the doctrine of Union law direct effects that, [b]y neglecting the international law of co-operation, the Court ended up with a narrow idea of international law, which explains how the Court could assume that international law is unfamiliar with the idea of direct effect and the involvement of the individual. Ibid at 779.

[45] In this regard Spiermann notes that the Court has tended to focus on discriminatory practices as constitutive of the economic freedoms, as opposed to barriers to trade that may not be discriminatory (although in recent years that Court has increasingly allowed challenges to substantive impediments to trade that are not directly discriminatory, maintaining however the requirement that they must operate in an indirectly discriminatory manner towards goods / workers / services from *other* States so maintaining a necessary cross border element). He notes in this regard that '[t]he content of a ban on nationality discrimination is purely negative, saying that the state is not allowed to treat aliens in any way worse than its own national. The EEC Treaty thereby opened a door. But in order to generate a real opportunity for aliens to go through that door, it was arguably necessary, or at least conducive to the objective of making a common market to supplement the negative ban on discrimination with various positive principles; or to put it more crudely, to offer the aliens a pat on the back when they appeared on the doorstep.' Ibid at 781-782.

[46] Either internally within Member States or in the relationship between the Union and the Member States. See on these issues, R Rawlings, 'Legal Politics: The United Kingdom and Ratification of the Treaty on European Union (Part One)' [1994] P.L. 214; A M Burley and W Mattli, 'Europe before the Court: A Political Theory of

V. Conclusions

How do these conclusions assist in understanding the role of the Court of Justice as a court of international jurisdiction charged with promoting the values of the Union? I have argued that two essentially competing current have informed the Court's articulation of Union law demands. The first concerns the legal commitments present as a matter of international law within the Union Treaty agreements. The Union Treaties represent a binding commitment under international law to profoundly restructure of Member States' governance around the ideals of an integrated European Union. The second concerns the fact that the full realisation of the objectives contained n the Treaties cannot realistically be achieved through the immediate assertion of legal demands alone given that even a limited expression of the overall Treaty objectives entails an acceptance of supranational legal authority previously unknown within the Member States.

The Court of Justice has therefore evolved the principle of Union law supremacy and hence the values of European integration through a consistently creative expression of its jurisdiction. Domestic courts have been persuaded to cede a 'corresponding' jurisdiction – to decide the limits and qualification of the supremacy of Union law – to the Court. This has paradoxically provided domestic courts with a derivative Union jurisdiction that embraces the Court's reading of the supremacy doctrine, empowering domestic judicial actors to a degree that arguably redefines the separation of powers amongst domestic governmental branches in favour of the judiciary.[47] In sum, the Court's exercise of an international law jurisdiction has triggered a constitution restructuring of domestic judicial authority.

Understanding the Union order as one of international law explains the *ideal* possibilities found in the Treaties as guiding factors in relation to Union institutional practices. The Treaties provide a perennial invitation to the Member States and Union institutions to pursue the ideal of an integrated Europe and to enhance and develop their commitment to a Kantian vision of European legal ordering that this commitment implies.[48]

Legal Integration' (1993) 47 I.O. 45; K J Alter, 'Who are the 'Masters of the Treaty'?: European Governments and the European Court of Justice' (1998) 52 I.O. 121; D Nicol, *EC Membership and the Judicialisation of British Politics* (Oxford University Press, Oxford 2001).

[47] Amstutz M, 'Marleasing and the Emergence of Interlegality in Legal Reasoning'.

[48] 'According to Kant, the states must finally enter into a cosmopolitan constitution

The almost unlimited potential of the Union project to effect, as a matter of the international law, an institutional restructuring of the Member States allied to the commitments signed up to in the Treaties has meant that the Court has necessarily developed the ideals of European integration in a qualified and limited fashion.

due to the constant wars and "form a state which is not a cosmopolitan commonwealth under a single ruler, but a lawful federation under a commonly accepted international right.' Přibáň, 'The Juridification of European Identity' at 52 citing I Kant, 'On the Common Saying: "This may be true in Theory but it does not apply in Practice"' in G H Reiss (ed), H.B.Nisbet (trans), *Kant's Political Writings* (Cambridge University Press, Cambridge 1971) at 90.

THE CONVENTION FOR THE PROTECTION OF ALL PERSONS FROM ENFORCED DISAPPEARANCE: MOVING HUMAN RIGHTS PROTECTION AHEAD

Ioanna Pervou*

The entry into force of the Convention for the Protection of all Persons from Enforced Disappearance at the end of 2010 signified the most important step in the struggle against enforced disappearances and marked a development in international human rights law. This article provides a historical overview of the phenomenon and tracks the background of the Convention's adoption. It analyses and evaluates the definition adopted by the Convention. It also probes into practices applied against terrorism and suggests that they should be classified as enforced disappearances under the Convention. Overall, it is argued that the Convention's application can be expected to cement detainees' protection.

TABLE OF CONTENTS

I. INTRODUCTION

Enforced or involuntary disappearances are a persisting phenomenon globally. The international community has been addressing it for more than forty years, not always successfully. All previous attempts have been stumbling at states' reluctance to share information and admit the exercise of disappearances, as well as to punish themselves (i.e. their organs and

* Doctor Juris Candidate Aristotle University of Thessaloniki (Academy of Athens Scholarship), LL.M. *Cantab*. The author is grateful to Dr. Alison Kesby for all her comments and structural proposals. The author is also indebted to Ass. Prof. Miltiadis Sarigiannidis for his support, advice and encouragement.

agents involved therein). Thereby there was not any international legal framework for years and only fragmented regional efforts had been recorded. Still, disappearances' complexity has turned into a retarding factor for the complete legal response to the phenomenon. Unlike torture and extraordinary executions, disappearances cannot be easily conceptualized and further captured in a definition. As a result, only the systematic study of the phenomenon's historical background can determine its specificity. States, or more precisely governments, have developed a number of practices which varied slightly, yet significantly, to erase the traces of those they considered opponents; opponents being determined mainly by their political beliefs or solely by their race. As a matter of fact, any legal response should cover all forms a disappearance could take and offer the maximum protection to any potential victim. The legal imprint of the phenomenon should correspond to all factual combinations and at the same time avoid a descriptive character. These are the pillars the 2007 United Nations Convention on Enforced Disappearances lies on. The international community has learnt its lesson well and took a holistic approach on the phenomenon of enforced disappearances. The definition provided takes into account the historical aspects of disappearances and aims to comprise all contemporary methods applied by states currently. In this sense, the Convention's definition serves a double goal: first, to demonstrate the distinct character of the phenomenon and protect all persons from the standardized methods reported so far and secondly to prevent the emergence of novel practices of disappearances. Regarding the first goal, there is little doubt for the Convention's success; however as disappearances remain in reality there are still challenges to be resolved. In this respect, the practices states apply to investigate terrorist acts incorporate elements of disappearances and raise the question whether they could fall within the Convention's protective scope and be classified as enforced disappearances. Anti-terrorist methods are a test for the Convention's applicability; uncharted waters which need to be further explored. All told, the Convention's value depends primarily on its applicability to current developments in public international law.

II. A HISTORICAL OVERVIEW

The term 'enforced disappearances' (desapararición forzada) was introduced by Latin American NGOs[1] to encapsulate a phenomenon that occurred in South America in the second half of the 20th century. For some authors this term is 'a euphemism'[2] for describing a series of severe human rights violations. Moreover, enforced disappearances are considered a recent addition to the human rights agenda,[3] and it has attracted global concern principally because of the large number of

victims.

Enforced disappearances were recorded for the first time during World War II, when thousands were disappeared due to the policy of the 'Night and Fog Decree' (known as the 'Keitel Order').[4] In that case, Adolf Hitler ordered the transfer of people who were deemed dangerous for the security of the Third Reich to the concentration camps in Germany. Vanishing without leaving a trace and providing information was thought to be an appropriate measure for the intimidation of the potential enemies of the Reich.

This Gestapo policy later spiraled in Latin America taking the form of a systematic, governmental practice; this practice aimed at the suppression of political opposition, since it was considered a threat to national security.[5] More specifically, during the 1960s and the 1970s, and especially within the political context of the Cold War, military juntas seized power in most Latin American countries. The majority of those military juntas were serving the establishment and preservation of a capitalist system based upon foreign investments. These dictatorships are usually referred to as 'bureaucratic – authoritarian' regimes. This term emphasizes the fact that the Latin American dictators did not aim to dissolve public institutions, but on the contrary to use them in favour of their regime. Consequently, at some level the term 'enforced disappearances' is also a synonym for the incessant use of military force to obliterate any form of opposition and to ensure public order. In this setting, enforced disappearances proved to be an effective measure for the sustainability of the military juntas.

Overall, the juntas sought to terrorize people in order to establish civil obedience. Therefore, the authorities turned against civilians regardless of their ideology making disappearances a part of everyday life in Latin America. Governments elaborated a very specific and detailed mode of operation. The victims were usually carried off from their homes in the presence of their families. Then they were transferred to secret detention centers where they were tortured to death. Torture was not an interrogation method but rather a means of dehumanizing the detainees, before death. The military applied unlimited torture, although they did not aim to extract information or obtain confessions from the victims. They viewed torture as part of their mission to cleanse society politically; opponents had to be punished before executed. The authorities refused to inform the victims' relatives of their fate and denied their detention. Most of the cases ended in extrajudicial executions and very few victims survived and reappeared.[6] This situation raised the concern of the Inter-American Commission on Human Rights (IACommHR). After making

reference to Argentina's 'disappeared' in 1978 it began to report on disappearances in Guatemala and Chile, encountering unsurprisingly the regimes' unwillingness to cooperate.[7]

Nevertheless enforced disappearances soon spread beyond Latin America. By way of example, during the 1970s many individuals' status was also unknown in Cyprus, as a result of the 1974 Turkish military intervention. However, these victims are not referred to as 'disappeared persons' but as 'missing' or 'persons unaccounted for', to demonstrate the difference between the causes of their disappearance.[8] This pattern suggests that enforced disappearances were mainly 'attributable to political reasons'.[9] The Philippines is another example where disappearances served as a tool against political opposition. During the Marcos Dictatorship (1971-1986) the country suffered from innumerous disappearances which were systematically applied from 1976 onwards. Hence, enforced disappearances are not a regional phenomenon or one rooted only in regimes perpetrating atrocities. It also occurs in countries with long-standing internal conflicts.[10] In other cases, enforced disappearances are used by governments to decimate indigenous populations[11] or they are associated to gender-based violence supported by the authorities.[12]

Consequently, the phenomenon has troubled the international community as it continues to proliferate and because the perpetrators usually remain unpunished. For over forty years, the international community struggles for a viable legal response to enforced disappearances. Its responses, though, have not always been to the benefit of the victims, or as influential as victims and their relatives would expect. Finally, the 2007 Convention seems to satisfy both states and civil society and to enhance all persons protection from disappearances.

III. INTERNATIONAL LEGAL RESPONSES TO ENFORCED DISAPPEARANCES

In the beginning, the international community treated the phenomenon of enforced disappearances *in casu* by appointing ad hoc Working Groups to monitor the application of human rights standards in Chile and Cyprus. Soon it became clear that a holistic approach was necessary and the General Assembly adopted resolution 33/173 on 'Disappeared Persons'[13] requesting the Human Rights Commission to 'consider the question of disappeared persons with a view to making appropriate recommendations'.[14] Thereinafter, the ECOSOC requested both the Commission on Human Rights and the Sub-Commission on Prevention of Discrimination and Protection of Minorities 'to consider the subject and

make recommendations to the Commission on Human Rights'.[15] The Sub-Commission recommended the establishment of a group of experts to collect 'all the information ... and to make the necessary contacts with the Governments and the families concerned'.[16] The United Nations Working Group on Enforced or Involuntary Disappearances (UNWGEID) was established with a resolution adopted without a vote by the Commission on Human Rights.[17]

The establishment of the UNWGEID was not an easy task. As soon as disappearances became part of the UN human rights agenda, tensions grew between states over the appropriate way to address the phenomenon. The drafting of a legally binding instrument was out of the question for almost all delegations, because the phenomenon was relatively new in the international plane and there was a lack of knowledge about the issue.[18]

However, the majority of states had realized that there should be an international response to enforced disappearances and so they suggested the establishment of a thematic mechanism, the UNWGEID. By contrast, states that applied the policy of enforced disappearances (like Argentina and Uruguay) opposed the creation of the mechanism; instead, they preferred the adoption of resolutions which would only acknowledge the existence of enforced disappearances. At that point it was the dedication of human rights NGOs to achieve a long-term solution that proved instrumental. Having secured political and diplomatic support by the American delegation,[19]they tried to rouse public concern on enforced disappearances and pressure governments to reach an agreement. To this end they organized campaigns on enforced disappearances and released particular details on the applied governmental practices.[20] These activities in conjunction with incessant lobbying paved the way for the establishment of the UNWGEID. Even states that were initially opposing to its creation, finally conceded to it, as it was the only way to avoid further criticism of their policies. Overall, certain states considered the UNWGEID an important step against enforced disappearances, whereas others saw it as the least problematic approach on the topic.[21] Still, even under these circumstances, the creation of the UNWGEID reveals a conscious and alarmed international community.

In this context, the UNWGEID's mandate depended much on international politics. Indeed, the Group had a narrow but clear mandate to deal with disappearances that involved a degree of governmental involvement and liability. At the same time, it decided not to address disappearances associated with armed conflicts.[22] It also 'decided to approach its tasks in a humanitarian spirit',[23] meaning that in the context of Cold War it should not get involved in or criticize the member-

states' domestic politics. It would instead seek governmental cooperation in order to function as a third party between the families of the disappeared and the liable governments. Thereafter the UNWGEID was heavily criticized for its dubious approach to the problem both by several human rights NGOs and authors. However, it was not the chosen approach that caused dissatisfaction, but the low rate of the cases that were resolved and the growing number of enforced disappearances worldwide.[24]

Apart from the UNWGEID's attempts to provide answers for the victims' fate, their relatives continued seeking the truth, either individually or through associations they had created. In a number of cases, their quest for justice led them to submit communications to the Human Rights Committee (HRC) under Article 5 of the Optional Protocol to the International Covenant on Civil and Political Rights (ICCPR).

The views of the HRC set the foundations for the protection of individuals. In the first communication considered (*Eduardo Bleier v Uruguay*)[25] it found breaches of Articles 6, 7 and 10 of the ICCPR and held state authorities responsible for the fate of the individual.[26] The importance of HRC's conclusions lies in the reversal of the burden of proof that it established in cases of disappeared people.[27] The HRC held constantly that state parties have by definition more access to the necessary information than the individuals and their insufficient responses turn in favour of the complainants. The HRC also stated that the parties' undertaking to provide the Committee with the requested information follows their positive obligation to conduct full investigations as to the fate of the disappeared.[28] Unfortunately, the HRC's lack of enforcement capacity has proved to be an obstacle difficult to surmount. The 'naming and shaming' strategy that has been developed by the HRC seems to be an inadequate tool to deter enforced disappearances, let alone that this process was still in its infancy at that time. Therefore, it signaled a considerable advance when the Inter-American system for the protection of human rights accepted the challenge to cope with enforced disappearances and bring the perpetrators to justice.

The Inter-American Court of Human Rights (IACHR) proved to be a leading authority in disappearances. On 23 July 1988 the IACHR held a 'landmark ruling'[29] concerning the *Velásquez Rodriguez Case*.[30] The case is of paramount importance not only regarding the domain of enforced disappearances, but also for the protection of human rights in the Inter-American system in general, since it was the first time that the IACHR applied its compulsory jurisdiction in a contested case.[31] Moreover, the case is pivotal because the Court established

special evidential standards regarding the practice of enforced disappearances and for the first time a state was held responsible for performing disappearances. More specifically, the Court lowered the required threshold of evidence as it acknowledged that one of disappearances' main aims is to efface all evidence[32] and held that 'circumstantial evidence, indicia, and presumptions may be considered, so long as they lead to conclusions consistent with the facts'.[33] In a unanimous ruling it found Honduras responsible for the disappearance of the victim. In doing so it declared the violation of several rights of the ACHR, as disappearances were not stipulated *per se* in the Convention. Furthermore, the Court held that the violation of these rights was in direct conjunction with the obligation of state parties of the ACHR to organize their legal orders in a way that it guarantees the protection of human rights (Article 1(1)). Consequently, it was established that enforced disappearances were violating the ACHR's values, *in toto*.[34]

The ruling created a leading precedent for the Inter-American legal order and confirmed the awareness of this regional community on enforced disappearances. The Court used exactly the same argumentation in the cases to follow and pinpointed that their common feature is the purpose of weakening political opponents[35] and intimidating the population.[36] It regarded disappearances as part of a general and systematic practice applied by governments. The Court, though, drew away from this line of argumentation in some of the ensuing cases of enforced disappearances it dealt with, with the view to strengthen the procedural aspects of the trials. In the case of *Caballero-Delgado and Santana v. Colombia* the Court affirmed that disappearances may arise on an occasional basis and not as part of a systematic practice.[37] In this case, uncontested evidence was presented on behalf of the victims regarding their disappearance and subsequent execution. The existence of strong evidence determined the Court's decision to a great extent. By contrast, in cases where evidence was insufficient to either indicate a governmental practice on disappearances or the victims' mistreatment and suffering, the Court did not pronounce a violation of the ACHR.[38]

The plethora of reported cases revealed the phenomenon's diffusion as well as the cruelty it entailed. The attempts of the HRC and the IACHR to deal with enforced disappearances stumbled at the lack of established international standards that would provide for a common understanding and legal basis. It gradually became evident that the phenomenon of enforced disappearances could not be captured by reference to already existing norms. The need for a more thorough and effective approach, based on a set of basic legal principles was evidently obvious.

