ARTICLES

DRAWING A LINE IN THE PATENT SUBJECT-MATTER SANDS: DOES EUROPE PROVIDE A SOLUTION TO THE SOFTWARE AND BUSINESS METHOD PATENT PROBLEM?

Susan J. Marsnik & Robert E. Thomas

Abstract: In June 2010, the Supreme Court issued its decision in Bilski v. Kappos, a case that had the potential to rewrite the landscape for determining what types of computer-related and business method inventions would receive patent protection. Just six weeks earlier, the European Patent Office’s Enlarged Board of Appeal delivered a decision on the same subject matter that had the potential to produce similar change in Europe. Yet, given these two opportunities to overhaul imperfect patent systems, neither decision provided more than incremental change. This Article explains why neither jurisdiction is able or willing to produce comprehensive reform in this area, and seeks to illuminate the nature of patent reform that is possible on the two continents.


Thomas K. Cheng

Abstract: This Article undertakes a comparative study of corporate veil piercing doctrines under U.S. corporation and English company law. The Article highlights some fundamental differences between the doctrines in terms of jurisprudential approaches, treatment of specific case types, and other related issues. The Article demonstrates that despite these substantial differences, many English corporate veil cases in fact share a similar
analytical approach to the instrumentality doctrine under U.S. law. Therefore, it is possible to construct an English instrumentality doctrine that will bring structure and clarity to the English corporate veil doctrine. The Article concludes with a revival of the much-maligned single economic unit theory, first propounded by Lord Denning. This reformulated theory will provide a more systematic approach to veil piercing cases involving corporate groups.

NOTES

MILITARY COMMISSIONS, CRIMINAL COURT, AND THE CHRISTMAS DAY Bomber

Ian Kennedy

[pages 413–438]

Abstract: In 2009, Umar Farouk Abdulmutallab attempted to detonate an explosive device on a plane landing in Detroit on Christmas Day. The attack was unsuccessful, but it spurred an important domestic debate regarding U.S. anti-terrorist programs and policies. In particular, the event fueled an argument over the proper forum for the interrogation and prosecution of terrorist suspects captured in the United States. Focusing on national security issues, some contended that treating Abdulmutallab as a criminal defendant in an Article III court, rather than subjecting him to a military commission, was imprudent and dangerous, while others insisted that it was entirely appropriate and responsible. This Note will probe this debate by comparing the two tribunals as each relates to the legal protections for suspects during interrogation. The Note argues that although some differences do exist, it is quite plausible that treating Abdulmutallab and other captured terrorist suspects as criminal defendants in Article III court does not adversely impact intelligence gathering and national security.

GERMAN HOMESCHOOLERS AS “PARTICULAR SOCIAL GROUP”: EVALUATION UNDER CURRENT U.S. ASYLUM JURISPRUDENCE

Miki Kawashima Matrician

[pages 439–462]

Abstract: Thirty years after the enactment of the Refugee Act of 1980, the Board of Immigration Appeals and U.S. courts and have not reached consensus on a uniform definition for the protected category of “particular social group.” The lack of consensus has created much confusion and
inconsistent results for applicants seeking asylum in the United States. This Note examines one family’s grant of asylum as a vehicle for analyzing the two main approaches to “particular social group” and argues that the current treatment of the two standards as mutually exclusive by the BIA and the federal courts is inconsistent with the U.N. Guidelines. The Note concludes that U.S. jurisprudence on “particular social group” should mirror the approach of the U.N. Guidelines, which envisions broader protection under that category.

COMMENTS

DEFINING INVESTOR CONFIDENCE: AVOIDING INTERPRETIVE UNCERTAINTY IN Chevron Corp. v. Ecuador

Jason Burke

[pages 463–476]

Abstract: In an increasingly globalized world, foreign direct investment is becoming an incredibly important tool for investors in developed nations and the developed nations in which they are investing. Investors have increasingly been seeking protections for their investments in foreign nations. This is why approximately 2400 bilateral investment treaties were signed between various nations between 1994 and 2006. When conflicts arise, the job of interpreting these treaties often falls to investment arbitration tribunals. Indeed, in 2010, an arbitration tribunal (Tribunal) operating under the United Nations Commission on International Trade Law (UNCITRAL) rules adjudicated a dispute between Chevron and the Republic of Ecuador (Ecuador) and interpreted the bilateral investment treaty between the U.S. and Ecuador (BIT). This Comment argues that the Tribunal’s interpretation of the BIT was the most reflective of the investor’s expectations and thus encouraged further investment. As incentivizing this investment is the very purpose of the BIT, the Tribunal reached the best possible conclusion as to its meaning.

THE INTERNET IS NOT A LAWLESS PRAIRIE: DATA PROTECTION AND PRIVACY IN ITALY

Noah C.N. Hampson

[pages 477–489]

Abstract: As the Internet has become more pervasive, so too have concerns about the security of personal data online. The rapid evolution of
Internet technology has outpaced the legislative process, leaving courts to resolve complex and important questions of policy. Their answers to these questions can have dramatic implications for the future of the Internet as a platform for global communication. The judicial decisions in this area are frequently issued ad hoc by judges who may be unfamiliar with the technology at issue and unaware of the potential ramifications of their rulings. The February 2010 conviction in Italy of three Google executives for violations of data protection laws sparked widespread controversy and criticism on this basis. This Comment argues that the Italian court’s decision is a prominent example of the broader trend of inexpert judicial regulation of the Internet.
DRAWING A LINE IN THE PATENT SUBJECT-MATTER SANDS: DOES EUROPE PROVIDE A SOLUTION TO THE SOFTWARE AND BUSINESS METHOD PATENT PROBLEM?

SUSAN J. MARSNIK*
ROBERT E. THOMAS**

Abstract: In June 2010, the Supreme Court issued its decision in Bilski v. Kappos, a case that had the potential to rewrite the landscape for determining what types of computer-related and business method inventions would receive patent protection. Just six weeks earlier, the European Patent Office’s Enlarged Board of Appeal delivered a decision on the same subject matter that had the potential to produce similar change in Europe. Yet, given these two opportunities to overhaul imperfect patent systems, neither decision provided more than incremental change. This Article explains why neither jurisdiction is able or willing to produce comprehensive reform in this area, and seeks to illuminate the nature of patent reform that is possible on the two continents.

INTRODUCTION .....................................................................................228
I. SOFTWARE AND BUSINESS METHOD PATENTS .................................233
   A. The Exclusion of Abstract Ideas From Statutory Subject Matter ........241
   B. The United States Supreme Court’s Cautionary Approach to Software
      Patents ..............................................................................................247
   C. Federal Court Treatment of Computer Software ..............................254
   D. The Bilski Case .............................................................................262
III. THE EUROPEAN PATENT OFFICE TREATMENT OF SOFTWARE PATENTS
    ...........................................................................................................266
   A. The European Patent System ..........................................................267
   B. Origins of the “Technical” Requirement .........................................272
   C. The EPO Technical Board of Appeal Decisions ..............................276
      1. The Technical Contributions Approach .....................................278

* Associate Professor of Ethics and Business Law, Opus College of Business, University of St. Thomas. With thanks to Leah Menden and Samuel Sands for their research assistance.
** Associate Professor of Legal Studies and Huber Hurst Fellow, Warrington College of Business, University of Florida.
The status of business methods and software as patent subject matter is one of the most controversial debates in patent law.¹ These patents, or a subset thereof, are considered to be overly broad in scope,² unlikely to advance the prior

---


³ John R. Allison et al., Extreme Value or Trolls on Top? The Characteristics of the Most-Litigated Patents, 158 U. PA. L. REV. 1, 3 (2009) (providing empirical evidence that the most
art due to lax disclosure requirements, and generally undeserving of patent protection. Due to the questioned legitimacy of these patents, there has been resistance to the unqualified ratification of business method and software patent subject-matter status on both sides of the Atlantic. Nevertheless, solutions to the perceived problems of business method and software patents have been ephemeral. Like shifting sand, proposals and approaches have been accepted and rejected at a seemingly increasing rate on both continents.

Litigated patents are held by non-practicing entities—so called “trolls” —and disproportionately cover software inventions).

4 See Burk & Lemley, supra note 2, at 1688–89 (arguing that weak disclosure requirements allow broad claims that can stifle “subsequent incremental improvements”).

5 See Rochelle Cooper Dreyfuss, Are Business Methods Patents Bad for Business?, 16 SANTA CLARA COMPUTER & HIGH TECH L.J. 263, 275–77 (2000) (arguing that the limited benefits from business method patents do not exceed their social costs); Thomas, supra note 1, at 210–11, 218 (extending Dreyfuss’ social welfare analysis to software patents).


The shifts have been so dramatic that the Court of Appeals for the Federal Circuit (CAFC) reaffirmed a seemingly discredited\(^8\) subject-matter test for business method and software patent claims.\(^9\) This reaffirmation left the Supreme Court with the unenviable task of either ratifying a test that previous courts refused to embrace or developing a viable solution that has eluded courts for decades.\(^10\) During oral arguments in *Bilski v. Kappos*, the Supreme Court seemed willing to explore all avenues for a possible solution to this conundrum.\(^11\) Justice Ginsburg queried whether the tied-to-technology requirement upon which European patents rest could provide a workable subject-matter test for business methods and software patents in the United States.\(^12\) Justice Ginsburg’s query suggests that at least some members of the Supreme Court did not find any of their options inviting and held some hope that European law might provide a remedy that the Supreme Court could use to rescue the United States’ patent system from its subject-matter malaise.

In this Article, we address Justice Ginsburg’s query by examining United States and European patent law to determine whether there is viable legal or policy support for a patent subject-matter test that provides the patent law community with clear guidelines for distinguishing “deserving” patents from “undeserving” patents. Now that the Supreme Court has issued a decision that basically orders the CAFC to “reboot” its patent process subject-matter approach—with little more than the exclusion against abstract ideas and 1980s-era Supreme Court precedents—the need for clear guidance in this area of patent law is even more pressing.\(^13\) The European Patent Convention (EPC)\(^14\) appears to contain provisions that address this problem. Courts have consistently

---


\(^9\) *Id.* at 963–66 (reaffirming use of the machine or physical transformation test to determine whether processes are valid statutory patent subject matter).


\(^12\) See *id*.

\(^13\) See *Bilski*, 130 S. Ct. at 3225–26, 3229.

interpreted the EPC to require that all patents be “technical,”\textsuperscript{15} and to exclude explicitly business methods and “programs for computers” from patent subject-matter coverage in Article 52(2).\textsuperscript{16} The strong exclusion is chimeric, however, offering less hope for a solution than previously considered U.S. approaches.

The clarity of the EPC exclusion is severely muddled by modifying language in Article 52(3)\textsuperscript{17} and by the complexity and fragmentation of the European patent system.\textsuperscript{18} Unlike the centralized United States patent system, national patent systems in Europe coexist with the European Patent Office (EPO), both of which apply and interpret the EPC.\textsuperscript{19} There are no pan-European courts that correspond to the CAFC and the U.S. Supreme Court.\textsuperscript{20} In Europe both the EPO and national courts have jurisdiction over patent subject-matter appeals.\textsuperscript{21} As one commentator put it, “[t]he resulting edifice is byzantine in complexity.”\textsuperscript{22} Although national courts strive to harmonize their decisions with EPO Technical Board of Appeal decisions, significant conflicts have, nonetheless, developed. For example, the potential exists for a computer software patent granted by the EPO to be invalidated under current U.K. law because of that country’s more restrictive interpretation of the EPC.\textsuperscript{23} Realization of this potential would render one of the EPC’s main tenets meaningless: namely, that patents granted at the EPO are valid in contracting states as if they were granted by the national office.\textsuperscript{24}

Elevating the probability of legal discord is the EPO Enlarged Board of Appeal’s denial of the existence of conflicting decisions within the EPO and the Board’s refusal to clarify the meaning of “technical” as

\begin{itemize}
\item \textsuperscript{15} Thomas & DiMatteo, \textit{supra} note 1, at 17 (“While there is no explicit requirement in the EPC for technical character or a ‘technical contribution,’ the patent courts initially interpreted the EPC as including such a requirement.”).
\item \textsuperscript{16} EPC 2000, \textit{supra} note 14, at 271–72.
\item \textsuperscript{17} See id. at 272 (limiting Art. 52(2) exclusions to the extent that a patent or patent application relates to the excluded subject matter or activities “as such”).
\item \textsuperscript{18} \textsc{William Cornish} & \textsc{David Llewelyn}, \textit{Intellectual Property: Patents, Copyright, Trade Marks and Allied Rights}, 114 (6th ed. 2007) (describing the EPO and UK patent systems and their interaction); \textit{infra} Part IV(B).
\item \textsuperscript{19} See EPC 2000, \textit{supra} note 14, at 258–59, 294–96.
\item \textsuperscript{20} See id. at 259–71 (outlining the structure of European institutions of patent review).
\item \textsuperscript{21} \textsc{Cf. Cornish} & \textsc{Llewelyn}, \textit{supra} note 18, at 114 (describing the EPO and U.K. patent systems and their interaction).
\item \textsuperscript{22} \textit{Id.}
\item \textsuperscript{23} \textit{See Guy Tritton et al., Intellectual Property in Europe} 86 (3d ed. 2008).
\item \textsuperscript{24} Shemtov, \textit{supra} note 7, at 514.
\end{itemize}
applied to European patent subject matter. Therefore, in addition to shaping patent subject-matter policy for business methods and software, Europe must also resolve conflicting approaches within the EPC and between countries with disparate and sometimes inconsistent approaches to patent policy. Europe cannot take these steps until there is a European patent court with jurisdiction to settle these interpretive differences.

Another impediment to European reform is the extremely liberal U.S. approach to software and business method patents. Implementing clear guidelines and boundaries could handicap European inventors and businesses relative to their U.S. counterparts, which enjoy the liberal U.S. treatment of software and business method claims. Hence, clarity and well-defined constraints in Europe may occur only after the United States has shifted its patent policy to include clear limitations on software and business method patents. Thus, it appears highly unlikely that Europe can provide the U.S. patent system with the guidance it apparently seeks.

This Article’s analysis of U.S. and European approaches to the patentability of business methods and software supports this rather pessimistic conclusion. Conceptually, the only significant differences between these invention types and mental processes that do not receive patent protection are speed, capacity, and accuracy. There is nothing that computer software controlling an electronic device cannot perform that a human mind cannot also perform using paper and pencil and sufficient time. Nevertheless, these differences coupled with technologically advanced computer equipment make software valuable and capable of performing tasks that cannot be done practically by the human mind alone.

Part I of the Article discusses the nature of software and business methods. Parts II, III, and IV, respectively, examine the development and present state of patent law in the United States, European Patent Office, and United Kingdom. The Article concludes by identifying the implications of this discussion. Our analysis shows that there is no unequivocal legal support for a patent subject-matter rule that excludes

---

26 See infra Part I.
27 See infra Part II.
28 See infra Part III.
29 See infra Part IV.
all inventions that include software or business methods under either patent system.\textsuperscript{30} There is also no unequivocal support for recognizing software and business methods as patent subject matter.\textsuperscript{31} The problem lies in the absence of clear, unambiguous legislative direction on either continent.\textsuperscript{32} The manifestation of this lack of a solid policy foundation in the United States has been the repeated adoption and rejection of different judicial approaches.\textsuperscript{33} In Europe, on the other hand, the lack of legislative direction combined with the absence of a centralized court system has resulted in conflicting and shifting approaches in the judicial treatment of software patent subject matter.\textsuperscript{34}

I. Software and Business Method Patents

To appreciate the dilemma that courts and policy makers face in addressing the software and business method patentability question, understanding the nature of software is critical. Therefore, we begin our analysis by discussing the general nature of software and business method patents and their treatment under U.S. law. Conceptually, software and business methods are closely related.\textsuperscript{35} Both are abstract processes that do not independently produce tangible results.\textsuperscript{36} An ad-

\textsuperscript{30} See infra Parts II–IV.


\textsuperscript{32} See infra text accompanying notes 166, 366–406.

\textsuperscript{33} Compare Bilski, 130 S. Ct. at 3231 (rejecting prior test for patentable subject matter and permitting the Federal Circuit to develop a new test based on the abstract ideas exclusion), with State St. Bank & Trust Co. v. Signature Fin. Grp. Inc., 149 F.3d 1368, 1373 (Fed. Cir. 1998) (recognizing the patentability of a computer algorithm that produces a “useful, concrete and tangible result”).

\textsuperscript{34} See, e.g., VICOM, [1987] O.J.E.P.O. at 14 (noting uncertainty in claims involving business methods or software and additional variables).