The international community took prompt action leading to the adoption of the 1992 UN Declaration on the Protection of all Persons from Enforced Disappearances (hereinafter 1992 Declaration).[39] Some of the factors which finally led to the adoption of the Declaration include public awareness and the constant pressure from NGOs towards the drafting of Conventions on Enforced Disappearances, the adoption of the Convention Against Torture (CAT 1984) and the recommendations made by the UNWGEID.[40] Soon after the Declaration, the Organization of American States (OAS) adopted the Convention on Forced Disappearance of Persons.[41] It was the first legally binding document, which tackled the phenomenon directly, tailored to the idiosyncrasy of the regional plane it was designed for. Overall, this Convention enhanced the Inter-American system for the protection of human rights and promoted democratization in Latin America countries.

All the same, there were a few occasions where international courts displayed very 'limited understanding of the phenomenon'[42] as they applied formalistic criteria. This happened particularly when the European Court of Human Rights (ECHR) first dealt with enforced disappearances in 1998. As discussed below, the ECHR's jurisprudence proved inconsistent with the approach of the HRC and the IACHR so far. It may be said that different circumstances ask for a different approach, yet there were some instances when the Court's judgments were simply deficient. The vast majority of the cases tried by the ECHR were related to Turkey. They were associated with the internal disturbances in the south-eastern region of the country, which is mostly populated by Kurds.[43] The Court, though, did not acknowledge that disappearances were systematically practiced there and tried them on an *ad hoc* basis. Its approach diverged from the one already established by the IACHR.[44] Moreover, in contrast to the approach of the IACHR and the HRC, the ECHR decided to apply high evidentiary standards. Thus, it did not accept that a reversal in the burden of proof was necessary and did not lower the evidentiary threshold as it demanded proof 'beyond reasonable doubt'.[45] As a result, the applicants had to present information they could not access, which is in effect 'a sort of *probatio diabolica*'.[46] On top of that, the ECHR created a quantitative formula when it came to violations of the right to life.[47] It held that the victims could be presumed dead only when a considerable period of time had passed without any news from the disappeared[48] which left unanswered questions as to the status of people who disappeared not long ago. It was not until recently, that the ECHR adopted a more flexible approach and acknowledged the relation between enforced disappearance and the threat of death. In the case of*Baysayeva v. Russia* it held that disappearances are life-threatening, when the victims disappear under violent

circumstances.[49] However, there are some positive aspects in the ECHR's jurisprudence, especially when it comes to member-states' duties under the European Convention of Human Rights (ECHR). The Court held that states have a duty to effectively investigate every case of disappearance. This duty emanates from the general obligation established under Article 2 of the ECHR to 'protect the right to life by law'. Therefore, member-states have to conduct prompt and profound investigations on the fate of the disappeared as soon as they take notice of it; their failure to do so constitutes a breach to the ECHR. In other words, the Court did not easily pronounce a violation of the right to life resulting from a disappearance, however it required states investigate the alleged violation,[50] compensating for its hesitance to presume the victims' deaths.

Despite serious developments on the international level – such as the adoption of the 1992 Declaration and the progressive evolution of jurisprudence – there were still unsolved issues, which hindered a satisfactory response to the phenomenon. Such inconsistencies could not be easily surmounted without a universal instrument which would directly address the main issues of the phenomenon. In 1998 the Sub-Commission on the Promotion and Protection of Human Rights adopted a Draft International Convention on the Protection of all Persons from Forced Disappearance[51] (1998 Draft), which shed light on key aspects of the phenomenon. Moving forward and capitalising upon previous efforts, the Commission on Human Rights adopted without a vote resolution 2001/46, according to which an independent expert (Prof. Manfred Nowak, a former member of the UNWGEID) had to examine the existing international human rights' framework on enforced disappearances and report on the necessity of a 'legally binding normative instrument'.[52] Reaffirming the strong concern of the international community, the resolution established an 'Inter-sessional Open-ended Working Group (ISWG) to elaborate a draft legally binding instrument for the protection of all persons from enforced disappearance',[53] having taken into consideration the recommendations of the expert. Professor Nowak concluded that a legally binding instrument was essential for establishing protection against disappearances, since there existed gaps regarding, *inter alia*, the definition of the term, the perpetrators' punishment and the phenomenon's prevention. Thus, he proposed three possible forms:

> 'a separate human rights treaty such as the draft convention, an optional protocol to the International Covenant on Civil and Political Rights, or an optional protocol to the Convention against Torture.'[54]

It was then upon the ISWG to decide the form of the document. The Working Group decided that a separate treaty would be the most appropriate form[55] and in 2005 it submitted a draft to the Commission on Human Rights.[56]

Finally, the International Convention for the Protection of All Persons from Enforced Disappearance (Convention) was adopted by the newly created Human Rights Council and consecutively by the Third Committee and the General Assembly.[57] It entered into force on 23 December 2010 and it now counts 32 members and 91 signatories. As the European Union representative stated during the GA Plenary Session, the Convention 'sends a strong political signal from the international community that this shameful and still widespread practice must come to an end'.[58]

Indeed, the Convention fills serious gaps in the protection against disappearances. The creation of the right not to be subjected to enforced disappearance alongside the definition of disappearance are probably the most important achievements of the Convention. The creation of a comprehensive, protective legal framework requires the phenomenon's crystallization. Accordingly, the Convention's definition is the starting point for understanding all persons' right not to be subjected to enforced disappearance, but also for grasping the concepts and values of the Convention.

IV. ENFORCED DISAPPEARANCES: THE QUEST FOR A WIDELY ACCEPTED AND COMPREHENSIVE DEFINITION

The definition of enforced disappearances proved to be an issue of legal and political controversy.[59] During the last years, disappearances have been labeled by a diversity of characterizations, which were used almost indistinguishably.[60] As a result, once disappearances attracted international concern, shedding light on the content of the term proved to be a difficult task.

Human rights NGOs were the first to respond to the need for 'conceptual clarity'.[61] From the early 1980s, NGOs engaged in a strenuous quest for a definition; yet all definitions have been descriptive and followed an analytical approach of the practices developed by governments. The UN was also in search of a definition, but it did not come to any until the adoption of the 1992 Declaration. That definition came as the precursor of the one which states parties concluded in the 2006 United Nations'

Convention, however it was not a part of the Declaration itself, but only a preambular clause. The 1992 Declaration has proven instrumental, since it 'proclaims' itself 'as a body of principles for all states'[62] and is the first international document to declare that 'any act of enforced disappearance is an offence to human dignity', thus violating a series of human rights. Moreover, the act of enforced disappearance also constitutes an offence under criminal law. This conceptualization clearly illustrates the reasons why the 1992 Declaration was a landmark; it established a definition at the international level and stipulated that enforced disappearances should be treated as offences under the domestic legal orders. The indispensability of the definition is due to the fact that an offence (enforced disappearance) had been acknowledged. Thus a domestic legal order should prevent and punish such acts. Given this evident correlation between the definition and the offence and their parallel lives, it could be argued, that despite its preambular placement, the definition has been functionally incorporated into the main part of the Declaration. Moreover, the 1992 Declaration remains important even after the adoption of the Convention, since

> [i]t sets forth a set of rules that all the Member States of the United Nations, without the requirement of a ratification are called upon to apply as a minimum to prevent and suppress the practice,[63] and because it might be also considered of customary value.[64]

The definition given is principally based on the 'working definition' adopted by the UNWGEID in its annual reports.[65] The UNWGEID in its general comments on the Declaration stated the elements that a definition of enforced disappearance should include three minimum cumulative elements, which are:

> a) deprivation of liberty against the will of the person concerned,
> b) involvement of governmental officials, at least indirectly by acquiescence,
> c) refusal to disclose the fate and whereabouts of the person concerned.[66]

These elements are interconnected and reveal the historicity and complexity of the crime.[67] However, this offence presupposes that there exists governmental involvement or at least awareness of the perpetration of the crime, obviously rendering the above provision useless, if there is no political will to punish the crime and suppress this practice.[68]

The significance of the progress made in 1992 is beyond doubt. Apart from establishing 'the autonomous nature of the crime'[69] it gave impetus to a

broader debate upon the issue. Therefore the next step deemed essential was the creation of a right *per se*, not to be subjected to enforced disappearance. Article 1 of the Convention established a new right, using a negative formulation, and after intensive negotiations the state-parties finally concluded a definition (Article 2).

The term 'enforced disappearance' was the focal point during the negotiations as it would naturally entail specific state obligations. The definition given in Article 2 of the Convention follows the pattern of the 1992 Declaration, but the formulation of the term is substantially different, although phrasal alterations seem slight. Article 2 reads as follows:

> "[e]nforced disappearance" is considered to be the arrest, detention, abduction or any other form of deprivation of liberty by agents of the State or by persons or groups of persons acting with the authorization, support or acquiescence of the State, followed by a refusal to acknowledge the deprivation of liberty or by concealment of the fate or whereabouts of the disappeared person, which will place such a person outside the protection of the law.

The term is analyzed in three elements according to the dominant approach. It is explicitly stated that the deprivation of liberty is the first element of an enforced disappearance. The final phrasing is the outcome of negotiations, a compromise between the proposals made during the sessions of the ISWG. The initial wording, proposed in the Working Group by the Chairperson-Rapporteur referred to 'the deprivation of a person's liberty, in whatever form'.[70] The follow up debate brought to light two different trends. Some states endorsed the Chairperson's suggestion, with a view to ensuring full protection, whereas other states considered the phrasing 'imprecise',[71] thus calling for the use of more specified terms. Although, the use of specified terms such as arrest, detention and abduction would be 'by way of example',[72] meaning that the listing in the definition is not exhaustive, it seems that the delegations sought for clarity in the definition in order to limit ambiguity.

However, these terms do not only serve as examples in the context of the Convention. Their explicit enumeration, signals that they constitute essential components of a 'disappearance'.[73] They form part of the crime, an element of it (*actus reus*).[74] Finally, these terms are of considerable conceptual value, because they are invoked in the most important international human rights instruments and their content has been enriched through the interpretation of international courts and tribunals over the years.

Another contested issue during the negotiations was that of lawfulness. Some delegations expressed the opinion that only cases of unlawful deprivation of liberty should be included.[75] But the majority of states did not welcome this approach opining that it would dramatically limit the term's field of application. Besides, jurisprudence,[76] as well as the UNWGEID's experience had also shown that the deprivation of liberty was essentially related to the third element of a disappearance, namely refusal to acknowledge the deprivation and concealment of the fate or whereabouts of the victim, indicating that lawful arrests or detentions could turn into disappearances.[77]

The third element of the term[78] depicts the denial of the proper national authorities to cooperate with the relatives or the counsel of the victim and inform them about his/her fate. It is not surprising that national authorities might refuse the deprivation of liberty itself[79] or details about it and as a result erase all traces of the victim. This element is not only related to factual circumstances but it is also critical in achieving one of the two aims of the Convention, that of prevention and it should be read together with Articles 18 to 20. Article 18 refers to the right to information of the people with a legitimate interest and lists the accessible information, whereas Article 20 frames the exception, spelling out when the state can refuse the provision of information about the detained person. This article caused disagreement until the end of negotiations of the ISWG, as it was thought to distort the instrument and render it ineffective. The drafters of the Convention were aware of the fact that this provision could serve as a 'Trojan Horse', capable of bringing the Convention's implementation to a standstill; thus they explicitly restricted its scope of application through the establishment of both affirmative and negative requirements. In any event, though, Article 2 enjoys normative supremacy over Article 20 and in case of conflict it prevails (Article 20(1)).

However, ambiguity does not arise from the priority to be accorded the two provisions, but rather from which 'conduct' is deemed permissible. During the negotiations, some delegations opted for exceptions to the right to information on the grounds of witness protection, threats to national security, and the protection of the detainee's integrity, whereas others proposed for the postponement of the information provision instead of refusal.[80]Overall, it seems that articles 2, 18 and 20, on their proper interpretation protect the right to information. As a result, a systemic interpretation would ensure maximum protection for the victim and his/her relatives as well.

The last part to examine is the one related to the status of the perpetrators (the second element of the definition). According to Article 2, the

perpetrators should be 'agents of the State or persons or groups of persons acting with the authorization, support or acquiescence of the State'. Thus enforced disappearances are committed only by state actors, either direct or indirect. Non-state actors are excluded from the definition and Article 2 cannot be applied to them, not even by analogy, since there is special provision for them (Article 3). This approach caused dissatisfaction during the negotiations and proved to be a hard case, as members of the ISWG agreed only in the last section. It has been argued that the UNWGEID and the ISWG have adopted a 'traditional notion' on this topic, since they left out non-state actors.[81]

However, a careful reading of ILC's Articles on the Responsibility of States for Internationally Wrongful Acts (ARSIWA) reveals what non-state actors stand for in the law of state responsibility and which entities are finally excluded from the Convention's definition. As Special Rapporteur James Crawford points out, human rights supporters have long demanded the abandonment of a 'firm distinction between the State and the private sector' when the international law of state responsibility is applied to human rights instruments.[82] The request, though, for 'extension of state responsibility in the private sector' is considered to be 'undue'.[83] According to the ARSIWA in a number of cases, non-state actors' acts are attributed to the state, namely when they serve as agents of the state, when they function under the direction or the control of a state and lastly when armed opposition groups are guided by the state.[84] Even if the ARSIWA guarantee in this manner that there shall be no impunity for non-state actors by equating them to indirect state actors, there are still arguments to explain why this does not correspond specifically to the practice of disappearances. The most convincing amongst them, is that in disappearances it is almost impossible to prove who committed the crime and further on, whether there was state involvement or not.[85] The issue of non-state actors provoked serious discord during the sessions. The Chairperson managed to reach a compromise with the inclusion of Article 3. This article was cautiously phrased, as it refers to 'acts defined in article 2' and not to enforced disappearances, implying that the acts are characterized as such only when there is state involvement. The provision acknowledges states' discretion in this field.[86] In the meantime it lowers the victim's protection as these acts fall outside the *ratione materiae*.[87] The most serious concerns were expressed by the associations of the families of disappeared persons and by NGOs who argued that the provision's scope may be distorted by governments in an attempt to justify their policies.

Overall, the above debate can be condensed into two conflicting propositions. First, state involvement is a *sine qua non* condition of

enforced disappearance. Secondly, it is very difficult to prove state involvement in disappearances, especially when indirect. Thus, what seems appropriate (so as to avoid doctrinal aberrations and meet the phenomenon's particularity) is to adopt a wide interpretation of the ARSIWA and lower the applicable evidential threshold in cases where there are allegations of indirect state involvement. This view has already been introduced by the IACHR in the case *Masacre de Pueblo Bello v. Colombia*, where the Court held Colombia responsible for disappearances carried out by paramilitary groups.[88]

Still, the treatment of non-state actors committing enforced disappearances is not the most complex part of the definition. The last phrase of Article 2 'which place such a person outside the protection of the law' constitutes one of the major weaknesses of the whole text. More specifically, states disagreed on whether the placement of the victim outside the protection of the law was a fourth element of the definition (the subjective part of the crime, meaning that the intention of the perpetrators should be accordingly evidenced), or a mere consequence of any act of enforced disappearance. States having experienced enforced disappearances held that the placing of the victim outside the protection of the law was an 'inherent consequence' of an enforced disappearance.[89] On the other side, a number of states urged for an additional forth constitutive element to the definition. They explained that it would be incompatible with their domestic penal systems to introduce a crime which would not ask for the establishment of the perpetrator's intention. Apart from that, they also referred to the definition provided by the Rome Statute of the International Criminal Court (ICC Statute) where intent is a critical element.[90] The debates left the issue unresolved and thus the Chairperson of the ISWG stated that states-parties 'were fully entitled to make an interpretive declaration on the matter at the time of ratification.'[91]

However, the interpretation of this phrase is found in documents prior to the Convention, and it leaves no doubt about the meaning of the text. According to UNWGEID the placement of a person outside the protection of the law is an aftermath of a disappearance.[92] Nowak and the ISWG during its early sessions also side with this view, as they identify only three constitutive elements in a disappearance and they expressly avoid reference to it as the fourth one. Nowak also underlies that it would be almost impossible to identify intent in the perpetrators' acts, as in most cases, they are trained to carry out specific tasks, for which they could only be held responsible.[93] Yet, the most explicit and clear statement which matches this view, is the one made in a UN experts' Joint Report on secret detention:

> [T]he definition does not require intent to put the person concerned outside the protection of the law as a defining element, but rather refers to it as an objective consequence of the denial, refusal or concealment of the whereabouts and fate of the person.[94]

The fact that states did not embrace the approach set up by the UNWGEID during the last decades indicates that this issue touched upon the important matter of reserved jurisdiction and domestic policy. The issue gains even more importance when it comes to the evidentiary standard set for disappearances. If 'putting the victim outside the protection of law' is to be considered an element of the definition, then the alleged victims should prove that the perpetrators had *dolus* in doing so, which undoubtedly makes the evidentiary threshold higher.[95]

However, this debate was totally misleading. The purport of this phrase is properly revealed when examined in combination with the provisions of the OAS Convention. The relevant phrase in the OAS Convention is: 'thereby impeding his or her recourse to the applicable legal remedies and procedural guarantees'.[96] Under the Inter-American system, 'outside the protection of the law' means that the victim is denied recourse to legal remedies. This is a material element which concerns the victim's case and not the perpetrator's defence, an aspect which was overlooked during the negotiations.

The fact that the ISWG Chairperson referred to it as the 'third and half element' of the definition, trying to reconcile all different views, did not remove ambiguity over this point. This is regrettable because it assigns priority to the crime of enforced disappearance passing over the right not to be subjected to it. In other words it sets guarantees for the protection of the potential perpetrator, minimizing the protection of the victim.

Notwithstanding the above controversies, the Convention's definition enjoys wide acceptance both by states and by human rights' NGOs. It is also a positive development regarding the identity of perpetrators, even if non-state actors are excluded from the definition. Overall, the definition is deemed a success, not only because it is comprehensive but also because it offers quite a broad definition which in turn may well correspond to a wide variety of methods that governments apply. The fact that it recognizes that any kind of deprivation of liberty may result in a disappearance is very important especially with regards to new methods to which governments resort to.

In this respect, contemporary practices of the 'War on Terror' have unfolded new aspects of the issue. In the name of national security many states launched anti-terrorist campaigns and moved towards strict legislation, thus increasing the risk enforced disappearances to occur.