\textsuperscript{35} See Allison & Tiller, supra note 1, at 1012 (arguing, in a section entitled “A Bit of Software Patent Déjà Vu,” that most criticisms of business method patents have already been applied toward software patents); Thomas, supra note 1, at 193 n.6 (arguing that “[b]usiness methods are part of the continuum [of patents] that includes software and computer-implemented inventions”).

ditional similarity is that business methods are often implemented through computer software. Under the “machine-or-transformation test” championed in In re Bilski, business methods that are not computer or machine implemented are unlikely to survive a subject-matter challenge. Whether business methods or computer software are statutory subject matter when “computer implemented” is a more difficult determination. When “computer implemented,” the business method is computer software. A major question that courts on both sides of the Atlantic have struggled with is whether implementing software or a business method through a computer or other machine is necessary or sufficient to render such claims valid statutory subject matter. Or, more generally, under what circumstances are business methods or computer programs capable of becoming patentable inventions?

Therefore, determining whether and under what circumstances software should be valid statutory subject matter under U.S. law, or statutorily excluded subject matter under European law, requires an understanding of the nature of computer software. Modern computers consist of tangible physical components including one or more processing units, graphical and input-output subsystems, memory chips, storage devices, and other support systems. Software provides the means that business method and software patents had crossed into the realm of thought and abstraction with nuts and bolts that are “vaporous and intangible”).


39 See Hand, supra note 37, at 470–71.


by which human users interact with and control these myriad computer systems.\textsuperscript{42} Software allows computer users to give instructions to computers to perform a variety of tasks.\textsuperscript{43} The term software also applies to information stored on, used, and manipulated by computers.\textsuperscript{44} The term software additionally includes the internal routines that allow different parts of the computer to interact with each other and to translate human input into instructions that computers understand.\textsuperscript{45}

Conceptually, there is little to distinguish software patent claims from claims involving abstract ideas, algorithms, and mental processes that courts consistently rejected prior to the 1998 case before the CAFC, \textit{State Street Bank v. Signature Financial Group}.\textsuperscript{46} Computers and the human mind operate similarly by processing algorithms.\textsuperscript{47} Mathematical algorithms, computer programs, and mental processes are ways of “defining abstract relationships among concepts and [with] defining rules about how those concepts should be manipulated.”\textsuperscript{48} Thus, for example, the area of a rectangle can be characterized by the relationship between the base and height of the geometric figure and defined as the product of those two elements.\textsuperscript{49} There is no conceptual difference between calculating this area in one’s mind, with pencil and paper, and calculating this area through the use of a programmed computing device. Many psychologists model human thought processes as a series of computational steps.\textsuperscript{50} According to these psychologists, human thought proc-

\begin{footnotesize}
\begin{enumerate}
\item[G42] Gregory A. Stobbs, \textit{Software Patents}, § 2.02 (2000) (“[S]oftware is what empowers a computer to handle information and to control information flow.”).
\item[G43] Haynes, \textit{supra} note 41, at 247 (“A computer is useless without software.”).
\item[G44] Stobbs, \textit{supra} note 42, § 2.02 (“[S]oftware is information that is fed into the input, placed in storage, and then delivered from storage to the computer.”).
\item[G45] Andrew Rodau, \textit{Computer Software: Does Article 2 of the Uniform Commercial Code Apply?} 35 Emory L.J. 853, 867–68, 868 n.57 (1986) (noting that the term “software” is confusing because it applies to many different aspects of the computing process including internal and external computer functions).
\item[G46] See 149 F.3d 1368, 1373 (Fed. Cir. 1998).
\item[G48] Samuelson, \textit{supra} note 47, at 1123.
\item[G49] See \textit{id}.
\item[G50] \textit{Id} (quoting Professor Newell as stating, “humans think by means of algorithms. Sequences of mental steps and algorithms are the same thing.”).
\end{enumerate}
\end{footnotesize}
esses, mathematical equations, and computer programs are algorithms that have no conceptual difference.\textsuperscript{51}

The difference between computer processing and human mental processes is thus more quantitative than qualitative. The computer is faster, more capacious, and more accurate than the human mind.\textsuperscript{52} These improvements allow computers to accomplish tasks that cannot be accomplished by humans working alone. Managing and manipulating scientific experiments, rocket launches, and graphical representations all require the speed, precision, and tremendous storage capacity of computers.\textsuperscript{53} The 1981 Supreme Court case \textit{Diamond v. Diehr} provides a useful illustration of the value of computer technology.\textsuperscript{54} The \textit{Diehr} patent claim provided a novel way of curing artificial rubber.\textsuperscript{55} The mathematical formula for completing this process, called the Arrhenius equation, was well known before submission of the \textit{Diehr} application.\textsuperscript{56} Applying the formula in an industrial context was difficult, however, because it required continual monitoring and adjustments to determine the precise time to terminate the curing process.\textsuperscript{57} Without computer aid, humans could not collect process data and perform the required constant calculations to determine the optimal time to terminate the curing process.\textsuperscript{58} The \textit{Diehr} patent claim included a computer that was capable of accurately completing the repetitive calculations required by the Arrhenius formula and applying adjustments to the industrial process.\textsuperscript{59}

Mixed processes such as the \textit{Diehr} method for curing rubber are at the crux of the subject-matter problem. EPC law and U.S. commentators who wish to limit patent subject matter agree that pure mental

\textsuperscript{51} See \textit{id.} at 1123–24 (“\textquote{A}n algorithm for representing how a legal problem can be solved is just as \textquote{mathematical} as an algorithm for addition, for finding the lowest common divisor for two numbers.”).

\textsuperscript{52} Laura R. Ford, \textit{Alchemy and Patentability: Technology, \textquote{Useful Arts,} and the Chimerical Mind-Machine}, 42 \textit{Cal. W. L. Rev.} 49, 53 (2005) (explaining how modern computers are able to carry out tasks traditionally performed by human mental processes at a speed and level of accuracy that far exceeds human capabilities).


\textsuperscript{54} 450 U.S. 175, 177–78 (1981).

\textsuperscript{55} \textit{id.} at 177.

\textsuperscript{56} \textit{See id.}

\textsuperscript{57} \textit{id.} at 178.

\textsuperscript{58} \textit{See id.} at 178–79.

\textsuperscript{59} \textit{See id.}
processes and inventions without industrial or technical applications are not valid patent subject matter. European courts and these commentators have trouble answering the question of how much more than a mental process is needed for an invention to receive a patent. In Diehr, the Court ruled that an invention that employed a mathematical algorithm was patent subject matter in large part because it was part of an industrial process. Such industrial processes, while employing software and mathematical algorithms, satisfy the Bilski machine or physical transformation test because the process as a whole takes industrial raw materials as input and produces a finished manufacture as an output.

Computers primarily interact with information when not connected to an industrial process. Most computer end-users employ application programs, either purchased or custom-made, to perform desired tasks. Common computer uses include word processing, database management, statistical and financial analysis, photograph and video editing, and game playing. A program that solely applies a mathematical formula to data and delivers a result does not qualify as statutory subject matter. Nevertheless, if the computer running the program is connected to a plotter or computer monitor that draws a graph of the calculation results, at least one court has held that the physical output is enough to render the entire process valid statutory subject matter.

If providing physical manifestations of data analysis is sufficient to qualify a process containing a computer program as statutory subject matter, however, there are virtually no constraints on software subject-

---

60 See EPC 2000, supra note 14, at 108 (excluding business methods and computer programs as patentable subject matter); Thomas F. Cotter, A Burkean Perspective on Patent Eligibility, 22 Berkeley Tech. L.J. 855, 860–61 (2007) (arguing that the disfavored mental steps doctrine still has merit).

61 See, e.g., Case T-208/84, Computer-Related Invention/VICOM, [1987] O.J.E.P.O. 14, 14–23 (Technical Bd. Appeal 3.5.01, July 15, 1986), available at http://archive.epo.org/epo/pubs/oj1987/p001_046.pdf (interpreting the “as such” modifier in EPC Art. 52 in the context of a computer-aided design program whose only contribution over the corresponding mental process was speed, and remanding case to Examining Division to consider redrafted claims).