V. ENFORCED DISAPPEARANCES THROUGH THE ANTI-TERRORIST SPECTRUM

The 'War on Terror' raised new issues for the law of enforced disappearances, regarding mainly the application of two key practices: incommunicado detentions and extraordinary renditions.

The aftermath of the 9/11 terrorist attacks reinforced the public interest concerning disappearances. On the grounds of the 'Global War on Terrorism' against the so called 'Axis of Evil', some states took austere legislative measures authorizing human rights' restrictions as a safeguard to national security, whereas other states went further and promulgated a state of emergency.

It is not the first time that such a policy has been implemented by states; since 1960s the language used by the Latin American authoritarian regimes identified military or paramilitary groups as subversives or terrorists.[97]Yet now, the situation is different due to the fact that this policy is adopted by democratically elected governments and generally by countries which are often referred to as liberal democracies. In addition, the operations carried out after 9/11 against terrorism are unprecedented in terms of their intensity and state cooperation in intelligence sharing.

One of the effects of these draconian laws was the substantial increase in suspects' detentions which were mainly secret or incommunicado. States embarked on new techniques as well, which resulted in the lowering of the applicable human rights standards. Thus, enforced disappearances came to the fore once more, as a result of these circumstances.

1. *Incommunicado Detentions*

It has been already mentioned that a deprivation of liberty is just one of the three constitutive elements of an enforced disappearance. Also, according to the definition in the 2007 Convention every kind of deprivation of liberty might turn into a disappearance. Indeed, some methods place the detainee under an incredibly high risk and result almost always in a disappearance. 'Incommunicado detention' is, in these terms, a means of erasing all traces of the victim. The term describes the detainee's absolute confinement from the outside world. The victim is not allowed to

communicate with people other than his/her captors.

The implications of incommunicado detention are several and relate mostly to the victim's protection. The victim is unable to notify his family of this new situation and the reasons for his custody and also cannot consult a lawyer. His confinement indicates a further denial by the victim's captors to bring him/her before the judiciary.[98] Because of these restrictions, the detainee's treatment is in the captors' absolute discretion and may 'invite other forms of coercion'.[99]

Incommunicado detention is not a novel practice and has already been addressed by the international community.[100] The question though, is whether an incommunicado detention may amount to, or result in an enforced disappearance, given the human rights' curtailments that states have already introduced and are willing to undertake under their anti-terrorist campaigns. The UNWGEID has stressed the potential relationship between the two since 2003.[101] According to UNWGEID's Reports and the Convention, incommunicado detention falls under the states' obligation to take preventive measures against disappearances and to refrain from using any methods that endanger a detainee's security. The Convention does not mention incommunicado detention *expressis verbis*, but it can be argued that it implies it in article 17(2)(d) (read in conjunction with article 17(1) which refers to secret detention):

> [...][e]ach State Party shall, in its legislation: Guarantee that any person deprived of liberty shall be authorized to communicate with and be visited by his or her family, counsel or any other person of his or her choice, subject only to the conditions established by law, or, if he or she is a foreigner, to communicate with his or her consular authorities, in accordance with applicable international law.

The Convention's rationale denotes that the drafters understood incommunicado detention as a particular form that secret detention may attain (incommunicado detention is an aspect of secret detention: article 17(2) seeks to address the violations a person suffers when secretly detained). Contrary to this view, some commentators place incommunicado just before disappearances in the scale of severity. In other words, an enforced disappearance is considered 'a heinous form of incommunicado detention'.[102] Apparently this view misconceives the complexity of an enforced disappearance; however, the value of equating these two practices is obvious only exceptionally when it comes to the newly developed anti-terrorist policies.

In the post 9/11 era, incommunicado detentions are standard tools to confront terrorism and to avert future attacks. The scope of the undertaken measures is not the individual's extermination (as it was in the 1960s), but the weakening of the terrorist organization's structures. Thus, the captors aim at the extraction of the best available information. To that end, a detainee's confinement enables the authorities to apply severe interrogation techniques affecting his/her treatment,[103] but not concluding in torture that will cause irreparable damage or in extrajudicial executions. This, of course, does not guarantee humane treatment for detainees. Recent statistics prove that ill-treatment is almost inevitable during incommunicado detention[104] and detainees are exposed both to physical and mental suffering.

Furthermore, confinement enables the authorities to leave the suspects incommunicado for a prolonged period (this tactic apart from the detainees' debilitation also ensures that there is no communication with other suspected terrorists), which constitutes per se cruel inhuman and degrading treatment[105] and is a *prima facie* violation of the ICCPR[106] and of the IACHR.[107] The data available shows that terrorist suspects are usually held incommunicado for months or even years,[108] while the HRC held as early as in 1979 that even 6 weeks of incommunicado detention is a breach of the Covenant.[109] In addition, the ECHR found a fourteen days incommunicado detention to be exceptionally long, even when there is a state of public emergency because of a terrorist threat.[110] Prisoners that are under prolonged incommunicado detention are usually referred to as 'Ghost Detainees'. This term describes eloquently the detainees' absolute alienation from the society and that their very existence depends solely on the information they possess.

It seems that prolonged confinement and severe interrogation methods applied to alleged terrorists fulfill all the requirements of Article 2 of the Convention and set up the causal link between incommunicado detentions and enforced disappearances. Obviously, the current conditions of terrorism suspects' detentions leave no doubt that these individuals are eventually disappeared under human rights standards for so long as they are in confinement. The key element is that the state also refuses to acknowledge their detention and whereabouts. The fact that at some point they may be put on trial or get released does not affect their characterization as disappeared, since neither the 2007 Convention nor international jurisprudence ask for the victim's death or interminable capture.[111] This, also, does not reduce their next of kin's anguish over their fate. Their relatives cannot be aware of the patterns that intelligence services follow and therefore they fear for the detainees' life.

The extensive application of incommunicado detentions in the 'War on Terror' might also mark a turn in enforced disappearances' jurisprudence. So far, the international human rights courts have in the majority of cases presumed the victim's death because there were allegations of ill-treatment. However, contemporary enforced disappearances occur under different conditions. As victims are likely to reappear, courts should be more cautious in presuming their death; this clearly indicates that disappearances could be dissociated from the right to life. Although, it could be argued that a shift in international jurisprudence comes dimly into sight,[112] it is rather premature to deem a general change since jurisprudence is only now evolving on the issue.

2. *Extraordinary Renditions*

The anti-terrorism measures have reasonably incited domestic criticism in the states that adopted them. Human rights NGOs and the mass media stressed the legal contraventions they entailed and further enumerated their inconsistencies both with domestic laws and international obligations. As a result, public opinion started opposing some of the adopted rules, despite the fact that terrorism remains on top of the agenda regarding national security and is still considered as a potential danger. Therefore, governments faced constant pressure to disclose information about the detention conditions of terrorist suspects while their refusal to do so exacerbated domestic reactions. Some states in an attempt to evade accusations for human rights violations (at least regarding domestic legal standards) turned to other methods. That was the critical point when extraordinary renditions became a commonly applied tool in the War on Terror, also affecting enforced disappearances.

> 'Extraordinary rendition' is neither a legal term nor an entirely new one.[113] In regard to enforced disappearances, it is well suggested that extraordinary renditions have been used by governments since the 1970s. At that time, intelligence services of several Latin American countries had created a network of information-sharing for alleged 'subversives/terrorists'. This networking is also known as Operation Condor (Operación Condor).[114]However, this is a rather primitive form of the methods that states have developed after 9/11.

Indeed, extraordinary rendition is now used to describe the transfer of alleged terrorists from the country where they are apprehended to states with underdeveloped and poor human rights protection. In other words, it is a forcible transboundary movement,[115] a complex method which

requires the cooperation of at least three countries: the captor, the accomplice and the extractor state. The suspect is usually caught in the borders or in airports of a country (the accomplice state) by secret agents of another country (the captor state). The victim is then taken to a third country, where he is held in custody and interrogated (the extractor state). In most cases, the interrogation takes place in secret detention centers, over which the captor state's secret services exercise a significant degree of control. Extraordinary renditions have not been standardized up till now as there is not any standard pattern followed.[116] Despite several variations that have been recorded so far, there is a common feature in all such incidents: the element of extraterritoriality vis-à-vis the captor state. The suspects are apprehended, detained and interrogated abroad, yet on behalf of the captor state; moreover, the victims are foreign nationals.[117]

Captor states try to accomplish two goals through extraordinary renditions. First and above all, they prefer increased harshness during interrogations to yield the maximum benefit on intelligence gathering grounds. However, constitutional and legal guarantees in combination with effective enforcement mechanisms almost prohibit the use of severe techniques in their territory, as victims may ask for judicial protection. This explains the second goal, which is to fully deprive the transferred from access to their judicial system where they can challenge their treatment during detention.[118] In other words, the captor state tries by all means to avoid its domestic legislation and to create a 'legal *lacuna*'.

The detainees' lives are in jeopardy since extraordinary rendition reduces their legal protection to the bare-minimum, permitting grave human rights violations. Extraordinary renditions should be undoubtedly placed among practices to disappear individuals, as they are 'designed to evade public and judicial scrutiny, to hide the identity of the perpetrators and the fate of the victims'.[119] They also constitute 'a degrading and dehumanizing practice for the victims',[120] because they are aware of their inability to reach both the outside world and also the judiciary. It is this inability that equates extraordinary renditions to enforced disappearances and differentiates them from mere international abductions.[121]

The UNWGEID came round to this view in 2004 and further mentioned that the practice constitutes a breach to the 1992 Declaration.[122] It also came across extraordinary renditions when examining the complaint of Maher Arar. This is one of the very few complaints that have gained publicity until now, mainly due to efforts made by Canadian human rights NGOs and by the victim's wife. The victim, a national of Canada, was detained in an American airport while returning from Tunisia.[123] He

was then transferred to Syria to be interrogated on his alleged links with Al-Qaeda, where he was kept nearly for a year. After his release Arar brought his claims before American Courts, only to be rejected on jurisdictional grounds.[124] This is indicative of the juridical difficulties the practice entails. So far, national courts have rejected similar claims based either on lack of jurisdiction or on aspects of national security.[125] Apart from some exceptions,[126] domestic jurisprudence has generally arrived at unsatisfactory judgments for the victims.

The Convention's application in the case of incommunicado detentions and extraordinary renditions is beneficial for the protection of detainees. It has already been alluded, that states tried to limit their human rights obligations by derogating from major international instruments on grounds of public emergency. Such derogations affected mostly the right to liberty and the due process guarantees attached to it. As noted above, they were enforced with laws which permitted prolonged incommunicado detentions and unlawful renditions.[127] At this point, the 2007 Convention may prove to be a useful tool, enhancing the victims' protection. The characterization of the discussed methods as enforced disappearances has at least two obvious advantages. First of all, their complex nature will be acknowledged. It is a more realistic and systematic approach which affirms the danger for the victims since they cannot inform anyone of their current status and seek help. Secondly, under the 2007 Convention the right not to be subjected to enforced disappearance is non-derogable. Article 1(2) of the Convention contains an absolute prohibition on enforced disappearances precluding derogations under any possible justification. Prolonged incommunicado detentions and extraordinary renditions will therefore be utterly outlawed since the Convention leaves no space for a gray area in this field. Although human rights commentators link these two practices to enforced disappearances in general, they haven't examined them under the Convention's framework, although the latter provides a straightforward and sound response to the current international concerns. Hence, it constitutes a valuable underpinning for the individuals' protection.

VI. CONCLUSION

Enforced disappearances are a widespread phenomenon that involves extreme suffering for the victims, and consequently it has attracted much international attention. Public ignorance alongside conceptual difficulties are just a few of the reasons which delayed the adoption of a legally binding instrument to regulate enforced disappearances. However, the international UN human rights regime reserved a unique approach to

enforced disappearances. Indeed, the preference towards creating the UNWGEID over drafting a Convention seemed more reasonable in the 1980s; almost 30 years later, enforced disappearances have generated considerable concern on a global scale, and the international community is finally ready enforce a legally binding instrument.

The adoption of the 2007 Convention is by far the most prominent response to enforced disappearances. The effectiveness and success of the instrument cannot be measured, since it has entered into force only a few months ago. Nevertheless, it is fair to say that the Convention is a rather powerful instrument as it provides a comprehensive definition for enforced disappearances and pronounces a right of all peoples not to be subjected to enforced disappearances.

The phenomenon's particularity lies in the authorities' refusal to disclose information on the victims' fate or whereabouts; this refusal renders them essentially helpless. This aspect is well treated by the definition and the Convention as a whole; more specifically, any practice that is mainly characterized by an attempt to efface traces of the victim can be classified as an act of enforced disappearance. This approach signals that methods such as extraordinary renditions and incommunicado detentions can be characterized as enforced disappearances. Therefore, the Convention's scope of application is not confined within the limits of methods that have already emerged. It may also cover new practices that are not yet standardized. Overall, the 2007 Convention signifies considerable progress in the field of international human rights law. This is affirmed through the provided definition, which is flexible enough to respond to present demands and also to adapt to future legal challenges.

References

[1] ECOSOC 'Civil and Political Rights, Including Questions of: Disappearances and Summary Executions' UN Doc E/CN.4/2002/71 (8 January 2002) para 8.
[2] Jackson Nyamuya Maogoto, 'Now You See, Now You Don't: The State's Duty to Punish Disappearances and Extra-Judicial Executions' (2002) 3 AustlInt'lLJ 176, 182.
[3] ECOSOC (n 1), para 7.
[4] Christopher K Hall, 'Enforced disappearance of persons' in Otto Triffterer (ed), *Commentary on the Rome Statute of the International Criminal Court, Observers' Notes, Article by Article* (Nomos Verlagsgesellscaft 1991)151.
[5] L Roniger and M Sznajder, *The Legacy of Human-Rights Violations in the Southern Cone, Argentina, Chile and Uruguay* (OUP 1999) 7-28.
[6] National Commission on the Disappearance of Persons (CONADEP), '*Nunca Más* (Never Again) Report', 1984.

[7] Organization of American States, 'Report on the Status of Human Rights in Guatemala' Inter-American Commission on Human Rights, OEA/SerL./V/II53, Doc 22, rev 2 (Washington DC 13 October 1981) Chapter II; Organization of American States, 'Report on the Status of Human Rights in Chile' Inter-American Commission on Human Rights, OEA/SerL/V/II.34, Doc 21, corr 1 (Washington DC 25 October 1974) Chapter IX. The Pinochet regime did not provide information to the Commission, despite its numerous visits and special reports. Chile remained for many years a land of violence.
[8] ECOSOC (n 1) para 12.
[9] Ibid para 13.
[10] UNCHR 'Report of the Working Group on Enforced or Involuntary Disappearances' (21 January 2004) UN Doc E/CN.4/2004/58, paras 322-323.
[11] *Coronel and others v Colombia*, Human Rights Committee (HRC) (2002) UN Doc CCPR/C/76/D/778/1997; *Vicente and others v Colombia*, HRC (1997) UN Doc CCPR/C/56/D/612/1995.
[12] *González et al ('Cotton Field') v Mexico*, Inter-American Court of Human Rights Series A No 205 (16 November 2009); UN Committee on the Elimination of Discrimination against Women 'Report on Mexico produced by the Committee on the Elimination of Discrimination against Women under article 8 of the Optional Protocol to the Convention and reply from the Government of Mexico' (27 January 2005) UN Doc CEDAW/C/2005/OP.8/Mexico, paras 61-110.
[13] UNGA Res 33/173 (20 December 1978).
[14] Ibid operative clause 2.
[15] ECOSOC Res 1979/38 (10 May 1979).
[16] UNCHR, 'Report of the Working Group on Enforced or Involuntary Disappearances' (1981) UN Doc E/CN.4/1435, Annex III.
[17] UNCHR 'Question of Missing and Disappeared Persons' (1980) UN Doc E/CN.4/Res/1980/20, operative clause 1.
[18] Rosemary Foot, *Rights Beyond Borders, The Global Community and the Struggle over Human Rights in China* (OUP 2000) 40.
[19] William Korey, *NGOs and the Universal Declaration of Human Rights, "A Curious Grapevine"* (Palgrave 2001) 250.
[20] Ann Marie Clark, *Diplomacy of Conscience Amnesty International and Changing Human Rights Norms* (Princeton University Press 2001) 17.
[21] Jeroen Gutter, *Thematic Procedures of the United Nations Commission on Human Rights and International Law: in Search of a Sense of Community* (School of Human Rights Research 2006) 82-89.
[22] UNCHR (n 16) para 3.
[23] Ibid para 30.
[24] JD Livermore and BC Ramcharan, '"Enforced or Involuntary Disappearances": An Evaluation of a Decade of United Nations Action' (1989-1990) 6 CHRYB 217, 217-230.
[25] *Bleier v Uruguay,* HRC (1982) UN Doc CCPR/C/15/D/30/1978.
[26] Ibid para 13.
[27] *Quinteros v Uruguay*, HRC (1983) UN Doc CCPR/C/OP/2, para 11.
[28] *Arévalo v Colombia*, HRC (1989) UN Doc CCPR/C/37/P/181/1984, para 10; UNHRC, Sixteenth Session 1982, 'General Comment No. 06: The right to life (art.6)' (30 April 1982) para 4.
[29] Linda Drucker, 'Governmental Liability for "Disappearances": A Landmark Ruling by the Inter-American Court of Human Rights' (1988-1989) 25 Stan J Int'l L 289.

[30] *Velásquez-Rodriguez v Honduras*, Inter-American Court of Human Rights Series C No 4 (29 July 1988).
[31] Claudio Grossman, 'Disappearances in Honduras: The Need for Direct Victim Representation in Human Rights Litigation' (1991-1992) 15 HastingsInt'l&CompLRev 363.
[32] *Velásquez-Rodriguez v Honduras* (n 30) para 131; *Fairén-Garbi and Solís-Corrales v Honduras*, Inter-American Court of Human Rights Series C No 6 (15 March 1989) para 127.
[33] *Velásquez-Rodriguez v Honduras* (n 30) paras 130, 188.
[34] Ibid (n 30) para 158.
[35] *Godínez-Cruz v Honduras*, Inter-American Court of Human Rights Series C No 5 (20 January 1989) para 154b(i).
[36] *Caracazo v Venezuela*, IACHR Series C No 58 (11 November 1999); *Blanco Romero v Venezuela*, Inter-American Court of Human Rights Series C No 137 (28 November 2005).
[37] *Caballero-Delgado and Santana v Colombia*, Inter-American Court of Human Rights Series C No 22 (8 December 1995).
[38] *Fairén-Garbi and Solís-Corrales v Honduras* (n 32) paras 157-158.
[39] UNGA Res 47/133 (18 December 1992) UN Doc A/47/49. The UNWGEID was assigned to report any obstacles during its application and to monitor its implementation in the member-states (an expansion of its mandate), UNCHR 'Report of the Working Group on Enforced or Involuntary Disappearances' (1994) UN Doc E/CN.4/1994/26, para 74.
[40] ECOSOC (n 1) para 44.
[41] OAS, 'Inter-American Convention on Forced Disappearance of Persons' General Assembly Res AG Res 1256 (XXIV-O/94) (Belem Do Para 6 September 1994).
[42] T Scovazzi and G Citroni, The Struggle against Enforced Disappearance and the 2007 United Nations Convention (Martinus Nijhoff Publisher and VSP 2007) 191.
[43] *Çakici v Turkey*, App no 23657/94 (ECHR, 8 July 1999); *Ismail Ertak v Turkey*, App no 20764/92 (ECHR, 9 May 2000); *Tas v Turkey*, App no 24396/94 (ECHR, 14 November 2000); *Çiçek v Turkey*, App no 25704/94 (ECHR, 27 February 2001).
[44] Irum Taqi, 'Adjudicating Disappearance Cases in Turkey: An Argument for Adopting the Inter-American Court of Human Rights' Approach' (2000-2001) 24 Fordham Int'l LJ 940, 966-967.
[45] *Kurt v Turkey*, App no 24276/94 (ECHR, 25 May 1998) paras 99, 107.
[46] Scovazzi and Citroni (n 42) 190.
[47] Carla Buckley, 'The European Convention on Human Rights and the Right to Life in Turkey' (2001) 1 HumRtsLRev 35, 55-64.
[48] In the *case of Timurtaş v Turkey*, the ECHR set a very high threshold of six and a half years, to presume the victim's death. *Timurtaş v Turkey*, App no 23531/94 (ECHR, 13 June 2000) paras 82-86.
[49] *Baysayeva v Russia*, App no 74237/01 (ECHR, 5 May 2007) para 119.
[50] *Mahmut Kaya v Turkey*, App no 22535/93 (ECHR, 28 March 2000) paras 342-343.
[51] UNCHR 'Report of the sessional working group on the administration of justice (1998) UN Doc E/CN.4/Sub.2/1998/19, Annex.
[52] UNCHR Res 46 (2001) UN Doc E/CN.4/2001/46, para 12.
[53] ibid.
[54] ECOSOC (n 1) para 97.

[55] UNCHR, 'Report of the Intersessional Open-Ended Working Group to Elaborate a Draft Legally Binding Normative Instrument for the Protection of All Persons from Enforced Disappearances' (2004) UN Doc E/CN.4/2004/59.
[56] UNCHR, 'Report of the Intersessional Open-Ended Working Group to Elaborate a Draft Legally Binding Normative Instrument for the Protection of All Persons from Enforced Disappearances' (2005) UN Doc E/CN.4/2005/66.
[57] UNGA 'Implementation of General Assembly Resolution 60/251 of 15 March 2006 Entitled "Human Rights Council"' (23 June 2006) 1st Session (2006) UN Doc A/HRC/1/L.2; UNGA 'Report of the Third Committee' (6 December 2006) 61st Session (2006) UN Doc A/61/448, para 27; UNGA Res 61/177 (12 January 2007) UN Doc A/Res/61/177.
[58] UNGA 'Verbatim Record' (20 December 2006) UN Doc A/61/PV.82, 3.
[59] UNCHR, 'Report of the Working Group on enforced or involuntary disappearances' (1998) UN Doc E/CN.4/1999/62.
[60] Characterizations varied from forced, enforced, involuntary, to political or even systematic. HM Kleinman, 'Disappearances in Latin America: A Human Rights Perspective' (1986-1987) 19 NYU J Int'l L & Pol 1033, 1033.
[61] Jan Egeland, 'Political "Disappearances" – A Challenge for Humanitarian Law' (1982) 51 Nordisk Tidsskrift Int'lRet 189, 189.
[62] UNGA Res 47/133 (n 39) operative clause 1.
[63] UNHCHR 'Enforced or Involuntary Disappearances' (Geneva 2009) UN Doc Fact Sheet No.6/Rev.3, 6.
[64] Gabriella Citroni and Tullio Scovazzi, 'Recent Developments in International Law to Combat Enforced Disappearances' (2009) 3*Revista Internacional de Direito e Cidadania* 89, 93.
[65] ECOSOC (n 1) 70.
[66] UNCHR, 'Report of the Working Group on enforced or involuntary disappearances' (1996) UN Doc E/CN.4/1996/38, para 55.
[67] UNCHR, 'Report of the Working Group on enforced or involuntary disappearances' (1983) UN Doc E/CN.4/1984/21, para 178.
[68] Report of the Working Group on enforced or involuntary disappearances (n 66) para 51.
[69] Citroni and Scovazzi (n 64) 92-93.
[70] UNCHR (n 56) para 17.
[71] UNCHR (n 55) para 20.
[72] UNCHR (n 56) para 19.
[73] *Gangaram-Panday v Suriname*, Inter-American Court of Human Rights Series C No 16 (21 January1994).
[74] This approach differs from the one the ECHR follows. For the ECHR, a disappearance is an 'aggravated form of arbitrary detention'. For that reason, the Court finds it important that the disappeared were last seen in custody by governmental authorities. The Convention focuses on deprivation of liberty as well, but it does not treat it as a prerequisite for a disappearance to occur. *Orhan v Turkey*, App no 25656/94 (ECHR, 18 June 2002) paras 265, 278; *Tekdag v Turkey*, App no 27699/95 (ECHR, 15 January 2004) paras 66, 68.
[75] UNCHR, 'Report of the Working Group on enforced or involuntary disappearances' (2004) UN Doc E/CN.4/2004/59, para. 21.
[76] *Cantoral-Benavides v Peru*, Inter-American Court of Human Rights Series C No 69 (18 August 2000) paras 90-91; *Neira-Alegría et al v Peru*, Inter-American Court of Human Rights Series C No 20 (19 January 1995).

[77] UNWGEID, 'General Comment on the Definition of Enforced Disappearance' para 7 <http://www2.ohchr.org/english/issues/disappear/links.htm> accessed 3 June 2012.
[78] For the sake of coherent argumentation it is discussed before the second one.
[79] *Castillo-Páez v Peru*, Inter-American Court of Human Rights Series C No 34 (3 November 1997) para 58.
[80] UNCHR, 'Report of the Intersessional Open-Ended Working Group to Elaborate a Draft Legally Binding Normative Instrument for the Protection of all Persons from Enforced Disappearances' (2006) UN Doc E/CN.4/2006/57, paras 16-29.
[81] Civil and Political Rights, Including Questions of: Disappearances and Summary Executions (n 1) para 73.
[82] James Crawford, 'Human Rights and State Responsibility' (12th Raymond & Beverly Sackler Distinguished Lecture Series, Thomas J. Dodd Research Centre, University of Connecticut, 25 October 2006) para 1 <http://doddcenter.uconn.edu/dd/events/sackler/Crawford%20transcript.pdf> accessed on 3 June 2012.
[83] Ibid para 5.
[84] Ibid para 2.
[85] The case is more complicated in countries where both state actors and non-state actors (usually opposition groups) resort to disappearances, so as to weaken the opponents.
[86] UNCHR, 'Report of the Intersessional Open-Ended Working Group to Elaborate a Draft Legally Binding Normative Instrument for the Protection of all Persons from Enforced Disappearances' (2003) UN Doc E/CN.4/2003/71, para 35.
[87] Susan McCrory, 'The International Convention for the Protection of All Persons from Enforced Disappearance' (2007) 7 Hum Rts L Rev 545, 551.
[88] *Masacre de Pueblo Bello v Colombia*, Inter-American Court of Human Rights Series C No 140 (31 January 2006) paras 111-153.
[89] UNCHR (n 80) para 91.
[90] Federico Andreu-Guzmán, 'The Draft International Convention on the Protection of All Persons from Forced Disappearance' (2001) 62-63 *ICJ Review – Impunity, Crimes against Humanity and Forced Disappearance* 73, 79-80.
[91] UNCHR (n 80) para 93. See also, the statement made by the United Kingdom after the Convention's adoption: UNGA 'Verbatim Record' (n 58) 2.
[92] UNWGEID (n 77) para 7; UNCHR, 'Report of the Working Group on Enforced or Involuntary Disappearances' (2002) UN Doc E/CN.4/2002/79 para 365.
[93] ECOSOC (n 1) paras 73-74.
[94] UNHRC, 'Joint Study on Secret Detention of the Special Rapporteur on Torture & Other Cruel, Inhuman or Degrading Treatment or Punishment, the Special Rapporteur on the Promotion and Protection of Human Rights & Fundamental Freedoms while Countering Terrorism, the Working Group on Arbitrary Detention & the Working Group on Enforced or Involuntary Disappearances' (2010) UN Doc A/HRC/13/42, para 28.
[95] *Bazorkina v Russia*, App no 69481/01 (ECHR, 27 June 2006) para 167.
[96] OAS (n 41) article 2.
[97] *Bámaca-Velásquez v Guatemala*, Inter-American Court of Human Rights Series C No 70 (25 November 2000) para 121(b), (d); *The 19 Tradesmen v Guatemala*, Inter-American Court of Human Rights Series C No 108 (5 July 2004) para 84(a), (h).
[98] Nigel S Rodley, *Treatment of Prisoners under International Law* (3rd edn, OUP 1999) 333-334.

[99] M Bator and J Vorenberg, 'Arrest, Detention, Interrogation and the Rights to Counsel: Basic Problems and Possible Legislative Solutions' (1966) 66 CLMLR 62, 71-72.
[100] Economic and Social Council 'Standard Minimum Rules for the Treatment of Prisoners' Res 663C (XXIV) (31 July 1957) Annex 1, principles 37-38; UNGA, 'Body of Principles for the Protection of All Persons under Any Form of Detention or Imprisonment' (9 December 1988) UN Doc A/Res/43/173, Annex, Principles 16, 19; Tullio Treves, 'The UN Body of Principles for the Protection of the Detained or Imprisoned Persons' (1990) 84 AmJInt'lL 578, 581-582.
[101] UNCHR, 'Report of the Working Group on enforced or involuntary disappearances' (2003) UN Doc E/CN.4/2003/70, 3.
[102] Sarah Joseph et al, *Seeking Remedies for Torture Victims, A Handbook on the Individual Complaints Procedures of the UN Treaty Bodies* (World Organization Against Torture 2006) 170.
[103] S Marks and A Clapham, *International Human Rights Lexicon* (OUP 2005) 77.
[104] Roland D Crelinstern, 'Violence in Prisons/Torture' in Wilhelm Heitmeyer and John Hagan (eds) *International Handbook of Violence Research* (Kluwer Academic Publishers 2003) 189.
[105] UNCHR, Forty-fourth Session 1992, 'General Comment No. 20: Replaces general comment 7 concerning prohibition of torture and cruel treatment or punishment (Art.7)' (10 March 1992) para 6.
[106] *Lopez Burgos v Uruguay*, HRC (1981) UN Doc *CCPR*/C/1 3/D/52/1979 paras 11.5-11.6.
[107] *Suárez-Rosero v Ecuador*, Inter-American Court of Human Rights Series C No 35 (12 November 1997) paras 48-52.
[108] International Committee of the Red Cross, 'Report on the Treatment of 14 'High Value Detainees' in CIA Custody' 2007 (Washington 14 February 2007) 7, 8 < http://assets.nybooks.com/media/doc/2010/04/22/icrc-report.pdf> accessed 3 June 2012.
[109] *Caldas v Uruguay*, HRC (1979) UN Doc *CCPR*/C/19/D/43/1979 para 14.
[110] *Aksoy v Turkey*, App no 21987/93 (ECHR 26 November 1996) para 78.
[111] *Mónaco v Argentina*, HRC (1995) UN Doc *CCPR*/C/53/D/400/1990 paras 2.1, 10.4; *Loayza-Tamayo v. Peru*, IACHR Series C No 33 (17 September 1997) para 46(c), (e), (f).
[112] *Jegatheeswara Sarma v Sri Lanka*, HRC (2003) UN Doc *CCPR*/C/78/D/950/2000 para 9.6.
[113] Marguerite Feitlowitz, *A Lexicon of Terror: Argentina and the Legacies of Torture* (OUP 1998) 71; Paul Michell, 'English-Speaking Justice: Evolving Responses to Transnational Forcible Abduction after Alvarez-Machain' (1996) 29 Cornell Int'lLJ383, 389-390. Extraordinary renditions stood at the beginning as a synonym to the 'Ker-Frisbie Doctrine', a modern version of the *male captus bene detendus* maxim.
[114] *Goiburú and others v Paraguay*, Inter-American Court of Human Rights Series C No 153 (22 September 2006).
[115] Joan Fitzpatrick, 'Rendition and Transfer in the War against Terrorism: Guantánamo and Beyond' (2002-2003) 25 LoyLAInt'l&CompLRev 457, 461.
[116] UNHRC, 'Report of the Working Group on Enforced or Involuntary Disapearances' (2007) UN Doc A/HRC/4/41, para 454.
[117] Scovazzi and Citroni (n 42) 42.
[118] Mario Silva, 'Extraordinary Rendition: A Challenge to Canadian and United States Legal Obligations under the Convention against Torture' (2008-2009) 39 CalWInt'lLJ 313, 317.

[119] Joint Hearing Before the Subcommittee on International Organizations, Human Rights: An Oversight and the Subcommittee on Europe of the Committee on Foreign Affairs House of Representatives 110th Congress First Session, *Extraordinary Rendition in U.S. Counterterrorism Policy: The Impact on Transatlantic Relations* (US Government Printing Office 2007) 4.
[120] Council of Europe (CoE), *Secret detentions and unlawful inter-state transfers of detainees in Europe (1st report Marty)* (CoE Publishing 2008) 45.
[121] The CoE referred to the CIA Rendition Program as 'a spider web of disappearances'. CoE Parliamentary Assembly Ordinary Session, *Official Report of Debates,* (3rd PartVol III CoE Publishing 2006) 607.
[122] UNCHR, 'Report of the Working Group on enforced or involuntary disappearances' (2005) UN Doc E/CN.4/2005/65, paras 13, 368.
[123] Joint Hearing Before the Subcommittee on International Organizations, Human Rights: An Oversight and the Subcommittee on Europe of the Committee on Foreign Affairs House of Representatives 110th Congress First Session (n 119) 357.
[124] *Maher Arar v John Ashcroft et al*, 532 F3d 157 (2D CIR 30 June 20) paras 192-193.
[125] American Courts have dismissed most of these cases, relying upon the 'state secrets privilege'. *El-Masri v Tenet et al*, US 437 F Supp 2d 530, 541 (ED Va 12 May 2006). The case is now pending in the Inter-American Court of Human Rights.
[126] In November 2009, an Italian Court convicted *in absentia* 23 CIA agents for the extraordinary rendition of Abu Omar.
[127] However, after a careful reading of the applicable international law, the validity of these derogations may be challenged. UNCHR, 'Report of the Working Group on Arbitrary Detention' (2005) UN Doc E/CN.4/2005/6, para 76; OAS, 'Report on Terrorism and Human Rights' Special Reports of the IACommHR OEA SerL/V/II.116, Doc 5 rev 1 corr, para 24 (Washington 22 October 2002); UNCHR, 'General Comment No. 29, States of Emergency (Article 4)' (31 August 2001) UN DocCCPR/C/21/Rev.1/Add.11, para 13(a), (b).

RECOGNITION OF CONTRACTS AS INVESTMENTS IN INTERNATIONAL INVESTMENT ARBITRATION

Velimir Zivkovic*

The issue of recognition of contractual rights as protected investments in international investment arbitration, primarily under the auspices of ICSID, has sparked divergent approaches in case law. Treatment of certain contracts and the criteria used differ, which leads to unwelcome consequence of lowering legal certainty in a very sensitive issue. The aim of this paper is to contribute to enhancement and clarification of legal reasoning in this area, with a special focus on the criteria to be used and on sales contracts which are particularly controversial in practice. This is done through the analysis of the current state of affairs which is followed by a proposition of a new model of criteria which could present a beneficial compromise between the existing models and increase certainty.

TABLE OF CONTENTS

I. INTRODUCTION

The issue of what can be recognized as an investment and given the corresponding protection under the rules of international investment law is both long lasting and highly contentious. While categories of investments which can be deemed 'easily recognizable' do exist, they only form a (relatively) stable core of the term. Its outer limits are far from settled.

* LL.M, MJur (Oxon). PhD Candidate in Business Law, University of Belgrade Faculty of Law and a Technical Expert at OPTIMUS Center for Good Governance in Belgrade.

Contractual rights are a good example of the shifting boundaries of investment protection. Historically, the recognition of possibility of international law regime to deal with contracts including private entities, started with the*Serbian Loans* case of the Permanent Court of International Justice in 1929,[1] was an impetus that eventually grew to creation of investment dispute settlement mechanisms we recognize today.[2] Even before that, Permanent Court of Arbitration recognized the possibility to expropriate contractual rights as assets in *Norwegian Shipowners' Claims* case.[3]

Despite widespread contemporary acceptance of intangible assets (and indeed contractual rights) as protected investments in international investment law, the key question - *which* contractual rights are to be protected? - is not decisively settled. Actual examples from the practice of investment protection offer intriguing examples of dilemmas that need to be solved – is commercialization of tobacco products a form of investment? What about contracts for retrieving shipwrecked artefacts and selling them later? Or maybe expenditures made prior to actually obtaining a contract with a host state? The debate about these issues is, of course, far from a purely academic one. Recognizing that a certain contractual right (or as is commonly abbreviated, 'contract') is a protected investment can mean a world of difference for a foreign investor in terms of available protection. It can mean a difference between litigating in a possibly slow and/or biased court system of a host State and having recourse to arbitral proceedings before a specialised and well-known international institution such as International Centre for Settlement of Investment Disputes (ICSID). The topic of this paper is to contribute to enhancement and clarification of legal reasoning in this area, with a special focus on the criteria to be used and also on sales contracts, a category which is particularly controversial in practice.