62 See Diehr, 450 U.S. at 192–93.

63 See id.

64 See Stobbs, supra note 42, § 2.02.

65 See id. § 1.04.

66 See id. § 2.02.

67 Benson, 409 U.S. at 71–72 (noting that granting a patent on such a claim would amount to granting a patent on a mathematical formula or pure idea).

matter patentability. The nature of computers and software engineering means that most software can include output capabilities. Modern programming and computer design employ a modular approach to manage the complexity of large programs. Analogous to automobiles—which are built with tires, engines, brakes, and other components that perform specific discrete tasks—computer application programs are also built from multiple components that perform discrete tasks. Software functions may specify how the program handles data input-output routines or may provide methods for performing mathematical and statistical operations. These functions are offloaded to specialized systems that handle input-output and other functions. The programmer only needs to know the commands required to evoke the subsystems to produce desired results: there is no need to know exactly how the specialized subsystems achieve such results. The programmer’s ability to ignore redundant or highly specialized tasks is facilitated by the availability of off-the-shelf and generic function libraries. These libraries can be used to provide capabilities for any program that requires the included tasks. Programmers only need to understand the operation and syntax of library functions. In fact, it is possible to create a word processor and other common computer applications using function libraries and just enough programming code to integrate the functions into a cohesive whole. Most software programs thus run on “generic” computers and utilize standard methods for interacting with the external world that are available in off-the-shelf libraries, such as input-output, printing, and audio-visuals. Thus, it seems that programmers need to focus on solving discrete problems rather than engaging in elaborate software engineering, unless available libraries are inefficient or deficient in some significant respect. Therefore, much software innovation is at the abstract information or algorithmic level.

69 See Stobbs, supra note 42, § 2.02.
70 Thomas, supra note 1, at 219.
71 Id.
72 Id.
73 Id.
74 Id.; see, e.g., Walter Savitch, Absolute C++ 92 (1st ed. 2002) (“C++ comes with libraries of predefined functions that you can use in your programs.”).
75 See Thomas, supra note 1, at 219.
76 See id.
77 Id.
78 See id.
79 See id.
Additionally, software patents suffer from excessive breadth. Software patent applications do not include source code—the program written in a human language—and the United States Patent and Trademark Office (USPTO) often approves claims consisting of little more than a rudimentary flow chart. As a result, software patent holders lay claim to broad areas of software practice without well-identified claim boundaries, with virtually no implementation details, and with few clues about the quality of claim implementation. This practice appears inconsistent with patent law, which requires a claim’s application to provide sufficient detail such that someone of ordinary skill in the relevant art is able to practice the invention. This is the quintessential means by which knowledge is disseminated through the patent process.

The relaxation of disclosure requirements for software patent claims in the United States, codified in U.S. Code chapter 35, section 112, is due to necessity. Ironically, merely providing software source code would more fully satisfy the Section 112 disclosure requirements than the current practice. Source code consists of human-readable instructions that, when translated into a computer-readable format, give the computer detailed instructions that set forth the order in which the computer executes program steps and functions. Source code is equivalent to a detailed blueprint of the program’s construction and functions. Including software source code or detailed outlines of how a program operates would easily satisfy Section 112 enablement and best mode requirements.

Nevertheless, the CAFC has ruled that disclosing outcomes—or “functions”—without identifying detailed procedures—or “means”—meets the Section 112 burden for computer software. Section 112 disclosures, in addition to creating a publicly accessible record of the invention, provide a check over the breadth and scope of claims. To avoid overly expansive claims, patent law provides protection solely for

---

80 See Klemens, supra note 47, at 73.
81 See id., at 21–22.
82 See id.; Cohen & Lemley, supra note 2, at 24–25.
84 See Cohen & Lemley, supra note 2, at 17–19.
87 See Stobbs, supra note 37, § 2.06 (b)–(c).
88 Thomas, supra note 1, at 234.
89 See Fonar, 107 F.3d at 1548–49; see also Cohen & Lemley, supra note 2, at 24 n.87.
the means of achieving identified functions—the “means-plus-function test.”

But requiring software patents to meet the means-plus-function criteria limits their efficacy. In programming, there are usually many different ways to accomplish a desired result. Programmers may use different routines, approaches, and languages to accomplish the same programming task. Different programmers will address a given problem in many—often significantly—different ways reflecting style, emphasis, priorities and skills. Nonetheless, each of these programs produces the same result—or function—while employing a distinctly different means. Requiring software claims to satisfy the means-plus-function test would mean that software patents would only be infringed when the source or machine code of the alleged infringing program is identical to the patented program. The myriad ways to replicate the patented program’s function without duplicating code would not be infringing. Therefore, limiting software patents with a means-plus-function test would substantially reduce the value of patents meeting the means-plus-function standard.

The discussion in this section has identified some of the dilemmas software patents create. In particular, allowing limited software patent disclosure results in overly broad patent scope, which is anti-competitive. Although excluding computer software from statutory subject matter would appear to address these problems, the European experience indicates that such a prohibition is extremely difficult to implement. In the next section, we examine how U.S. courts have dealt with software patent claims and why the issue presents such a conundrum. The striking similarities between U.S. and European courts struggling with determining the patentability of software patent claims is particularly enlightening.

91 See id.
92 See Fonar, 107 F.3d at 1548–49 (discussing how source code is not sufficient to allow a software engineer to replicate the code because source code is machine-specific and it is more “important . . . to have a description of what the software has to do”).
93 See Klemens, supra note 47, at 43.
94 Thomas, supra note 1, at 235.
95 See Klemens, supra note 47, at 43 (“In view of the astounding number of choices available in such an exercise, the two programmers’ solutions could be vastly different.”)
96 Id.
97 See Thomas, supra note 1, at 234–35.
98 See id. at 235–36.
99 See Klemens, supra note 47, at 73; Burk & Lemley, supra note 2, at 1688–89.
100 See infra Parts III–IV.
II. U.S. Judicial Treatment of Software Patents

A. The Exclusion of Abstract Ideas From Statutory Subject Matter

Whether software and business methods can be classified as statutory subject matter depends on the location of the boundaries of patent protection. The constitutional authorization for patents envisioned boundaries by granting Congress the power “[t]o promote the Progress of . . . Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.”101 Initial U.S. patent legislation identified statutory subject matter, but did not enumerate any subject-matter exclusions.102 The courts accepted the responsibility of defining the limits of patent protection.103 They were cognizant of the dangers of allowing inventors to obtain overly expansive coverage of their claims and designed tests to curtail the scope of overreaching patent claims.104 In particular, courts understood that allowing overly broad statutory subject matter could impede industrial innovation.105 These tests precluded patent claims for abstract ideas, functions and effects, and mental steps,106 and the exclusions lasted well into the twentieth century.107 In 1972, the Supreme Court included “laws or principles of nature, mental processes, mathematical expressions and formulas, and abstract intellectual concepts” among excluded subject matter because “they are the basic tools of scientific and technological work.”108

The primary policy objective behind the early judicial approach to statutory subject matter was to limit the anti-competitive impact of patents while maintaining innovation incentives.109 Because granting exclusive rights is anti-competitive, the courts attempted to limit the subject area of patent coverage to technological-industrial innovations.110 Exclusive rights and the concomitant competitiveness losses are the

---

101 U.S. Const. art. I, § 8, cl. 8.
102 See Patent Act of 1790, ch. 7, § 1 (1790) (identifying statutory subject matter as “any useful art, manufacture, engine, machine, or device, or any improvement therein not before known or used.”).
104 See id.
105 See LeRoy v. Tatham, 55 U.S. 156, 175 (1852) (noting that excessive patent scope would retard innovation and development in a manner contrary to patent policy).
107 See id.
108 Id.
109 See supra notes 99–103 and accompanying text.
costs that society incurs in return for encouraging innovation in the useful arts and furthering rapid dissemination of new knowledge.\textsuperscript{111} If the patent grant is overly expansive, however, the anti-competitive effects of granting exclusive rights impede innovation without significantly increasing the volume of knowledge in the public domain.