The discussion ahead consists of three parts. The first part deals with the general issues of recognizing a contract as a protected 'investment' and also explores the broader issue of the criteria used to recognize protected investments in international investment arbitration and the proposed future model of recognizing investments. The second part is focused on sales contracts, their current status and the potential use of the suggested model when facing the issue if a particular sales contract is an investment or not. The third part proposes certain guidelines for the future regarding the discussed issues. These guidelines are based on the conclusions reached in previous parts.

II. RECOGNITION OF CONTRACTS AS INVESTMENTS

1. *General Remarks*

Contractual rights are often present in investment disputes. This is in accordance with the general trend of the changing nature of investments, which is evolving from the old natural resource exploitation and ownership of production facilities to more modern forms, such as service agreements.[4]

It is possible to make a list of commonly encountered forms of contract. The types of contracts usually considered as having a character of investment in bilateral investment treaties (BITs) when listed, ICSID practice and doctrinal writings are: construction, turnkey, management/service, production, profit-sharing, leasing, technology/know-how transfer, and joint-venture contracts.[5] Other important contracts are public concession agreements, but one should bear in mind that they by their nature include the host State and thus generally pose no particular problems in being identified as protected by international investment law. Some authors include loans in the group of protected contracts,[6] which is technically true (loans are, of course, contracts), but they are usually classified as a separate group of investments along with other financial instruments.

But the above is merely an informative list, a recapitulation of what can be found in legal instruments and case law. The key issue is not just to identify these contracts. The crucial question is *why* are these contracts recognized? Only if the criteria which led to this are known and understood properly it can be said that there will be enough predictability to ascertain if in a future case a contractual right is likely to be recognized as an investment. And at this point the analysis necessarily becomes somewhat broader.

The criteria proposed for recognizing contractual rights cannot be separated from the criteria which will be used in general to evaluate if there is an investment. Despite certain specificities (some will be suggested in the section dealing with sales contracts) the underlying core criteria will necessarily be the same for different types of investments. Thus, examining what these criteria are and how the overall approach can be improved has a wider relevance for the notion of 'investment' in international investment law. Of course, due effort will be made to frame the findings and conclusions within the context of contractual rights as much as possible.

2. *Current Criteria for Recognition*

Two preliminary notes should be made. Firstly, it is not the author's objective to try and ascertain criteria which led to recognition of certain contracts or other transactions as investments in particular BITs. Bearing in mind their vast number and a plethora of circumstances which might influence specific definitions in particular BITs, such task is indeed out of the scope of this paper. What can be said is that developments in practice do influence BITs and definitions therein, but such influence always has to be examined on a case-by-case basis.

Secondly, the discussion that follows is primarily centred on establishing ICSID jurisdiction and problems with the notion of 'investment' found in Article 25 of the Washington Convention.[7] One reason for this is that ICSID is (in terms of caseload and dispute values) the most important forum for resolving investment disputes.[8] The second one is that this sensitive area is currently marred by divergent jurisprudence.

Regarding non-ICSID arbitral tribunals, generally speaking two different situations can exist. In some cases the issues discussed are not so prominent, as there might be no need to deal with Article 25 and parties are generally free to arbitrate about whatever they agree upon (subject, of course, to potentially mandatory rules on arbitrability and similar provisions). In such cases the jurisdiction can be established, for example, merely by interpreting the BIT, which would be the only step that needs to be taken when establishing jurisdiction. In other situations, most prominently NAFTA cases, the situation might again revolve around establishing an objective meaning of the term 'investment'. In such cases, it is legitimate to ask whether or not the discussion of ICSID cases and jurisprudence formed therein might be of influence or even of precedential value? Regarding NAFTA, it can be said that there is a growing tendency to look upon investment treaty arbitration and awards made as a single phenomenon and not to insist on differences between jurisdictions.[9] Decisions of arbitral tribunals dealing with alleged breaches of NAFTA provisions confirm the tendency to give due consideration and careful examination to previous ICSID awards as well.[10] Bearing this in mind, it can be said that the discussion that follows can also be of wider (non-ICSID) relevance.

The starting point for dealing with the issues of ICSID jurisdiction is Article 25(1) of the ICSID Convention, which states:

> The jurisdiction of the Centre shall extend to any legal dispute arising directly out of an investment, between a Contracting State (or any constituent subdivision or agency of a Contracting State designated to the Centre by that State) and a national of another

> Contracting State, which the parties to the dispute consent in writing to submit to the Centre ...

The term 'investment' (unlike 'national', for example) is not defined in further text. Despite some differing opinions, it is commonly considered that this was not caused by a mere lack of agreement, but was an intentional compromise.[11]

However, this led to a situation in which it was not clear what role the term was supposed to play and what was its inherent meaning, if it had one in the first place. To fully understand the possible approaches to this issue, one should remember that 'investments' (including contractual rights or not) are already defined in BITs. What has not been settled so far is if this definition of, for example, a contractual right as an investment found in a BIT is also the one relevant for Article 25(1) (which would mean that Article 25 term has no inherent, objective meaning) or this definition only constitutes *consent* as required by Article 25(1) while 'investment' presents a separate jurisdictional hurdle. The debate became even more complicated because of differing opinions in case-law how is this objective, inherent meaning of 'investment' to be established if the tribunal considers it to indeed exist.

The case-law dealing with these problems is substantial and diverse. It is not possible within the scope of this work to go into all the interesting factual or theoretical subtleties of particular cases. What is possible is to rationalize the general approaches of various tribunals into three groups: a) 'deference to consent'; b) 'benchmark'; and c) 'cumulative' group. In essence, all these approaches are located along the line which starts at total subjectivity and deference to consent, and ends at the strictly objective approach with the need for cumulative fulfilment of additional (varying) objective preconditions in a manner resembling a checklist.

Decisions in the first group[12] practically equate consent and investment – if the BIT proclaims something to be an investment, this should suffice for all purposes of Article 25(1). The other two groups of decisions share a different starting premise – there is something more in Article 25(1) that needs to be fulfilled and that is the requirement of 'investment' which should be ascertained by some objective criteria.

Reasoning based on objective criteria is usually considered to have originated in *Fedax v. Venezuela*.[13] It seems that in early ICSID cases the tribunals were not much willing to deliberate about the issue.[14] The *Fedax*tribunal turned to the writings of Professor Schreuer and concluded that the basic features of investment (for Article

25(1) purposes) were a certain duration, a certain regularity of profit and return, assumption of risk, a substantial commitment, and significance for the host State's development.[15] These criteria were followed by the tribunal in the *Salini v. Morocco*[16] case (giving the name to the so-called Salini test), with the exception of the need for certain regularity of profit and return, while also noting the need to assess all these criteria globally, in light of their interdependence.[17]

However, variations that developed in this general approach are not irrelevant. Cases in the 'benchmark' group[18] suggest that the presence of certain criteria is providing only exemplary guidance for the tribunal - 'benchmarks or yardsticks'[19] to help the tribunal in deciding, while it stays as flexible as possible. Cases in the 'cumulative' group state the requirement that all the criteria *need* to be present in order to find an investment – but they disagree on what these criteria are. Cases in this group revolve around the Salini test, but the number of criteria is either three (Salini test minus 'contribution to the host State'criterion),[20] four,[21] five (adding back the 'regularity of profit and return' criterion)[22] or even six (adding legality and good faith to the Salini criteria).[23]

It is not hard to see that such a confusing state of jurisprudence is seriously infringing the predictability of outcomes and thus also legal certainty. In the end, achieving investment protection before ICSID for a potentially high value contract might depend on the doctrinal inclinations of the arbitral tribunal, and not on settled legal principles. Therefore, it is submitted that there should be a single approach in determining jurisdiction, and in the author's opinion that approach should be an objective, semi-cumulative, three criteria test that will be elaborated below. As a side note, one should be aware that the lack of formal binding precedent doctrine in ICSID arbitration might be an obstacle to ever achieving totally unified approach. However, with the attitude that was exhibited, for example, by the *Bayndir* and *Saba Fakes* tribunals and which endorses following established and consistent case-law in comparable situations,[24] homogeneity of case law can be largely achieved. Such development are already noted and supported in doctrine.[25]

3. *The Proposed Model*

It is first necessary to see why the approach should not be based on the total deference to consent. It might seem that such an approach has some compelling arguments to support it. Article 25 (1) of the ICSID Convention simply speaks of disputes arising out of an investment without

further qualifications about an investment. Article 25 (4) sanctions the freedom of the contracting State to exclude whole classes of disputes from their consent to jurisdiction and thus clearly confirms the principle of party autonomy. On the basis of such premises, it is not easy to see why then the parties should not be absolutely free to define what an investment is. If contracting Stateshave a strong interest in giving BIT/ICSID arbitration protection to a particular form of transaction, should Article 25 stand in their way? If such protection is under the circumstances important for the economic development of a certain country (for example, as a tool to attract particular foreign businesses) would that not mean that Article 25 would contravene the Preamble of the Convention (which sets economic development as a primary goal) and undermine its aims? In light of such questions, it can be seen why some tribunals accepted total deference to parties consent or why, for example, Professor Mortenson suggests that whatever parties considered an investment, as long as it is 'colorably economic', should be considered to be an investment for the purposes of Article 25.[26]

But this simply cannot stand. While it is true that establishing whether or not a certain transactions falls within what the parties agreed is an investment is a *necessary* condition of establishing ICSID jurisdiction, it should not be a *sufficient* condition. Two main arguments speak against unrestricted deference. Firstly, this would mean that the term 'investment' does not have and can never have any inherent meaning for the purposes of an institution intentionally created to deal with investments. Although one can accept that legal and economic definitions of an investment may differ, this cannot mean that they differ so much that former is actually *tabula rasa* to be written by the Contracting States over and over again while the latter has well-known (albeit sometimes blurry) borders. Contracting States of the ICSID Convention did not create ICSID in order to resolve all sorts of 'economic' or 'business' disputes, but only 'investment' disputes. This is not to say that creating a new, wide reaching dispute resolution centre aimed at 'business' disputes in general would be illegitimate or unwarranted, but simply that it is not what ICSID is. Trying to 'transform' it to something through (the lack of) jurisdictional thresholds should thus be prevented.

Secondly, one should consider what would be the practical consequences of accepting unrestricted deference. Wide acceptance of economic activities as investments could lead to many transitory and fringe activities suddenly becoming investments.[27] This would potentially (or even likely) lead to the opening of the floodgates and undesirable massive increase in investment litigation. As the trend of increase in cases is already a constant in international investment arbitration, pushing the

process even more could easily lead to the system becoming hopelessly overstretched and, ultimately, inefficient.

In conclusion, establishing if a transaction is an investment for the purposes of a relevant BIT is essential to establish if there is the *consent* required by the Article 25(1) but not more than that. In order to find this consent, a tribunal needs to interpret the relevant BIT in accordance with the rules of public international law and the circumstances of the particular case, but this analysis remains separate from finding of an 'investment' for the purposes of Article 25.

It is thus necessary to turn to establishing the inherent meaning of that term. It is common ground that certain criteria need to be established in order to 'fill' the term 'investment' with some meaning. Two key issues must be resolved: first, what these criteria should be and second, how one should characterize their nature and mutual interdependence.

As for the number and contents of the criteria to be applied, as seen above, the Salini test is the starting point. However, before dealing with the problem of the actual variant of the test that would be preferable, one can question whether the Salini criteria are to be taken as a starting point at all. Indeed, there are serious conceptual objections to the Salini test, specifically that it is ideologically coloured and also unsuitable to comprise portfolio[28] investments.[29] It would be quite legitimate to propose a new, maybe more appropriate test. Still, it is submitted that *Salini* should be a viable starting point. It is widely (albeit somewhat differently) applied in the case-law, which makes it an obvious choice for creating and maintaining a line of consistent jurisprudence. In addition to that, in the author's opinion, criteria of duration, contribution and risk really do form a core of what should be expected from an investment.

However, this is not the case with the other criteria sometimes proposed. Criteria of legality and good faith, mentioned above, cannot be reconciled with Article 25(1) and should be rejected for reasons well explained in case law.[30] Regularity of profit and return also seem inadequate as a criterion. If this regularity *must* be achieved, than this is an unjustifiably high threshold, as the foreign investor is left without protection if its investment failed for commercial reasons, and that should not be relevant in this context. And if it is *expected* regularity, then this can easily be assimilated with the criterion of risk.

But the situation is not so clear regarding the criterion of host State development. This criterion has strong proponents, some going so far as to consider it a 'crucial' one.[31] Yet, it should be rejected. It is inherently

open to different interpretations and also subject to so much (substantiated) criticism that it can hardly play a meaningful role. Some respected scholars are clearly against the idea of the need to show any particular contribution to the host State apart from general benefits that investments usually bring.[32] The term itself is very vague. Even if the discussion is limited to just economic development, as opposed to broad notion of 'development', an arbitral tribunal will face itself with numerous possible definitions of what 'economic development' actually is.[33]

There are also other problems related to this criterion. There is no agreement whether this contribution needs to be 'significant' or not. Or how is it supposed to be measured – by the increase of the GDP of the host State, or somehow differently. All these difficult issues were put forward before arbitral tribunals and, regrettably, received different answers.[34] Not to mention how much more complicated the landscape would become if human rights and similar non-economic variables were also included into the notion of development.[35] Finally, as some arbitral tribunals aptly noted, this criterion is not only difficult to establish but is practically covered by the remaining three,[36] leaving the arbitral tribunals prone to confuse it with other criteria.[37]

Unfortunately, the issue remains hotly contested in arbitral practice, as illustrated by the decisions of the annulment committees in *Malaysia Historical Salvors*[38] (arguing for non-jurisdictional and flexible character of this criterion) and *Patrick Mitchell*[39] (arguing for essentiality of this criterion).[40] Dissenting opinion of judge Mohamed Shahabuddeen in *Malaysia Historical Salvors*, for example, offers a good illustration of differing positions accepted on this point by developed and developing states.[41]

Although 'development' is the aim stated in the ICSID Convention Preamble, it is submitted that transforming an (optimistic and diplomatic) wording found there into any sort of jurisdictional requirement is not just unusual, but also unwarranted and excessively troublesome. The Preamble remains a useful tool for interpretation and for establishing the aims of the Convention. But these aims are quite sufficiently advanced by applying the remaining three Salini criteria. Thus, regarding the objective criteria to be used, the first three Salini criteria (contribution, certain duration and an element of risk) are the foundation that is needed.

This leads us to the second crucial issue. In the light of the existing jurisprudence, it seems necessary to decide if these three criteria are to be simply 'benchmarks and yardsticks' or their presence needs to be established in each and every case.

It is submitted that if the three above mentioned criteria are accepted as the core of what constitutes an investment, then the tribunal should not treat them as mere guiding examples. These criteria should be present in every case. But this does not mean that the tribunals should drift into an overly strict approach and impose some general minimum 'quantities' of each of the criteria that must always be present. It should be borne in mind, for example, that the very author whose writings were the source of the criteria, Professor Schreuer, warned and criticized against accepting these general features of investment as a strict jurisdictional test.[42]

What the tribunal should do is that it should be attentive to what the *Salini v. Morocco* tribunal stated in addition to setting out the test, and that is the need to interpret the criteria in totality and having regard to the circumstances of a particular case. The tribunal should not be able to find the existence of an investment if one element is lacking. It is truly hard to argue, for example, that a contract of negligible duration can seriously be considered an investment. On the other hand, what the tribunal should be free to do is to conduct a balancing exercise. It should be free, while taking into account all the circumstances of the case, to decide what extent of fulfilment of each of the criteria is enough.

In the author's opinion, this precludes prescriptive statements such as that investment must have a minimum duration of a certain number of years or any similar 'quantification' of investments. Such requirements, that to some extent resemble a Procrustean bed, simply cannot be reconciled with the flexible approach which is needed. Additionally, one should also bear in mind that one similar 'quantification' threshold for an investment to exist (in the form of minimum value) was explicitly rejected during the drafting of the ICSID Convention.[43]

In summary, a foreign investor seeking to protect his contractual rights as investments before an ICSID should expect two distinct steps in proving that jurisdictional thresholds are met. The first step is establishing that its contractual right is covered by the relevant BIT. Generally, because of the usually broad wording used in BITs, this should not prove to be excessively hard in most cases (some potential issues will be mentioned below when discussing sales contracts). After this step, which establishes consent for the purpose of Article 25, the investor should prove the fulfilment of three criteria – certain duration, contribution and existence of risk, as to fulfil the quintessential conditions for the existence of an 'investment' within the meaning of Article 25(1).

When deliberating about this issue, the tribunal should decisively

determine the existence of all three criteria, but it should be flexible in assessing the extent to which these need to be fulfilled. The tribunal should take into account all the relevant circumstances of the particular case, with potential diversity of these not permitting any all encompassing or very specific guidance. If the contract in question is one already recognized in case law as constituting an investment, this should provide a useful guidance and also support the investor's case, but that fact alone should not be decisive. Even if one supports the development of harmonious ICSID jurisprudence, this should not come at the expense of doing justice to the facts of each particular case.

It can be predicted that in most cases what the States envisaged in a BIT as an investment, and what can be an investment for the purposes of Article 25 will coincide.[44] But not always, and the following section offers a good illustration.

III. SALES CONTRACTS

1. *General Remarks*

As can be concluded from above, adding new types of contract to the list of usually recognized investments should not be considered to be a finished process. One can be even less sure if a particular contract will be recognized as an investment in individual cases with potentially very differing circumstances. Yet, in contrast to this assertion, there seems to be a widespread trend in legal instruments and jurisprudence that ordinary commercial contracts, primarily contracts of sale, cannot fall within the definition of investment.[45] This trend is largely supported by doctrinal writings as authors emphasize these contracts as examples of what would usually fall out of the scope of protection when discussing definitions of investments in various instruments and for various purposes.[46]

It is submitted that such assertions should be taken with caution and that this general proposition should not be considered valid in all situations. The *ratio* behind it certainly has merit, in that it aims to prevent unwarranted and highly undesirable stretching of investment protection too far. However, it can be argued that there are situations in which what might be perceived as a sales contract should be recognized to be an investment, mainly due to its close relation with a previous investment already made.

It should also be said that the term 'sales contracts' is here used as a generic term to denote all kinds of trading arrangements which have as their key feature exchange of goods for payment, as opposed to various

types of services/labour/production arrangements. In the author's opinion, this group can offer a good view as how the term 'investment' can and should continue to evolve.

An excellent illustration of the general trend of exclusion of sales contracts is provided by the *Global Trading and Globex v. Ukraine* case. The tribunal, after resorting to previous ICSID decisions, concluded that pure commercial transactions, such as simple purchase and sale contracts, cannot be considered as investments for the purpose of Article 25.[47] As for the transactions in question in that particular case, which were rather typical trans-boundary CIF sales, the tribunal stated:

> ... these are each *individual* contracts, of *limited duration*, for the purchase and sale of goods, on a *commercial basis* and *under normal CIF trading terms*, and which provide for delivery, the transfer of title, and final payment, before the goods are cleared for import into the recipient territory; and that neither *contracts of that kind*, nor the moneys expended by the supplier in financing its part in their performance, can by any reasonable process of interpretation be construed to be 'investments' for the purposes of the ICSID Convention.[48]

Another example is the often cited non-ICSID case of *Petrobart v. Kyrgyz Republic*,[49] centred on the sale of gas condensate under the ECT. In that case, as is suggested, an investment would not exist if examined by an ICSID tribunal applying *Fedax* and *Salini*.[50]

The reasoning in the above cases, which should be supported, is not only relevant because it clearly prohibits using ICSID for the purposes of commercial arbitration. It is also relevant because it contains useful starting points in determining which sales contracts actually *should be* investments for the purposes of Article 25(1). And this is a very interesting issue that (bearing in mind the vast number of sales contracts being concluded and performed every day around the world) deserves careful consideration.

2. *Sales Contracts which Deserve Recognition*

Apart from looking at the case law, it is also useful to carefully examine what legal scholars have to say about transactions that should not be considered investments in any case. Here we find such notions as 'non-recurring transactions such as simple sales (...)'[51], 'ordinary transaction for purposes of a sale (...)'[52] and 'ordinary sales (...) unless some special feature of the transaction could objectively support a subjective stipulation

by the parties to that effect.'[53] Thus, both case law and doctrine point in the same direction - there needs to exist something special, something that would elevate the contract of sale to something more than just 'ordinary' or 'simple' one in order for an investment to exist.

In the author's opinion, that special element should be the complexity of the transaction combined with its firm relation to an existing investment in the host State. It is clear that simple trans-boundary sales have no place here. Not only that the case law clearly shows why the rejection of such contracts is justified in the context of Article 25(1), even regardless of the exact objective test one can use, but even establishing consent can easily prove to be an insurmountable obstacle for the claimant. For example, BITs usually speak of investments made *in the territory* of the host State.[54] A claimant who is trying to prove that a trans-boundary commercial sale has any meaningful relation with the territory of the host State would indeed be highly unlikely to succeed, even if it can at somehow subsume this transaction into some broader notion of 'contractual rights' as potentially found in a BIT.

But let us now turn to a different situation. For example, a foreign investor establishes a production facility in the host State. It might be, let us say, a pharmaceuticals producer aiming at supplying the host State health system, or using some advantages of the business climate in the host State to use it as a base for exporting its products. There should be no difficulty in finding that the production facility is an investment. But what should be the status of sales contracts concluded by the foreign investor to market the products of the facility?

From the outset the situation seems to be different from the one involving trans-boundary contracts. The potential territoriality criterion should no longer be an obstacle in framing the transaction within the BIT definition of an investment. What is more, many BITs include the example of a 'claim associated with an investment' in their illustrative lists of potential investments. It is quite conceivable that the claimant could here find strong support in proving that the BIT covers its transaction, and thus provides necessary consent for arbitration.

But this still leaves the second step. There must be certain duration, risk, and contribution by the claimant.[55] It can safely be stated that an individual, one-off transaction should remain out of the scope of investment protection even in these factual circumstances. Its duration is still practically negligible in the context of investments. That can also be said about the assumed risk. Even with a flexible mindset that a tribunal should assume, simple sales which were thrown out through the door

regarding pure trans-boundary sales should in any case not be allowed to come back through the window in this different context.

However, the situation should change once the transaction under scrutiny becomes significantly more complex, despite remaining in its essence a sales contract. An example of such a transaction is a high-value, long-term supply contract. In the example of a pharmaceutical company, this can mean supplying medicines to the health system of the host State for a number of years. It would involve a large number of recurring transactions under the general umbrella of a contractual framework. It is submitted that in such circumstances the test to be used in the second step of establishing jurisdiction can indeed be satisfied.

Bearing in mind the specific circumstances of a factual situation as presented (and more on that will be said below), the three criteria can be fulfilled to such an extent that it can be hard to see how not to regard a contract as an investment. What if this long-term contract was actually a key motive for the foreign investor to come to the host State in the first place? If the contract fails because the host State breached standards of protection prescribed in the BIT, it is not only that the risk of losing the profit from that particular contract materialized, it is also possible that there is the risk that the whole initial, primary investment is now at stake. Similarly, when discussing contribution, it is not only that the investor contributed, let us say, a particular quantity of goods to fulfil a particular contract, but in a sense it 'contributed' the whole initial investment, which was made exactly to contribute to fulfilling the contract. But even if the contract is not crucial in the sense indicated above, it should be sufficient to show that the three criteria are fully fulfilled and that the contract can be readily distinguished from ordinary contracts as described in *Global Trade and Globex v. Ukraine* or *Joy Mining v. Egypt*.

This is especially true if one takes into account, as the tribunal should, a special circumstance which exists here, and that is the close connection of the transaction with a recognized investment. This is an excellent example of a situation where the concept described as 'general unity of an investment operation'[56] comes into play. This concept is based on a premise that an overall project may qualify as an investment even though certain individual transactions comprising it do not, and that disputes arising out of these related transactions (even though they are not investments in themselves) still can be seen as arising out of an investment.[57] Reasoning of the*CSOB* tribunal explains this well:

Hence, a dispute that is brought before the Centre must be deemed to arise directly out of an investment even when it is based on a transaction

which, standing alone, would not qualify as an investment under the Convention, provided that the particular transaction forms an integral part of an overall operation that qualifies as an investment.[58]

It should be noted that non-ICSID cases also support such a conclusion. For example, in the *Franz Sedelmayer*[59] case it was emphasized that it was the close relation to an already existing investment that was crucial in determining whether some other right was an investment too.[60]

Therefore, the claimant has a strong additional argument that the contract, in essence, 'emanates' from an established investment and that this primary investment serves as a 'leverage' to propel it into the scope of protected investments.

Of course, despite establishing a possible theoretical model for recognition of such sales contracts as investments, it is impossible in advance to define sufficient duration or sufficient risk, or sufficient level of complexity to differentiate simple from complex transactions. What should also be borne in mind is that different branches of industry operate in different conditions. It can well be the case that the manufacturer will not have a long term, well-defined contractual arrangement. Instead, it might have to rely on sales which are occasional and far apart, but of very high value and of crucial importance for its business. Is there still an investment if such an isolated sales contract comes under scrutiny of an arbitral tribunal? Can high value and importance be that special element that will differentiate it from an 'ordinary' sale? It is hard to answer in abstract terms. This illustrates how tribunals can face truly hard cases in practice. But the general approach should remain the same.

In conclusion, when dealing with sales contracts, arbitral tribunal should remain committed to the prevailing approach that ordinary commercial sales are not investments for the purpose of Article 25(1). But this should not be the general conclusion for all sales transactions. It should be qualified with an exception that more complex and longer lasting transactions associated with existing recognized investments warrant recognition as investments themselves.

IV. CONCLUDING REMARKS

The issues analysed in this paper show that the topic of contractual rights in investment law is a dynamic one. In dealing with these rights in arbitral practice what should be sought is an adequate balance between flexibility and predictability. It is thus useful to propose certain guidelines for the future that should help in achieving such aim. Some of these are of a more

general nature, while some deal with particular groups of issues examined above.

Two general remarks seem warranted. Firstly, the divergence in the ICSID jurisprudence regarding very important issues of jurisdiction is a reason for serious concern. While achieving uniformity of practice through introducing binding precedents is hardly practically feasible, or even desirable, ICSID tribunals should be aware of their role in remedying this situation. Striving for uniformity in a reasonable manner should be the aim pursued in practice. Secondly, arbitrators dealing with contractual rights as potential investments should keep an open mind and be receptive to the ever changing forms in which foreign investments take place. The historical development of the notion of investment is a good illustration how flexible this area can be. But this open-mindedness is also warranted by the very essence of the idea of investment protection. Excessive restrictiveness can only lower the incentives for investing and in that way infringe the main goal – economic development.

As for recognizing which contracts are investments, apart from further harmonization of definitions in legal instruments which would certainly be beneficial, the way forward seems to be in accepting a common approach for determination if the conditions found in Article 25(1) are fulfilled. The approach that should be accepted is based on distinction between establishing consent (to be found in a BIT) and establishing if there is an investment, as both are distinctly required by Article 25(1). While the first issue remains largely in the area of general treaty interpretation, more guidelines can be given for the second element. The test to be applied should be based on the criteria of duration, risk and contribution. These should be fulfilled to a sufficient extent in every case, cumulatively, but the tribunal should be free to determine what the sufficient extent is. In committing this balancing exercise, it should pay attention to the specific circumstances of each particular case.

Regarding sales contracts, the existing general view that ordinary sales are not investments for the purpose of Article 25 should remain predominant. But this reasoning cannot be extended to all sales contracts. When a sales contract (which is, as the first condition, protected under the BIT) forms a part of a broader investment enterprise, clustered around a recognized investment, and by its other features also complies with the established test for recognition under Article 25(1), then it should be considered to be an investment and protected accordingly.

It is, of course, not easy to achieve the observance of these guidelines in practice. But it is something to be aimed for. It is the author's opinion that

application of the above guidelines would promote fair, balanced and reasonably predictable outcomes in deciding various issues that come before investment arbitration tribunals. And such outcomes would increase the protection of both legal and economic interests of investors and host States.

REFERENCES

[1] Payment of Various Serbian Loans Issued in France (France v Yugoslavia), PCIJ Rep Series No 20.
[2] Thomas Wälde, 'The Serbian Loans Case: A Precedent for Investment Treaty Protection of Foreign Debt?' in Todd Weiler (ed), International Investment Law and Arbitration: Leading Cases from the ICSID, NAFTA, Bilateral Treaties and Customary International Law (Cameron May 2005) 393.
[3] Permanent Court of Arbitration Norwegian Shipowners' Claims (Award,13 October 1922) 1 RIAA 307. See also August Reinisch, 'Expropriation' in Peter Muchlinski, Federico Ortino and Christoph Schreuer (eds), The Oxford Handbook of International Investment Law (OUP 2008) 411-412.
[4] Muthucumaraswamy Sornarajah, The International Law on Foreign Investment (3rd edn, CUP 2010) 11-6.
[5] Cristoph Schreuer, The ICSID Convention: A Commentary (CUP 2001) 138-139; United States Model Bilateral Investment Treaty 2012 < http://www.state.gov/documents/organization/188371.pdf > accessed 13 May 2012 (US Model BIT 2012) art 1; Sergey Ripinsky and Kevin Williams, Damages in International Investment Law (British Institute of International and Comparative Law 2008) 102; John P. Given, 'Malaysia Historical Salvors Sdn., Bhd. v. Malaysia: An End to the Liberal Definition of "Investment" in ICSID Arbitrations?' 31 Loyola of Los Angeles International and Comparative Law Review (2009) 467, 475; Prabhash Ranjan, 'Definition of Investment in Bilateral Investment Treaties of South Asian Countries and Regulatory Discretion' 26 Journal of International Arbitration 219, 225. See also OECD, The Multilateral Agreement on Investment: Draft Consolidated Text, DAFFE/MAI(98)7/REV1 11 (MAI) ch 2, art 2.
[6] Given (n 5) 475-6.
[7] The Washington Convention on the Settlement of Investment Disputes between States and Nationals of other States (adopted 18 March 1965, entered into force 14 October 1966) 575 UNTS (1965) 159 (The ICSID Convention).
[8] Lucy Reed, Jan Paulsson and Nigel Blackaby, Guide to ICSID Arbitration (Kluwer Law International 2010) 7-8; Jan Peter Sasse, An Economic Analysis of Bilateral Investment Treaties (Gabler Verlag 2011) 59.
[9] See Jeffery P. Commission, 'Precedent in Investment Treaty Arbitration – A Citation Analysis of a Developing Jurisprudence' (2007) 24(2) Journal of International Arbitration 129.
[10] Grand River Enterprises Six Nations, Ltd., et al. v. United States of America, NAFTA/UNCITRAL Arbitration Rules Arbitration (Award, 12 January 2011), <http://www.arbitrationlaw.com/files/free_pdfs/Grand%20River%20v%20US%20-%20Award.pdf> accessed 15 May 2012, para. 61; The Canadian Cattlemen for Fair Trade v. United States of America, NAFTA/UNCITRAL Arbitration Rules Arbitration (Award on Jurisdiction, 28 January 2008),

<http://italaw.com/documents/CCFT-USAAward_000.pdf> accessed 15 May 2012, paras 49-51.

[11] Engela C. Schlemmer, 'Investment, Investor, Nationality, and Shareholders', in Peter Muchlinski, Federico Ortino and Cristoph Schreuer (eds), The Oxford Handbook of International Investment Law (OUP 2008) 49, 62-3; Julian D. Mortenson, 'The Meaning of "Investment": ICSID's Travaux and the Domain of International Investment Law' (2010) 51 Harvard International Law Journal 257, 280ff.

[12] Azurix Corp. v. Argentine Republic, ICSID Case No. ARB/01/12 (Decision on Jurisdiction, 8 December 2003) 43 ILM (2004) 262; Enron Corporation and Ponderosa Assets, L.P. v. Argentine Republic, ICSID Case No. ARB/01/3 (Decision on Jurisdiction, 14 January 2004) <http://www.asil.org/ilib/Enron.pdf> accessed 15 May 2012; Fraport AG Frankfurt Airport Services Worldwide v. Republic of the Philippines, ICSID Case No. ARB/03/25 (Award, 16 August 2007) <http://ita.law.uvic.ca/documents/FraportAward.pdf> accessed 15 May 2012; Biwater Gauff (Tanzania) Limited v. United Republic of Tanzania, ICSID Case No. ARB/05/22 (Award, 24 July 2008) <http://icsid.worldbank.org/ICSID/FrontServlet?requestType=CasesRH&actionVal=showDoc&docId=DC1589_En&caseId=C67> accessed 15 May 2012. Generally, also in: Pantechniki S.A. Contractors & Engineers v. Republic of Albania, ICSID Case No. ARB/07/21 (Award, 30 July 2009) <http://icsid.worldbank.org/ICSID/FrontServlet?requestType=CasesRH&actionVal=showDoc&docId=DC1133_En&caseId=C113> accessed 15 May 2012.

[13] Fedax N.V. v. Republic of Venezuela, ICSID Case No. ARB/96/3 (Decision on Objections to Jurisdiction, 11 July 1997) 37 ILM (1998) 1378.

[14] Campbell McLachlan, Laurence Shore and Matthew Weiniger, International Investment Arbitration: Substantive Principles (OUP 2007) 169-170; Reed, Paulsson and Blackaby (n 8) 68.

[15] Fedax v. Venezuela (n 13) para. 43.

[16] Salini Costruttori S.p.A. and Italstrade S.p.A. v. Kingdom of Morocco, ICSID Case No. ARB/00/4 (Decision on Jurisdiction, 23 July 2001) 42 ILM (2003) 609.

[17] Ibid., para 52. The Salini v. Morocco dispute arose out of claims which two Italian companies, Salini and Italstrade, had against a Moroccan entity (ADM) financed by the Moroccan government regarding the construction of highways for which these companies where contracted by ADM. Moroccan government, inter alia, tried to negate the definition of contractual rights (which these companies had) as investments for ICSID purposes, which was eventually rejected by the tribunal. Salini criteria as described above were defined in the process.

[18] Československa obchodní banka, a.s. v. Slovak Republic, ICSID Case No. ARB/97/4 (Decision on Objections to Jurisdiction, 24 May 1999) 14 ICSID Review (1999) 251; MCI Power Group LC and New Turbine v. Ecuador, ICSID Case No. ARB/03/6(Award, 31 July 2007) <http://ita.law.uvic.ca/documents/MCIEcuador.pdf> accessed 15 May 2012; RSM Production Corporation v. Grenada, ICSID Case No ARB/05/14 (Award, 13 March 2009) <http://ita.law.uvic.ca/documents/RSMvGrenadaAward.pdf> accessed 15 May 2012.

[19] RSM v. Grenada (n 18) para. 241.