Early cases illustrate this judicial sensitivity. In 1852, the U.S. Supreme Court provided detailed rules and rationales for limited patent protection in \textit{LeRoy v. Tatham} based on patent policy from the United States and England.\textsuperscript{112} The Court, in correcting the trial court’s jury instructions in a patent infringement case, enunciated policy principles that limited the scope of patent law.\textsuperscript{113} The trial court had instructed the jury that a patent claim for producing lead pipe should be validated regardless of whether the machine used was novel because the innovation consisted of “bringing a newly discovered principle into practical application.”\textsuperscript{114} The Court stated as a foundational rule that inventors could not patent abstract principles and natural laws.\textsuperscript{115} To prevent inventors from overreaching with respect to abstract principles and laws of nature, the Court stated what has since been termed the means-plus-function test.\textsuperscript{116} In essence, an inventor could not obtain a patent on the result or effect of a process—in this case, the production of lead pipe by exploiting a particular property of lead—but could obtain exclusivity solely on the means by which the process achieved its effect.\textsuperscript{117} The Court reasoned that this limitation was needed to keep the abstract principle in the public domain.\textsuperscript{118} Allowing such exclusivity would discourage rather than promote advancement in the useful “arts and manufactures.”\textsuperscript{119}

\begin{flushright}
\textsuperscript{111} See Dan L. Burk, \textit{The Role of Patent Law in Knowledge Codification}, 23 \textit{Berkeley Tech. L.J.} 1009, 1010 (2008) (“The rationale for patenting long favored in judicial opinion is the ‘\textit{quid pro quo}’ theory: that patents are a bargain of sorts, between the inventor and the public, exchanging public disclosure of the claimed invention in return for the grant of a period of exclusive rights.”).
\textsuperscript{112} See \textit{LeRoy}, 55 U.S. at 175–76.
\textsuperscript{113} \textit{Id.}
\textsuperscript{114} \textit{Id.}
\textsuperscript{115} \textit{Id.}
\textsuperscript{116} \textit{Id.}
\textsuperscript{117} \textit{Id.}
\textsuperscript{118} \textit{Id.}
\textsuperscript{119} \textit{Id.}
\end{flushright}
The ruling in *LeRoy* was not unanimous.\textsuperscript{120} In a strong dissent, three justices argued that inventors who discover a way to exploit a natural law for a particular purpose should be entitled to exclusivity over all uses of the natural law to achieve the useful outcome.\textsuperscript{121} The dissent argued that the true innovation was in recognizing the application of the natural principle for a useful purpose, and thus, limiting exclusivity to the method or mode was contrary to patent policy.\textsuperscript{122} The implicit assumption in the dissent was that inventors should be rewarded for their ingenuity. The majority opinion’s limitation of exclusivity provided inadequate incentives or rewards.\textsuperscript{123} The dissenting opinion also dismissed the majority’s concern about social welfare losses, arguing that exclusivity would be limited both in scope and term.\textsuperscript{124} Inventors would be free to use the abstract principle for any other function, and inventors could use the principle for any function at the conclusion of the patent term.\textsuperscript{125} Thus, according to the dissenting opinion, welfare losses would not be significant.

Despite the *LeRoy* dissent’s argument for broader patent scope, subsequent courts adopted the more restrictive subject-matter requirements of the *LeRoy* majority. The 1853 Supreme Court case *O’Reilly v. Morse* exemplified this trend.\textsuperscript{126} The Court in *O’Reilly* was particularly concerned with the over-expansive nature of claims based on abstract ideas.\textsuperscript{127} Henry O’Reilly challenged the validity of patents held by Samuel Morse on the invention and improvement of telegraph technology.\textsuperscript{128} The specific patent in question contained eight claims, the eighth of which was the subject of the controversy.\textsuperscript{129} Morse claimed as follows:

Eighth. I do not propose to limit myself to the specific machinery, or parts of machinery, described in the foregoing specifications and claims; the essence of my invention being the use of the motive power of the electric or galvanic current, which I call electro-magnetism, however developed, for

\begin{itemize}
  \item \textsuperscript{120} See id. at 177.
  \item \textsuperscript{121} Id. at 187.
  \item \textsuperscript{122} Id.
  \item \textsuperscript{123} See *LeRoy*, 55 U.S. at 187.
  \item \textsuperscript{124} See id.
  \item \textsuperscript{125} Id.
  \item \textsuperscript{126} See O’Reilly v. Morse, 56 U.S. 62 passim (1853).
  \item \textsuperscript{127} See id. at 135.
  \item \textsuperscript{128} See id. at 63–65.
  \item \textsuperscript{129} See id. at 85–86.
\end{itemize}
making or printing intelligible characters, letters, or signs, at any distances, being a new application of that power, of which I claim to be the first inventor or discoverer.\textsuperscript{130}

The Court invalidated the claim due to its excessive breadth.\textsuperscript{131} The Court reasoned that validating the claim would mean that “it matters not by what process or machinery the result is accomplished.”\textsuperscript{132} Morse would nonetheless have the exclusive right to that new invention or improvement. The Court, identifying social welfare-reducing implications of granting Morse exclusivity, stated that a competitor’s improved “invention may be less complicated—less liable to get out of order—less expensive in construction, and in its operation. But yet if it is covered by this patent the inventor could not use it nor the public have the benefit of it without [Morse’s permission].”\textsuperscript{133} Therefore, if the Court had granted Morse such broad exclusive rights, other inventors would have no incentive to conduct research in this area because the Morse patent would prevent the inventor from exploiting and profiting from the broad invention without Morse’s permission.\textsuperscript{134} Thus, unless Morse had the capability and willingness to duplicate the efforts of such inventors, there would likely be less innovation in the field of the Morse patent for the duration of the patent term, leaving society worse off.\textsuperscript{135}

Additional social welfare losses result from broad patent scope. As the \textit{O’Reilly} Court noted, not only would it have a chilling effect on competing inventors, but a patent grant for Morse’s eighth claim would reduce the flow of knowledge into the public domain.\textsuperscript{136} Morse would be free to improve and advance the subject matter of the patent without revealing such advancements to society. He would “need place no description of the new manner, process, or machinery, upon the records of the patent office.”\textsuperscript{137} Moreover, at the end of the patent term, “the public must apply to [Morse] to learn what it is.”\textsuperscript{138} Therefore, validating the expansive eighth claim could provide Morse with both patent protection during the patent term and trade secret protection at patent expiration. This result, while benefiting Morse greatly, would

\begin{itemize}
\item \textsuperscript{130} \textit{Id.} at 86.
\item \textsuperscript{131} \textit{Id.} at 113.
\item \textsuperscript{132} \textit{O’Reilly}, 56 U.S. at 113.
\item \textsuperscript{133} \textit{Id.}
\item \textsuperscript{134} \textit{Id.}
\item \textsuperscript{135} \textit{See id.}
\item \textsuperscript{136} \textit{See id.}
\item \textsuperscript{137} \textit{Id.}
\item \textsuperscript{138} \textit{O’Reilly}, 56 U.S. at 113.
\end{itemize}
leave society with no gain from incurring the costs of granting Morse exclusivity in this industrial area.\textsuperscript{139}

In 1876, the Supreme Court revisited the question of identifying the boundaries of patentable subject matter in \textit{Cochrane v. Deener}.\textsuperscript{140} \textit{Cochrane} remains relevant because of its characterization of patent processes. In the case, claimant Cochrane had multiple patents on a process and machinery to produce refined flour.\textsuperscript{141} His process patent claim covered the entire process of grinding and filtering flour with the use of air current to remove impurities.\textsuperscript{142} Similar to Morse’s broad process claim,\textsuperscript{143} Cochrane did not limit his claim to any particular machine. He claimed priority for any process employing the collective elements of his claim to refine flour.\textsuperscript{144} The defendant had an improved method of refining flour that used a different type of machinery to effectuate the same function.\textsuperscript{145} The defendant believed that this difference was sufficient to avoid infringement.\textsuperscript{146} Unlike \textit{O'Reilly}, however, the Court did not conclude that Cochrane’s claim was overly broad.\textsuperscript{147} In ruling that the defendant’s method was infringing, the Court concluded that any alternative process that duplicated the steps of the patented process would infringe that patent regardless of whether the steps were accomplished in a different manner.\textsuperscript{148} The particular method used to accomplish the function was irrelevant.\textsuperscript{149} This conclusion encouraged broad patent claims in its aftermath. Nevertheless, if a claim was too broadly drawn, the \textit{O'Reilly} ruling would disqualify the claim from statutory subject matter.\textsuperscript{150}

The \textit{Cochrane} Court addressed the patent scope concern by providing a narrow definition of patent processes. The Court reasoned as follows: “[A] process is a mode of treatment of certain materials to produce a given result. It is an act, or a series of acts, performed upon the subject matter to be transformed and reduced to a different state or thing.”\textsuperscript{151} Under this reasoning, a claim would describe a patent quali-

\textsuperscript{139} See \textit{id.}
\textsuperscript{140} Cochrane v. Deener, 94 U.S. 780, 786–88 (1876).
\textsuperscript{141} \textit{Id.} at 781–85.
\textsuperscript{142} \textit{Id.} at 785.
\textsuperscript{143} \textit{O'Reilly}, 56 U.S. at 85–86.
\textsuperscript{144} \textit{Cochrane}, 94 U.S. at 784–86.
\textsuperscript{145} \textit{Id.} at 785–86.
\textsuperscript{146} \textit{Id.} at 786.
\textsuperscript{147} See \textit{id.} at 787–88.
\textsuperscript{148} \textit{Id.} at 788.
\textsuperscript{149} See \textit{id.}
\textsuperscript{150} See \textit{O'Reilly}, 56 U.S. at 113.
\textsuperscript{151} \textit{Cochrane}, 94 U.S. at 788.
fying process so long as the claim was sufficiently detailed to describe a series of steps or actions taken to physically transform a material object to a different state or thing.\textsuperscript{152} Abstract ideas and mental processes would not satisfy this test, nor would computer programs that were not part of some larger process.\textsuperscript{153} Thus, the \textit{Cochrane} physical transformation test provided a safe harbor for claims that might be subject to challenge for being abstract or overly broad.