[20] Bayindir Insaat Turizm Ticaret Ve Sanayi A.S. v. Islamic Republic of Pakistan, ICSID Case No. ARB/03/29 (Decision on Jurisdiction, 14 November 2005)
<http://icsid.worldbank.org/ICSID/FrontServlet?requestType=CasesRH&actionVal=showDoc&docId=DC523_En&caseId=C27> accessed 15 May 2012; LESI, S.p.A. and

Astaldi, S.p.A. v. People's Democratic Republic of Algeria, ICSID Case No. ARB/05/3 (Decision on Jurisdiction, 12 July 2006) <http://icsid.worldbank.org/ICSID/FrontServlet?requestType=CasesRH&actionVal=showDoc&docId=DC528_Fr&caseId=C48> accessed 15 May 2012; Víctor Pey Casado and President Allende Foundation v. Republic of Chile, ICSID Case No. ARB/98/2 (Award, 8 May 2008) <http://ita.law.uvic.ca/documents/Peyaward.pdf> accessed 15 May 2012; Saba Fakes v. Republic of Turkey, ICSID Case No. ARB/07/20 (Award, 14 July 2009) <http://ita.law.uvic.ca/documents/Fakes_v_Turkey_Award.pdf> accessed 15 May 2012; For a non-ICSID case, see Romak S.A. v. The Republic of Uzbekistan, PCA Case No. AA280 (Award, 26 November 2009) <http://ita.law.uvic.ca/documents/ROMAK-UZBEKISTANAward26November2009.pdf> accessed 15 May 2012.

[21] Jan de Nul N.V. and Dredging International N.V. v. Arab Republic of Egypt, ICSID Case No. ARB/04/13 (Decision on Jurisdiction, 16 June 2006) <http://ita.law.uvic.ca/documents/JandeNuljurisdiction061606.pdf> accessed 15 May 2012.

[22] Joy Mining Machinery Limited v. Arab Republic of Egypt, ICSID Case No. ARB/03/11 (Award, 6 August 2004) 19 ICSID Review (2004) 486; Helnan International Hotels A/S v. Arab Republic of Egypt, ICSID Case No. ARB/05/19 (Decision on Objections to Jurisdiction, 17 October 2006) <http://icsid.worldbank.org/ICSID/FrontServlet?requestType=CasesRH&actionVal=showDoc&docId=DC773_En&caseId=C64> accessed 15 May 2012.

[23] Phoenix Action Ltd v. Czech Republic, ICSID Case No. ARB/06/5 (Award, 15 April 2009) <http://icsid.worldbank.org/ICSID/FrontServlet?requestType=CasesRH&actionVal=showDoc&docId=DC1033_En&caseId=C74> accessed 15 May 2012.

[24] See Saba Fakes v. Turkey (n 20) para 96.

[25] Andrea K. Bjorklund, 'Investment Treaty Arbitral Decisions as Jurisprudence Constante', UC Davis Legal Studies Research Paper No. 158 <http://ssrn.com/abstract=1319834> accessed 15 May 2012, 266-267; Cristoph Schreuer, 'Diversity and Harmonization of Treaty Interpretation in Investment Arbitration', (2006) 3 Transnational Dispute Management 10, 14; Jan Paulsson, 'International Arbitration and the Generation of Legal Norms: Treaty Arbitration and International Law', (Provisional Issue September 2006) Transnational Dispute Management 13; Thomas W. Wälde, 'The Present State of Research Carried Out By the English-Speaking Section of the Centre for Studies and Research', in New Aspects of International Investment Law (Brill 2006) 140ff; Alexis C. Brown, 'Presumption Meets Reality', (2001) 16 American University International Law Review 1013; Tai-Heng Cheng, 'Precedent and Control in Investment Treaty Arbitration', (2007) 30 Fordham International Law Journal 1048-1050.

[26] Mortenson (n 11) 315.

[27] See Zachary Douglas, The International Law of Investment Claims (CUP 2009) 163-164.

[28] Portfolio investments comprise investments constituted out of acquiring and ownership of shares in a foreign company without the aim of managing that company. See Sornarajah (n 4) 8-10.

[29] See, for example, Devanish Krishan, 'A Notion of ICSID Investment', in Todd Weiler (ed), Investment Treaty Arbitration and International Law (Vol. 1, JurisNet LLC 2008) 61.

[30] Saba Fakes v. Turkey (n 20) para 112.

[31] Amazu A. Asouzu, International Commercial Arbitration and African States – Practice, Participation and Institutional Development (CUP 2004) 263.
[32] Kenneth J. Vandevelde, 'The Economics of Bilateral Investment Treaties', 41 Harvard International Law Journal (2000) 469, 492.
[33] Ignacio D'Alessio, 'A Comment on ICSID's jurisdictional issues: Problems, Solutions and Recommendations for a Better Understanding in International Commercial Relations' (Research Paper, Vrije Universiteit Amsterdam 2008) <http://ssrn.com/abstract=1153988> accessed 15 May 2012, 35.
[34] Katia Yannaca-Small, 'Definition of "Investment": An Open-ended Search for a Balanced Approach', in Katia Yannaca-Small (ed), Arbitration Under International Investment Agreements: A Guide to the Key Issues (OUP 2010) 243, 259-261.
[35] Marek Jezewski, 'There is No Freedom without Solidarity: Towards a New Definition of Investment in International Economic Law' Working Paper No. 51/08, SIEL Inaugural Conference Online Proceedings <http://www.ssrn.com/link/SIEL-Inaugural-Conference.html> accessed 15 May 2012, 1, 12ff.
[36] LESI v. Algeria (n 20) para. 72; Bayindir v. Pakistan (n 20) para. 137.
[37] Given (n 5) 487.
[38] Malaysian Historical Salvors, SDN, BHD v. Malaysia, ICSID Case No. ARB/05/10 (Decision on the Application for Annulment, 16 April 2009) <http://icsid.worldbank.org/ICSID/FrontServlet?requestType=CasesRH&actionVal=showDoc&docId=DC1030_En&caseId=C247> accessed 15 May 2012.
[39] Patrick Mitchell v. Democratic Republic of the Congo, ICSID Case No. ARB/99/7 (Decision on the Application for the Annulment of the Award, 1 November 2006) <http://ita.law.uvic.ca/documents/mitchellannulment.pdf> accessed 15 May 2012.
[40] For more on this issue see Yannaca-Small (n 34) 259-61.
[41] Malaysian Historical Salvors, SDN, BHD v. Malaysia, ICSID Case No. ARB/05/10 (Dissenting Opinion of Judge Mohamed Shahabuddeen, 19 February 2009) <http://icsid.worldbank.org/ICSID/FrontServlet?requestType=CasesRH&actionVal=showDoc&docId=DC1031_En&caseId=C247> accessed 15 May 2012.
[42] Schreuer, The ICSID Convention: A Commentary (n 5) 140.
[43] Mortenson (n 11) 297-298.
[44] Global Trading Resource Corp. and Globex International, Inc. v. Ukraine, ICSID Case No. ARB/09/11 (Award, 1 December 2010) <http://icsid.worldbank.org/ICSID/FrontServlet?requestType=CasesRH&actionVal=showDoc&docId=DC1771_En&caseId=C660> accessed 15 May 2012, para. 44.
[45] North American Free Trade Agreement (The NAFTA) (signed on 17 December 1992, entered into force 1 January 1994) 32 ILM (1993) 289 (Parts I-III) 32 ILM (1993) 612 (Parts III-VI), Art. 1139; Joy Mining v. Arab Republic of Egypt (n 22) paras. 57-58; OECD, Definition of Investment and Investor in International Investment Agreements (2008) 50.
[46] Paul Szasz, 'A Practical Guide to the Convention on Settlement of Investment Disputes', (1968) 1 Cornell International Law Journal 1, 15; William Rand, Robert N. Hornick and Paul Friedland, 'ICSID's Emerging Jurisprudence: The Scope of ICSID's Jurisdiction', (1986) 19 New York University Journal of International Law and Politics 33, 36; UNCTAD, Scope and Definition (1999) 20; Schreuer, The ICSID Convention: A Commentary (n 5) 139; Raymond Doak Bishop, James R. Crawford and William M. Reisman, Foreign Investment Disputes: Cases, Materials and Commentary (Kluwer International 2005) 344; Rudolph Dolzer and Cristoph Schreuer, Principles of International Investment Law (OUP 2008) 60.

[47] Global Trading and Globex v. Ukraine (n 44) paras. 54-56.
[48] Ibid, para 56 (footnote omitted and emphasis added).
[49] Petrobart Ltd v Kyrgyz Republic, SCC Arbitration Institute Arbitration No. 126/2003 (Award, 29 March 2005) <http://www.asil.org/pdfs/Petrobart-Kyrgyz.pdf> accessed 15 May 2012.
[50] McLachlan, Shore and Weiniger (n 14) 170.
[51] Schreuer, The ICSID Convention: A Commentary (n 5) 139.
[52] Rudolph Dolzer, 'The Notion of Investment in Recent Practice', in Steve Charnovitz et al (eds), Law in the Service of Human Dignity: Essays in Honour of Florentino Feliciano (CUP 2005) 261, 263.
[53] Szasz (n 46) 15.
[54] Douglas (n 27) 171-2; Jeswald W. Salacuse, 'Towards a Global Treaty on Foreign Investment: The Search for a Grand Bargain', in Norbert Horn and Stefan M. Kröll (eds), Arbitrating Foreign Investment Disputes: Procedural and Substantive Legal Aspects (Kluwer Law International 2004) 51, 68-9.
[55] An excellent example of 'pure' sale of pharmaceuticals which failed to satisfy these elements (despite long duration and recurrent transactions) and was thus not found to be an investment can be found in The Republic of Italy v. The Republic of Cuba, Italy-Cuba BIT Ad Hoc Arbitration (Final Award,15 June 2008), < http://italaw.com/documents/Italy_v_Cuba_FinalAward2008.pdf> accessed 15 May 2012, paras 212-21.
[56] UNCTAD, Investor-State Dispute Settlement and Impact on Investment Rulemaking (2007) 21.
[57] Schreuer, The ICSID Convention: A Commentary (n 5) 141.
[58] CSOB v Slovakia (n 18) para 72.
[59] Franz Sedelmayer v Russian Federation, Germany-Russia BIT Ad Hoc Arbitration (Award, 7 July 1998) <http://ita.law.uvic.ca/documents/sedelmayer-russia.pdf> accessed 15 May 2012.
[60] McLachlan, Shore and Weiniger (n 14) 168.

BOOK REVIEW:

'SOVEREIGN EQUALITY AND MORAL DISAGREEMENT – PREMISES OF A PLURALIST INTERNATIONAL LEGAL ORDER' BY BRAD R. ROTH

(OXFORD UNIVERSITY PRESS, 2011)

J. Alexis Galán Ávila*

Despite the many developments that we are witnessing, very much in front of our eyes, in respect of the changing structures that have, for the last centuries, underpinned political and social structures and which are reflected in how international law is evolving, it is easy to forget that sovereignty (at least in its normative sense) still plays a very important role in the conducting of international affairs. The current European financial crisis offers us a clear-eyed instantiation of the state of affairs in which we are situated. What we witness is that the tension and paradoxes inhabiting world affairs cannot be wished away. In this sense, Brad R. Roth's new book is timely and necessary. It is a refreshing book insofar as it purports to provide a defense of sovereignty as the basic institution for the continuing existence and maintenance of a plural international legal order. Whereas it is common to read of how sovereignty is diminished and faltering, Roth's approach provides us with a veritable account on the continuing significance of sovereignty. The purpose is to warn those that advocate for a more forceful intervention from international law, even if they have good intentions, like human rights movements, of the dangers that throwing sovereign equality would entail. In this sense, my sympathies lie with him in respect of the core of his argument, which is that 'the continued significance of state sovereignty, not exclusively as newly harmonized with supranational authority in the form of "sovereignty as responsibility," but also in continued tension with supranational authority in the form of a presumptive *domaine réservé*.'[1] In other words, state sovereignty represents a common agreement among states whereby it provides them with the right to pursue their own good life in a situation in which there is ample moral and political disagreements concerning the 'right' good life. To do otherwise would mean to jeopardize the basic element of the international legal order and to give *carte blanche* to powerful states to impose their will upon weak states. There is a lot on

* PhD Researcher, Department of Law, European University Institute (Florence).

[1] Brad R. Roth, *Sovereign Equality and Moral Disagreement – Premises of a Pluralist International Legal Order* (Oxford University Press, 2011) 6.

which to comment in respect of this provocative book but I will focus on two issues with which he does not deal successfully or rather neglects to take into account in his defense of sovereignty. On the one hand, he wants to reconcile state sovereignty's prerogative in deciding what sort of good life the community wants to pursue, even if gross violations of human rights are involved. On the other hand, in focusing only on power as being coercive, Roth ends up ignoring other facets of power in which sovereignty as 'self-determination' is equally affected. These criticisms should not be seen as a clear indictment on the book rather than minor comments on an otherwise impressive piece of scholarship.

Before explaining my criticisms I will try to summarize Roth's argument. I hope I can do it justice as it is a subtle and intricated one and there is always the risk of omitting relevant parts in a short review. In a nutshell, he argues for sovereignty as the ability and right to pursue one's own conception of the good life by a political community, in contrast with discourses that simply pits sovereignty against international law or sovereignty as responsibility, represents the basic and constitutive institution of the international legal order. Sovereignty, then, in the discourse propounded by him, simply reflects the stubborn fact that the international community is so diverse and plural that there is no common understanding of what the good life is – that is, the notion of diverse political moralities. If this is not respected or its importance is desecrated, the potential result would be that the project of the international legal order would be imperiled because the basic agreement for respecting it would no longer exist. Likewise, powerful states would not have any normative restriction in simply imposing their will or conceptions upon weaker states by adducing that they are following the purposes of international law. Thus, (normative) sovereignty functions as a barrier that attempts to ensure the protection of political communities in developing their particular understandings of what the good life is (especially weak states)[2] In his own words, 'the sovereign equality of states [exists] as an institutional response to persistent disagreement about what constitutes a legitimate and just territorial public order.'[3] Because there is pervasive disagreement concerning what justice is, it is better to accept the right of those sovereign to dictate their own way of life, even if there is violence involved. The lack of knowledge from external actors about the specific conditions in a territory or about the values that a certain community upholds, and a possible imposition of alien understandings of public order from powerful states, dictates that the moral thing to do is to let the communities to solve their conflicts on their own. To sum up,

[2] Ibid., 127.
[3] Ibid., 273.

where foreign exertions jeopardize a political community's capacity to defend its interests or to arrive at decisions in keeping with the distinctive values of its members, the imperative to maintain that capacity ... justifies coming to the defense of that community's sovereign prerogatives, even where those prerogatives have been exercised unjustly.[4]

This argument is repeated throughout the entire book. Without entering into much detail, he bases this reading of sovereignty and the international legal order in a non-culture based pluralism, that is to say, in an internal liberal reading of communities as comprising individuals that decide how to constitute themselves, rejecting any 'essentialism."[5] And to further his case he attempts to show that this reading of sovereignty is still pervasive when analyzing the practice and discourses surrounding the issue of 'Responsibility to Protect,' including issues of internal order like secessions, coups and effective control and international criminal law.

The first thing that could spring to someone reading this book is that by emphasizing sovereignty as self-determination he is simply reverting to old notions in which the state was seen as a black box whereby it was of no interest for international law to discuss what is happening. In the book he assures us that this is not the case. He insists throughout the whole book that the acceptance of his sovereignty discourse cannot be equated to the condoning and acceptance of human rights violations. His reading of sovereignty does not intend 'to obscure the interests of human being grievously harmed by the abuse of sovereign prerogative.'[6] He merely wants to recalibrate the current discourses by pinpointing the morality that underpins sovereignty and to highlight that to disregard the pervasiveness of moral disagreement would have pernicious consequences.[7] Furthermore, the acceptance of the fact of distinctive and divergent ideas of the good life does not also imply the embracing of moral relativism. For him, 'a commitment to pluralism in no way implies agnosticism about the wrongfulness of the other's conduct, nor need it assert that the conduct, albeit wrong "for us," may be right "for them." Rather, pluralism ... accords... certain basic prerogatives and inviolabilities that withstand the other's wrongful, but non-aberrant, conduct.'[8] Despite Roth's assurances, it is neither clear nor evident from him how to restore the balance and when. My impression when reading the book is that in any confrontation

[4] Ibid., 122.
[5] Ibid., chapter 4.
[6] Ibid., 6.
[7] Ibid.
[8] Ibid., 125.

between possible 'standards' from external actors and 'internal' standards, the latter always have preference. The limit of pluralism lies, according to him, in the need to ensure that the international legal order respects human dignity and that a 'violation of the physical integrity of the person is, it would seem, incompatible with such respect.'[9] Now, from the reading of the book it would be hard to determine when that would occur. If we need to distrust universal applications of public order because of the level of generality of those norms, when do we know that human dignity is being violated? This determination would imply a reading of international legal order in a particular way. It seems that he has on mind an extreme case like the Rwandan genocide.[10] But if we followed him, in my opinion, we would not at that moment necessarily be able to determine what was happening because there might perhaps have been two different ideas of public order; at that moment, we did not have the relevant facts. I simply cannot see how he can reconcile the respect for human dignity when he clearly acknowledges that his sovereignty discourse will protect the perpetrators of human rights violations.[11] He is at pains in asserting that we can respect a community's self-determination with an evaluation of the violations.[12] Nevertheless, it is never clear or obvious when that would happen and he does not provide us with any guidance beyond a mere comment that the system surely needs some reforms.[13] Furthermore, any interpretation of human rights necessitates an understanding of what kind of public order is assumed. If we do not want to fall into the trap of determining a particular political morality that would impinge on a particular community, we need to accept at minimum a weak relativism that he seems to discard altogether. He expressly states that his unified account of sovereignty simply balances the external obligations with the internal standards. In any case, we can judge and evaluate if violations have been committed at the same time as we respect pluralism. This line of reasoning would go as follows: if in a country there is a civil war and violations have been committed we should respect that for several reasons – disagreements concerning the good life, we do not have sufficient knowledge of what is really happening and so forth. However, if there is lack of respect paid to human dignity we can assess it, or condemn it. But how? And with which consequences? The intervention of the ICC or the

[9] Ibid., 117.
[10] He states that to sustain '"Hutu Power" (the ideology of the Rwandan *genocidaires)* as an admissible answer to the question of public order would be a grotesque parody of the pro-sovereignty argument advanced above.' This suggests that there are somehow certain, however mimimal, limits for the pursuing of a public order, ibid., 117.
[11] Ibid., 6.
[12] See *inter alia* ibid., 41 or 67.
[13] Ibid., 129.

Security Council would simply undermine what Roth wants to uphold. I would go even further, if the SC would act by unanimity that could mean that there is simply a coincidence of interest but not that there is agreement on how good life should be managed. Even more worrisome, under which standards? If human rights norms are universal and the details are elusive then no possible condemnation can be done except by our own morality.[14] That is to say, we definitely condemn the action but because our own interpretation suggests this to us. Hence, we end up in a sort of weak relativism. We do not have to say that it is good for them what is not good for us but we undertake this assessment under our own conceptions. If we do not want to end up in this sort of relativism we need to have a sufficiently agreed public order and if that is not the case, according to Roth, then we have to either accept that sovereignty, which, as he puts it, should prevail, or we impose a certain notion of public order that involves the violation of human dignity. Human dignity and how to protect it inherently invokes a particular understanding of the good life. Thus, in my opinion, he ends up on the side of sovereign 'immunity' and human rights violations have to be in the majority of the circumstances an afterthought. There is nothing wrong with this position. In some occasions, the danger of intervening will be greater than not doing anything but Roth, contrary to what he argues,[15] has to accept that violations will be the norm, besides the weak relativism associated with his position and which it brings to the table.

There is equally a second problem with Roth's focus on overt uses of 'power' that ends up obviating other forces that undermine sovereign as reflection of self-determination. By focusing on 'extreme cases' of violence, intervention, gross human rights violations he ignores how power, by seemingly leaving sovereignty intact, can affect the poor and weak that he wants to defend. His account ignores how power does not necessarily lie in the classical Dahlian reading of A making B do something that he otherwise would not do.[16] Power is also about what has not happened. As Bachrach and Baratz noticed long time ago, sometimes there are certain issues that are never discussed or approved because the institutional structures are skewed towards the powerful, e.g., rules, voting, actors' position, and so forth.[17] What I mean is that the current design of human rights, for instance, as I just stated, already implied a particular conception of what the good life is. This would imply that by subscribing to the

[14] See his skepticisms towards universal norms at ibid., 95ff.
[15] See Ibid., 117.
[16] See R.A. Dahl, 'The Concept of Power,' 2:3 *Behavioral Science* 201 (1957).
[17] P. Bachrach and M.S. Baratz, 'Two Faces of Power,' 56:4 *The American Political Science Review* 947 (1962).

different human rights conventions, the communities have to construct a specific morality that would run counter to Roth's normative commitments. True, he would counter that they are reflections of consensus and therefore they represent a minimum public order that has been accepted. But then the supposed pluralism of Roth's would be totally undermined because states would end up with the obligation to construct something akin to a liberal-democratic state. This accusation could be a little unfair to Roth, after all, he is aware of how sometimes sovereignty has been a mere façade and that interventions, abuses, coercions inhabit more than it should the world of international affairs. Hence, he is not naïve in this regard.[18] But still he is surely conscious of what human rights bring with themselves. This is not by any metric a very novel claim. Unless for him human dignity is something else, I have to assume that he refers to current human rights treaties like the universal declaration of 1948. In my view, then, he has to simply acknowledge that there is less pluralism than there should be or the 'disempowering' role that human rights can have on communities wanting to have their own good life.

Moreover, he also seems to be unaware of a different facet of power which has been introduced by Foucault: the so-called 'productive power' which refers to the ability to impose certain discourses and conceptions through the production of knowledge.[19] Discourse brings with it certain prejudices, notions and readings of social life. Thus, it shapes our thoughts when we want to do something. This would imply that even if sovereignty seems to be respected, the actual 'ability' of those communities in self-determination would be rather limited. Hence, human rights could be considered one sort of discourse, and the same with bodies like the WTO or the IMF.[20] By producing certain knowledge they reduce the self-determination of the distinct political communities. In the end, sovereign equality, understood as the determination of how they should lead a good life, becomes a hollowed out concept. In this regards, sovereignty as a normative concept would unfortunately only be an empty shell.

Regardless of these inadequacies, Roth's book is a must read. We tend to forget in this age of 'constitutional language' that we are riffed with disagreements concerning our governance structures. This book is a timely

[18] See Roth (n 1) 53ff.
[19] See, for instance, M. Foucault, *The Order of Things: An Archeology of the Human Sciences* (Pantheon, 1971)
[20] For a good analysis of the role of knowledge in 'producing' power see A. Lang, 'Legal Regimes and Regimes of Knowledge: Governing Global Services Trade,' *LSE Law, Society and Economy Working Paper Series,* WPS 15-2009 (at http://www.lse.ac.uk/collections/law/wps/WPS2009-15_Lang.pdf) (Visited last time 22nd May 2012).

reminder of the dangers associated with benign discourses that can wind up unraveling something more essential for the maintenance of peace and security.

HOLDING YOUR OWN TO ACCOUNT: EU POLICY CONCERNING ITS MNES ABROAD.

BOOK REVIEW: 'MULTINATIONAL ENTERPRISES AND HUMAN RIGHTS: OBLIGATIONS UNDER EU AND INTERNATIONAL LAW' BY ALEXANDRA GATTO

(EDWARD ELGAR, CHELTENHAM, 2011, ISBN: 978 1 84844 034 0, £95)

Julien Topal

Alexandra Gatto's thesis wants to answer one question: how can the European Union (EU) "ensure that EU-based Multinational Enterprises (MNEs) respect human rights when operating in third countries?" (p.vii) This question is as salient as ever today, with increasing clarity that economic development needs to be embedded in broader societal values. Human rights breaches, complicity in government violence but also the detrimental effect of supply chains on local labor standards have put a magnifying glass to both negative and positive obligations for MNEs limiting those negative impacts. However, international law has been grappling with a way to conceptualize such obligations within its own confines and is confronted by a somewhat disabling doctrinal tradition, 'weak' governance in host countries and less than willing policy-makers.
It is against this background that Gatto has produced a broad scoped account of the developments in international law, through multi-stakeholder initiatives, and within the European Union (EU) concerning corporate human rights obligations. After arguing for a conception of 'limited corporate obligations under international law and a subsequent expansive set of corporate human rights obligations, the core of Gatto's argument concerns the ECs engagement in the field to date. As a self-referred 'normative' power, the EU should be a leader in the global human rights movement. Through in-depth analysis of EU Treaty Law, Regulations, and Directives Gatto however provides for convincing criticism of the extent to which the EC has utilized its competences. In practice thus there is much talk but only small practical benefits. Commendably this work thus is more than (yet) another argument on the possibilities of qualifying MNEs as holders of human rights obligation under international law. Her interest spans wider to include the legal opportunities of indirect measures under EU law of internal and external policy to ensure the human rights obligations of MNEs.

The book is made up of 4 parts, the last being the conclusion. Part I

provides general outline of the topic, containing chapters that provide a theoretical framework, MNEs as addressees of international law, and MNEs and human rights law. Part II and III represent the core of the book, the EUs law and policy concerning the human rights obligations of MNEs. These parts respectively take up the MNE-human rights relations within the EU as well as the way in which the relation plays out in the EUs external policy. In my opinion the innovative work is done in Part II and III, therefore I will focus on these and only shortly comment on Part I.

Gatto opens her book with a long (45 pages) chapter outlining her theoretical framework and some of the legal and conceptual issues such a framework has to capture. The chapter suffers somewhat under the amount of issues introduced In summary, however, Gatto gives good reasons that the complex nature of the issue of human rights obligations of MNEs multilevel and complimentary approaches. Her framework therefore includes non-binding measures, binding legal instruments, and complementary in applying direct as well as indirect (through EC internal market or commercial policy, or external policy geared at host states) measures. Lastly, MNEs should be ascribed human rights obligations according to the theory of indivisibility of human rights. This means that MNEs hold obligations or should be regulated in such a way that they contribute to all types of human rights. Following UN Special Representative John Ruggie's, Gatto specifies human rights into duties to respect, promote, protect and support. This is not the clearest part of the chapter, however I do not think that too much weighs on it in the end, due to Gatto's further developments in the book.

The subsequent 2 chapters develop the general international approaches to MNE human rights obligations. Chapter 2 touches on the 'obligatory' topic of the evolution of the concept of 'subject' and legal personality under international law since the Second World War. Topically, Gatto sketches the congruence between the progressive increase of corporate rights and the 'conservatism' concerning their obligations. Chapter 3 extends on the analysis by introducing two ways in which international human rights law has developed. These are discussed with an eye on applying these as 'models' for developing an account of corporate obligations. Gatto accounts for the increased focus on horizontal application of human rights law, i.e. the state responsibility to ensure human rights within its territory. Cases concerning investment projects that impede on indigenous people's tribal lands are one such recent novelty. The second development consists of the emergence of the individual on the international legal scene. Especially under criminal law and humanitarian law, individuals have come within the scope of international law. Neither of these 'models' are directly applicable to the

MNE however since it cannot be easily conceptualized as either a state-like entity or an individual. Up to date therefore, on the one hand, somewhat creative solutions stretching either the notion of state obligations or international criminal responsibility have been used to capture corporate human rights obligations. On the other hand, 'soft' law approaches have been used to determine more precisely how MNEs could be taken as holders of obligations under international law, culminating in, at least in Gatto's view, the UN Norms. In these two developments, Gatto sees a 'limited personality' for MNEs appearing – although this is not explicitly specified.

The core the book concerns a critical analysis of indirect corporate human rights obligations through efforts by the EC. Part II delves into the ways the EC has applied internal policy measures to pursue this goal. Part III accounts for the ECs external policy measure to strengthen host states capacity to ensure respect for human rights by MNEs.
The salience of this topic is clear: the EU has positioned itself as a 'normative' power that seeks to create a value-based system of global governance. Its role in embedding MNEs into societal values such as human rights can be seen as a test-case of this self-proclaimed power. Gatto concludes that the EC has not lived up to the hype. As she argues in light of internal measures, "there do not seem to be many obstacles to directly imposing human rights obligations upon European MNEs and applying them extraterritorially to European companies [...]." (p. 132) The fact that MNEs are recognized as subjects of EU law, that they are conceptualized as economic units (instead of the more diffuse concept of a legal unit), and that the EU has applied competition law extraterritorially already supports this claim. Chapter 5 neatly shows that the EC has not fully explored the opportunities that are legally within its reach by way of analysis of competences such as common commercial policy and company law, social policy and public procurement. Take the inclusion of social concerns in public procurement. The EC could make respect for human rights through the supply chain a condition of contracting. Such an approach flies in the face of the accepted doctrine of economic advantageousness as the sole basis for assigning contracts and might raise worries of discriminatory policy and protectionism. But while ECJ rulings allow for non-discriminatory social concern-inclusion and the leeway provided under the WTOs Plurilateral Government Procurement Agreement (GPA) the EC has held on to a restrictive interpretation of their policy space.

One important caveat is in place. The powers of the EC are 'conferred' powers; severely limiting the legal basis for ascribing the EC general

powers over human rights.[1] In other words, the EC itself does not have the competence to legislate directly on human rights. But Gatto convincingly shows however that there is enough of a legal basis within the ECs competences to ensure compliance with human rights within and by its own institutions.

The third part of Gatto's argument brings us to the legal basis and application of EC competences in its external relations. Here too, the EC has ample space to draw on implied powers. More so, however, than in the case of the European Common Market-policies, an argument of coherence can be made that policy should ensure MNE human rights compliance. As Gatto deduces from European Treaty law in light of Article 6 of the Treaty on European Union (TEU), all EC external policy should contribute to the respect for human rights (p. 200-201).
In this sphere the EC has two main approaches at hand to ensure respect for human rights by MNEs. The first measure is the use of a human rights clause in external agreements. These clauses introduce conditionality requiring a host state to commitment ('respect' for) to human rights within its territory. The EC also applies non-regulatory instruments, from incentivizing human rights policies in third countries to strengthening civil society. The question is however to what extent these two types of instruments effectively apply to the human rights obligations of MNEs. In chapter 8 Gatto discusses initiatives under these instruments and notes that notwithstanding the recognition of the importance of MNEs respecting human rights, none of these initiatives explicitly address MNEs. The concept of indirect approaches to corporate human rights obligations turns out somewhat empty – the fact that an improved host state's human rights record most probably also implies that corporate human rights breaches will be minimized is neither here nor there.
As an alternative route, the Common Commercial Policy (CCP) allows the EC to introduce unilateral trade measure to further human rights. Through the Generalized System of Preferences (GPS), an enabling clause that provides an exemptions of the Most Favorite Nation-clause of the GATT, the EC has offered preferential trade arrangements and capacity building to incentivize developing countries to ensure human rights. Tariff reductions and tax exemptions are offered to developing country in exchange to ratifications of human rights and labor standard treaties under this policy. The GSP is a potent incentive mechanism even though it

[1] Opinion 2/94 on the possible accession of the EC to the European Court of Human Rights (ECHR) provides the background for this caveat. Accession would imply the "entry of the Community into a distinct international institutional system, as well as integration of all the ECHR provisions into the Community legal order." (p. 115) Such constitutional change can only be brought about by a Treaty amendment.

suffers from some implementation and monitoring. Crucially however, Gatto notes, also in the case of the GSP the EC has not centered its attention on MNEs as such. Yet again thus, improvements of MNEs human rights records are expected to automatically follow those of the host country in which they operate.

Gatto's in-depth account of the potential and the shortcomings of EC policies, show some crucial weaknesses in the current role of the EC in furthering human rights globally. The oddity with this analysis is that the initiatives discussed, in Gatto's own words, do not (or barely) pay attention to the inclusion of human rights obligations of MNEs. The main actor, the MNE, is largely missing in action in the core parts of the argument. It is not surprising then that in her conclusions Gatto's recommendations aim at an improved focus on the issue of corporate human rights obligations. The problem within the EU does not concern a lack of legal possibilities, as Article 6 of the TEU confirms, but a lack of political will. Such unwilling attitude is exemplified by the limited use of competences and in the Corporate Social Responsibility (CSR)-program of the EC. To be complemented on its efforts in getting multiple stakeholders involved, the EC went off track in picking favorites with its Business Alliance for CSR-initiative and opting for a merely voluntary market based CSR-model. Gatto urges the EC therefore to develop its laws in sync with the evolving consensus in international law that human rights obligations of MNEs have to be secured. Secondly, she recommends the EC to utilize the legal competence it possesses to extend its policing powers to promote MNE obligations in third countries.

To conclude I want to make three critical observations. The first concerns the up-to-date-ness of the argument. Gatto book is based on her PhD that was submitted in 2007. There is thus a 4-year gap between the thesis and the current publication, in which a dynamic field as the one under scrutiny are bound to have taken place. The fact is however that very little to almost no updating has been done for the current publication. Did the implementation of the Lisbon Treaty change any relevant aspects of EC policy? And the Cotonou Agreement has seen revisions in 2005 and 2010 – Gatto's references end in the year 2000. Gatto urges the European Bank for Reconstruction and Development (EBRD) and the European Investment Bank (EIB) (p.158) to rigorously include social and envirnmental values in their project assessments. The EIB for instance, under its 2009 'EIB Statement of Environmental and Social Principles and Standards' follows 3 Environmental and Social Principles to assess projects for financing. [2] Similarly, the secondary literature latest reference is a

[2] See http://www.eib.org/attachments/strategies/eib_statement_esps_en.pdf.

forthcoming article (published in 2007) and there are only 13 references to sources dating after 2005. This does not disqualify Gatto's account but it does leave the reader wonder about what has come ever since she wrote up her account.

Second, a question left unaddressed concerns the reasons that the EC could possibly have to limit its actions towards specifying human rights obligations of MNEs. Besides potential conflict between the EC and member-states on competences and overarching interference in third countries, little probing into such considerations is offered. The EC might have good reasons to further going regulations – directly and indirectly. The availability of legal opportunities does not make for good or feasible regulations (yet). Gatto's argument therefore remains nothing more but also nothing less than a formal-legal argument that convincingly shows that there are few legal obstacles to the EC competences to further MNEs human rights obligations.

Of course, the disciplinary rationale of law expectedly seeks out legal argument. However, thirdly, the book could have benefited from a more specific analysis of how a more focused engagement with MNEs can play a constitutive role in the development of a state and the ensuring of human rights. Such an argument would look into the very specific issues pertaining to the corporate presence within a country. This comment connects to the conceptual foundations presented in part I of the thesis. Gatto holds strong to the idea of indivisibility of human rights and a gradational approach in ascribing them to companies, while at the same time she develops a differentiated register of types of obligations (from respect to promote) that should be promoted through both voluntary and legal means.[3]

Although commendable, such an expansive and rigorous account of human rights might not be most practicable in improving the ECs internal and external relations to promote human rights obligations of MNEs – nor does Gatto's core argument seemingly need it. In the conclusion of part I, Gatto builds this expansive set of human rights obligations on a notion of 'limited corporate personality' under international law. Potentially overreaching on this limited fundament, she ascribes MNEs both negative

[3] On p.188 Gatto seemingly lashes out against voluntary approaches that lack convincing monitoring and enforcement measures since they are necessarily harmless to companies and of no help to workers and communities. For someone who contends to support a 'mixed-bag' approach to MNE obligations it is a surprising critique that does not sit well either with the important transitional and indirect value in civil liability procedures these voluntary means can have according to the author herself (p.99 and 188 for instance).

and positive duties to respect, protect, fulfill, support, and promote human rights albeit according to a 'sliding scale' (p. 96; which I reckon is equal to the 'principle of graduation' (p.vii)), which makes the 'scope and intensity' of the duty dependent on the right at hand, the capacity to impact, the exercise of governmental authority by the MNE and the presence or absence of fault. Throughout the book Gatto apparently loosens up on this rigid framework when she specifically addresses harm/violation or development/poverty abatement potentials of MNEs. Her 'principle of gradation' for instance is made dependent on the 'distance to the victim,' the latter notion normally being associated with negative duties of harm.

Despite the commendable doctrine of the 'indivisibility' of human rights, the two broad categories of duties simply translate into very different policy measures and legal adjustments. A more differentiated approach will be more successful in practice since it allows for a better alignment of initiatives and specific human rights goals, and seeking out complementary approaches to make means connect to the desired end.