For some time, the \textit{Cochrane} process definition limited the scope of process patent claims.\textsuperscript{154} Interestingly, at least one recent commentator has dismissed this interpretation of \textit{Cochrane} as relying on incorrect dicta.\textsuperscript{155} The basis for such criticism, however, appears to stem solely from the fact that contemporary federal court decisions have rejected the \textit{Cochrane} physical transformation test.\textsuperscript{156} An alternative view is that the courts have adopted a policy mandate to limit the scope of statutory subject matter due to the anti-competitive nature of patents.\textsuperscript{157} Thus, limiting statutory subject matter to the types of inventions Congress conceived of when enacting the first patent act would be consistent both with this policy objective and congressional intent.\textsuperscript{158} The physical transformation test furthers this objective by limiting patent protection to industrial and manufacturing innovations.\textsuperscript{159} Whereas the test is flexible enough to accommodate new types of innovations within established categories, it does not allow protection for innovation in different or new categories of innovation such as business methods or computer software. Arguably, expansion of patent protection to new categories of innovation is in the sole domain of Congress.\textsuperscript{160} Thus, limiting the spread of patent protection is more emblematic of judicial

\textsuperscript{152} See id.
\textsuperscript{153} See, e.g., Diehr, 450 U.S. at 191–93 (ruling that a process patent that contained a mathematical algorithm as its sole novel element was valid subject matter because it was part of an industrial process for creating artificial rubber molds).
\textsuperscript{155} See Ambrose, supra note 154, at 907–08 (“[C]ourts interpreted dicta in the landmark case of \textit{Cochrane v. Deener} to mean that patentable processes must operate on physical substances, and the courts therefore denied patents to methods requiring only the use of the human mind and writing implements.”) (citations omitted).
\textsuperscript{156} See id.
\textsuperscript{157} See Thomas & DiMatteo, supra note 1, at 6.
\textsuperscript{158} See id. at 6–9.
\textsuperscript{159} See Diehr, 450 U.S. at 188–93; \textit{Cochrane}, 94 U.S. at 787–89; see also Thomas, supra note 1, at 193–97.
\textsuperscript{160} See U.S. Const, art. I, § 8, cl. 8. But see Thomas, supra note 1, at 194.
restraint than allowing the unfettered expansion of statutory subject matter that has occurred in recent decades.

The judicial policy of constraining statutory subject matter continued in the twentieth century. Until recent decades, patent claims that consisted of steps that take place in the human mind or require human intervention were excluded from statutory subject matter.161 “Mental steps jurisprudence,” which developed in a series of cases over several decades, construed valid subject matter as excluding claims that require human calculation, measurement, and interpretation.162 In 1951, the Court of Appeals in In re Abrams delineated the rules for applying the mental steps doctrine.163 The case identifies three possibilities: first, all steps of a process claim are mental; second, the claim consists of both mental and non-mental steps but the novelty lies entirely in the mental steps; and third, a mixed claim for which the novelty resides in non-mental steps and the mental steps are incidental parts of the process but are needed to limit or define the claim.164 Under the doctrine, only claims that fall in the third category qualify as statutory subject matter.165

B. The United States Supreme Court’s Cautionary Approach to Software Patents

The U.S. Congress has assiduously avoided addressing whether computer software is patentable, thereby leaving the judicial system to shape policy.166 With the rapid development of computer technology, the Supreme Court first considered the subject-matter question in 1972 in Gottschalk v. Benson.167 The claim considered was for a “method for converting binary-coded decimal (BCD) numerals into pure binary numerals.”168 This sweeping method was not tied to any particular machine or programming language. In fact, the claim method could be performed mentally or on paper without a computer.169 As a result, it was easy for the Supreme Court to dispose of the claim because it was

161 See Ambrose, supra note 154, at 903; Cotter, supra note 60, at 860–61; Samuelson, supra note 47, at 1037.
162 See Samuelson, supra note 47, at 1034.
163 See 188 F.2d 165, 166 (C.C.P.A. 1951).
164 See id.
165 See id.
166 See Benson, 409 U.S. at 73 (noting that “considerable problems are raised which only committees of Congress can manage, for broad powers of investigation are needed, including hearings which canvass the wide variety of views which those operating in this field entertain.”).
167 Id. at 64.
168 Id.
169 See id. at 66–67.
“so abstract and sweeping as to cover both known and unknown uses of the BCD to pure binary conversion.”

Although disposal of the BCD claim was straightforward, the Court was concerned about whether its opinion would be interpreted as standing for the proposition that computer software could never be patented. The Court stated that “[w]e do not hold that no process patent could ever qualify if it did not meet the requirements of our prior precedents,” and “[w]e do not hold] that the [Benson] decision precludes a patent for any program servicing a computer.” Nevertheless, the Court, observing that “pure” software patent claims had previously been denied, noted that there existed considerable practical confusion in dealing with some software patent claims camouflaged as “a process, or a machine or components thereof . . . rather than as a program itself.” In other words, whether or not Congress chose to extend statutory subject matter to cover pure software programs, congressional action was still required to give the USPTO and courts guidance on how to handle hybrid claims.

The Court was clearly uncomfortable with accepting the reins of policy makers. Rather than give the public guidance as to how to treat patent claims, the Benson Court issued an impassioned plea for Congress to accept its responsibilities to develop patent policy in the legislative forum. Thus, while identifying its machine or physical transformation test as the relevant precedent, Benson emphatically refused to assert that the test applies to all software patent claims. The Court intentionally left the decision ambiguous in the hopes of eventual congressional intervention.

With the failure of Congress to accept the Court’s invitation to take action, the Supreme Court issued its second ruling on software patentability. In 1977, in Parker v. Flook, the Court considered a process claim that calculated an alarm limit. In certain industrial applications, operators—whether human or machines—need to receive notifi-
cation when certain process variables reach or exceed a designated threshold—the alarm limit—in order to make adjustments to maintain efficiency or to avoid dangerous conditions.\textsuperscript{178} The \textit{Flook} claim, although machine independent, provided a method for continually adjusting the alarm rate based on changes in process variables.\textsuperscript{179} The only novel element of the process was the use of an algorithm for calculating alarm rates.\textsuperscript{180}

In rejecting the patentability of the \textit{Flook} claim, the Court did not directly consider whether computer software per se is patentable. The Court, acknowledging in a footnote that one could argue that Supreme Court precedent requires processes to change materials to a “different state or thing,” refused to apply those precedents to the \textit{Flook} claim.\textsuperscript{181} Moreover, the Court left open the question of whether it considered software as part of the category of unpatentable algorithms. Instead, the Court’s new test basically required examiners to remove the algorithm from consideration before evaluating the patentability of the claim.\textsuperscript{182} Under this test, courts and examiners must first consider the algorithm “well known” in the prior art to avoid biasing the overall evaluation of the claim.\textsuperscript{183} Then, subject to this constraint, the reviewer must determine whether the process claim as a whole is new and useful.\textsuperscript{184}

The respondent argued that this treatment conflated the Section 102 and 103 novelty and usefulness requirements with the Section 101 subject-matter determination.\textsuperscript{185} The Court deflected this claim by observing that certain discoveries such as laws of nature are outside what Congress intended to protect.\textsuperscript{186} Therefore, courts must refuse to allow patentability to be determined solely by non-patentable subject matter.\textsuperscript{187} Whereas inclusion of a law of nature or algorithm does not by itself disqualify a claim from patent protection, neither can it be the sole factor that is new and non-obvious.

In \textit{Flook} the Supreme Court attempted to create limits for software-based creations without violating its self-imposed constraint of not rul-

\begin{itemize}
  \item\textsuperscript{178} See id.
  \item\textsuperscript{179} See id. at 585–86.
  \item\textsuperscript{180} See id.
  \item\textsuperscript{181} See id. at 588 n.9, 594.
  \item\textsuperscript{182} See id. at 591–92.
  \item\textsuperscript{183} See \textit{Flook}, 437 U.S. at 591–92.
  \item\textsuperscript{184} See id. at 591.
  \item\textsuperscript{185} See id. at 592.
  \item\textsuperscript{186} See id. at 593.
  \item\textsuperscript{187} See id.
\end{itemize}
ing on the patentability of computer software. The Court’s ruling, that any software that could be classified as an algorithm could not provide the claim element that met patentability requirements, provided a bulwark against clever claim drafters getting patents that would be rejected if framed differently. Nevertheless, this approach increased rather than reduced ambiguity. Without defining “algorithm,” it was unclear whether the term encompassed all computer programs or just a subset. Nonetheless, Flook was successful in putting limits, albeit for a short time, on the scope of patentable process claims.

The Supreme Court quickly rejected the Flook constraints in 1981 in Diamond v. Diehr. The Diehr process claim consisted of “a process for molding raw, uncured synthetic rubber into cured precision products.” The contribution of this claim was the ability to measure the temperature of the rubber inside the press continually, and to recalculate the Arrhenius equation based on this data in order to determine the precise time to complete the curing process. Applying the Flook test to this claim required evaluating the artificial rubber-curing process under the assumption that the non-patentable Arrhenius equation was well known in the art. Employing this approach in his dissenting opinion, Justice Stevens concluded that the claim contained no innovation other than the continual monitoring of the process, and was thus not statutory subject matter.

The majority of the Supreme Court rejected this analysis based, ironically, on an application of the machine-or-physical-transformation test. The Court reconciled Benson and Flook by characterizing those claims as attempts to obtain patents on mathematical formulae. It distinguished the Diehr claim as a more efficient method for curing rubber, an industrial process of the type that patents were designed to

188 See id. at 595.
189 See Flook, 437 U.S. at 590.
190 See Diehr, 450 U.S. at 219 (Stevens, J., dissenting) (explaining how “the inclusion of the ambiguous concept of an ‘algorithm’ within the ‘law of nature’ category of unpatentable subject matter has given rise to the concern that almost any process might be so described and therefore held unpatentable”).
191 Compare Flook, 437 U.S. at 591–94, with Diehr, 450 U.S. at 192 (rejecting the constraints in Flook).
192 Diehr, 450 U.S. at 192.
193 Id. at 175.
194 Id. at 178–79.
195 See id. at 208–09 (Stevens, J., dissenting).
196 See id. at 209.
197 See id. at 192 (majority opinion).
198 Diehr, 450 U.S. at 185–87.
Critical to the Court’s analysis was that this process, considered in its entirety, transformed the state of a substance.\textsuperscript{199} The Court refused to use physical transformation as a necessary condition for patentability, but did recognize it as a sufficient condition to satisfy the subject-matter requirement.\textsuperscript{200}

Most significantly, the Court rejected the \textit{Flook} approach of not allowing patents for which a non-patentable algorithm was the sole source of innovation.\textsuperscript{201} This reversal made it significantly easier for patent lawyers to draft valid patent claims that were faster or more efficient than—but the same in all other respects as—existing technology.\textsuperscript{202} In a passionate dissent, Justice Stevens, joined by Justices Brennan, Marshall, and Blackmun, castigated the majority opinion for increasing uncertainty and ambiguity in the treatment of patent claims containing computer programs and mathematical algorithms.\textsuperscript{203} Of even greater importance, however, was the \textit{Diehr} dissent’s argument that the majority opinion eviscerated the holdings in \textit{Benson} and \textit{Flook} as well as “the settled line of authority reviewed in those opinions.”\textsuperscript{204} Whereas the Supreme Court in \textit{Benson} and \textit{Flook} attempted to provide some limits to the patentability of software, the dissenting justices in \textit{Diehr} argued that the majority’s approach essentially opened the floodgates to software patents.

Specifically, the dissenting justices saw little substantive difference between the alarm-limits patent claim invalidated in \textit{Flook} and the rubber-curing patent in \textit{Diehr}.\textsuperscript{205} None of the process steps or components were unusual except the application of the Arrhenius algorithm to temperature readings in order to choose an optimal time to end the curing process.\textsuperscript{206} The novelty in this process consisted solely of “updating the original estimated curing time by repetitively recalculating that time pursuant to a well-known mathematical formula in response to variations in temperature within the mold.”\textsuperscript{207} The only noticeable dif-

\begin{footnotes}
\footnotetext[199]{See id. at 184.}
\footnotetext[200]{See id. at 192.}
\footnotetext[201]{See id.}
\footnotetext[202]{See id. at 185–88.}
\footnotetext[203]{See id. at 187–88.}
\footnotetext[204]{See \textit{Diehr}, 450 U.S. at 193 (majority opinion), 219 (Stevens, J., dissenting).}
\footnotetext[205]{\textit{Id.} at 205 (Stevens, J., dissenting).}
\footnotetext[206]{\textit{Id.} at 209 (“Their method of updating the curing time calculation is strikingly reminiscent of the method of updating alarm limits that Dale Flook sought to patent.”).}
\footnotetext[207]{\textit{Id.} at 208 (“There is no suggestion that there is anything novel in the instrumentation of the mold, in actuating a timer when the press is closed, or in automatically opening the press when the computed time expires.”).}
\footnotetext[208]{\textit{Id.} at 209.}
\end{footnotes}
ference between this process and the alarms limit process in *Flook* is that the *Diehr* process automatically opened the rubber mold once designated conditions were made, whereas the alarms limit process claim did not include automatically setting off an alarm.\(^{209}\)

It is hard to reconcile the different results in *Flook* and *Diehr* given the strong similarity between the two cases. *Flook* stood for the proposition that adding insignificant post-solution activity was insufficient to make a mathematical algorithm patentable.\(^{210}\) Yet it is difficult to rationalize *Diehr*'s implicit conclusion that signaling the opening of a rubber mold after obtaining the solution of a mathematical algorithm does constitute a significant post-solution action.\(^{211}\) The use of the Arrhenius algorithm in an unequivocal industrial process clearly influenced the *Diehr* Court.\(^ {212}\) But perhaps the similar post-solution action of generating a signal in both cases makes the industrial application a distinction without substance.\(^ {213}\) With the *Diehr* Court’s “entire process” approach it appeared easier to get patent approval for an algorithm embedded in a useful process.\(^ {214}\) The most significant constraint was that the claim could not encompass all uses of the algorithm.\(^ {215}\) Nevertheless, if a particular industry, such as rubber manufacturing, relies on a particular algorithm, this constraint does not prevent an inventor from preempting the automatic or computerized use of the equation as happened in *Diehr*.\(^ {216}\) Thus, the *Diehr* decision opened the door to broad patent claims that relied on algorithms.

The final, oft-cited Supreme Court patent subject matter case of the twentieth century had nothing to do with computer algorithms or process claims. Nevertheless, the 1980 Supreme Court case *Diamond v. Chakrabarty*\(^ {217}\) is sometimes misquoted as standing for the proposition that U.S. patent law holds no boundaries for human inventiveness.\(^ {218}\) The defendant, Chakrabarty, invented a micro-organism that consumed

---

\(^{209}\) Cf. *Flook*, 437 U.S. at 585–86.

\(^{210}\) See id. at 590.

\(^{211}\) See *Diehr*, 450 U.S. at 184.

\(^{212}\) See id. at 185–88.

\(^{213}\) See id. at 184; *Flook*, 437 U.S. at 594–95.

\(^{214}\) See *Diehr*, 450 U.S. at 188–89.

\(^{215}\) See id. at 188; see also *Benson*, 409 U.S. at 71–73.

\(^{216}\) See *Diehr*, 450 U.S. at 191–93.


\(^{218}\) See, e.g., *State St. Bank & Trust Co. v. Signature Fin. Grp. Inc.*, 149 F.3d 1368, 1373 (Fed. Cir. 1998) (“Indeed, the Supreme Court has acknowledged that Congress intended § 101 to extend to ‘anything under the sun that is made by man.’” (quoting *Chakrabarty*, 447 U.S. at 309)).
crude oil, a trait that presumably would be useful in oil spill cleanups.\textsuperscript{219} The issue was whether or not a micro-organism was disqualified from statutory subject matter because it is a life form; or, alternatively, whether it qualified as statutory subject matter as either a manufacture or as a composition of matter.\textsuperscript{220} The Supreme Court reasoned that the micro-organism, although a life form, was not a product of nature but rather a creation of man, and as such, was statutory subject matter.\textsuperscript{221} The Supreme Court cited Committee Reports published in connection with the adoption of the 1952 Patent Act, which indicated that Congress intended statutory subject matter to “include anything under the sun that is made by man.”\textsuperscript{222}

Although the language is sweeping, the context of the Court’s quote makes it clear that the Court did not intend to apply the language to computer software. The Chakrabarty Court clearly stated that statutory subject-matter scope is not unlimited.\textsuperscript{223} Citing their Flook decision, the Court reiterated that laws of nature, physical phenomena, and abstract ideas are excluded from statutory subject matter.\textsuperscript{224} It is quite telling that the Court cited Flook as the most recent authority for these limitations.\textsuperscript{225} There is no contradiction between the two cases. Both cases purported to deal with new, unforeseen categories of creation: computer programs in Flook\textsuperscript{226} and man-made organisms in Chakrabarty.\textsuperscript{227} Although specific details may have been unforeseen, existing categories covered each of these claims. The Court classified the Flook computer program as a mathematical algorithm—an abstract idea\textsuperscript{228}—and the Chakrabarty micro-organism was either a manufacture or composition of matter.\textsuperscript{229} Thus, the claims’ classifications—either a process or a manufacture or composition of matter—dictated the respective claims’ treatments. A manufacture or composition of matter claim is clearly valid patent subject matter even if the invention had never been anticipated.\textsuperscript{230} Process claims, however, were subject to the more limited

\textsuperscript{219} Chakrabarty, 447 U.S. at 305.
\textsuperscript{220} See id. at 307.
\textsuperscript{221} See id. at 309–10.
\textsuperscript{222} See id. at 309.
\textsuperscript{223} See id. (“This is not to suggest that § 101 has no limits or that it embraces every discovery.”).
\textsuperscript{224} See id.
\textsuperscript{225} See Chakrabarty, 447 U.S. at 309.
\textsuperscript{226} See Flook, 437 U.S. at 585–86.
\textsuperscript{227} See Chakrabarty, 447 U.S. at 305.
\textsuperscript{228} See Flook, 437 U.S. at 594–96.
\textsuperscript{229} See Chakrabarty, 447 U.S. at 309–10.
\textsuperscript{230} See id.
treatment described in the Supreme Court’s *Benson-Flook-Diehr* line of cases.

C. Federal Court Treatment of Computer Software

Federal courts’ expansive treatment of computer and software patents in the computer age contrasts strikingly with the Supreme Court’s cautionary approach.\(^{231}\) Prior to the 1960s, the U.S. judiciary was consistent in limiting the scope of statutory subject matter.\(^{232}\) Nonetheless, the United States Court of Patent Appeals (CCPA) and its successor, the CAFC, delivered a series of decisions that rapidly removed the set of limitations that various courts adopted and followed during the nineteenth century and most of the twentieth century.\(^{233}\) The change can be attributed to a different interpretation of the policy basis behind the Patent Act. With respect to software, courts had interpreted Section 101 narrowly prior to the 1960s.\(^{234}\) Courts interpreted “useful arts” to include processes that were “technological” in nature.\(^{235}\) Federal courts did not question prohibitions against abstract ideas and laws of nature because the potential for such patents to deter progress was clear.\(^{236}\) By contrast, the policy bases for other statutory subject-matter exclusions, such as the mental steps exclusion, were rather opaque.\(^{237}\) Consequently, starting in the 1960s, federal courts—perhaps adopting a more expansive view of statutory subject matter, or perhaps demanding

\(^{231}\) See *Diehr*, 450 U.S. at 205 (Stevens, J., dissenting) (describing the approach of the Court of Customs and Patent Appeals to Section 101 statutory subject matter questions as “expansive”); *In re Musgrave*, 431 F.2d at 893–94 (Baldwin, J., concurring) (describing the majority opinion’s rejection of well-established subject matter limits as “radical”).

\(^{232}\) See *Diehr*, 450 U.S. at 195 (Stevens, J., dissenting) (observing that, “[p]rior to 1968, well-established principles of patent law probably would have prevented the issuance of a valid patent on almost any conceivable computer program”).

\(^{233}\) See *id*.

\(^{234}\) *Id*.

\(^{235}\) See Samuelson, *supra* note 47, at 1112 (quoting professor D. Chisum as stating “the general purpose of the statutory classes of subject matter is to limit patent protection to the field of applied technology, what the U.S. Constitution calls the ‘useful arts,’” and also quoting professor Chisum as noting that as practical and useful as they may be, “discoveries . . . in nontechnological arts, such as the liberal arts, the social sciences, theoretical mathematics, and business and management methodology” are not patentable).

\(^{236}\) See *Benson*, 409 U.S. at 67 (discussing how abstract ideas and laws of nature are excluded from statutory subject matter because “they are the basic tools of scientific and technological work”); *LeRoy*, 55 U.S. at 175 (“A principle, in the abstract, is a fundamental truth; an original cause; a motive; these cannot be patented, as no one can claim in either of them an exclusive right.”).

\(^{237}\) See, e.g., Ambrose, *supra* note 154, at 911.
greater rigor in shaping decisions—reviewed such exclusions critically and rejected them for lacking legislative support.

One of the first rules to fall under this enhanced level of scrutiny was the mental steps exclusion. In 1968, in *In re Prater*, the CCPA considered a claim rejection based on *In re Abrams*’ “three rules” for dealing with mental steps.\(^{238}\) The court noted that the defense attorney’s brief proposed the three rules and that the *Abrams* court had never adopted the three rules for testing claims containing mental steps.\(^ {239}\) The court continued by dismissing the *Cochrane* physical transformation test as dicta.\(^ {240}\) This decision and the CCPA’s reaffirmation of its reasoning in its rehearing of *In re Prater*, put into question the continued viability of the mental steps exclusion.

In 1970, in *In re Musgrave*, the CCPA answered this question by rejecting the mental steps doctrine and its point of novelty approach.\(^ {241}\) The CCPA reiterated its rejection of the *Abrams* mental steps rules.\(^ {242}\) The court continued by specifically rejecting the “point of novelty” analysis incorporated in the *Abrams* rule as “logically unsound.”\(^ {243}\) This rejection was particularly noteworthy because the Supreme Court in *Flook* had recently employed a point of novelty analysis in rejecting a patent claim containing an algorithm.\(^ {244}\) In addition, the CCPA explicitly rejected mental steps as a statutory subject-matter exclusion.\(^ {245}\) The CCPA reasoned that a claim that required subjective judgment would likely be rejected under other Patent Law provisions but would not be excluded from statutory subject-matter treatment.\(^ {246}\)

In dealing with claims that included algorithms, similar to those that the Supreme Court reviewed in *Flook* and *Diehr*,\(^ {247}\) the federal courts found substantial leeway in shaping the law. In 1992, in *Arrhythmia Research Technology v. Corazonix Corp.*,\(^ {248}\) the CAFC—the CCPA’s successor court—faced the task of determining the validity of a mixed patent claim that was remarkably similar to the mixed claims in *Flook*\(^ {249}\)

---

\(^{238}\) See *In re Prater*, 415 F.2d 1378, 1381, 1386 (C.C.P.A. 1968).

\(^{239}\) See id. at 1386.

\(^{240}\) See id. at 1388.

\(^{241}\) See *In re Musgrave*, 431 F.2d at 886.

\(^{242}\) See id. at 889.

\(^{243}\) Id.

\(^{244}\) See *Flook*, 437 U.S. at 594.

\(^{245}\) See *In re Musgrave*, 431 U.S. at 889, 893.

\(^{246}\) Id. at 893.

\(^{247}\) *Diehr*, 450 U.S. at 181, 185; *Flook*, 437 U.S. at 587.


\(^{249}\) 437 U.S. at 585–86.
The patent claim at issue in Arrhythmia Research dealt with a medical problem related to heart attack victims, who are at high risk of suffering from ventricular tachycardia, a condition which can lead to a large diminution in the flow of blood from the heart. Although drugs are effective in treating the condition, these drugs have serious side effects and optimally should be taken only when absolutely needed. Based on well-known relationships between patients’ electrocardiographic signals and heart conditions, treating physicians could identify patients who were at particularly high risk for ventricular tachycardia. Specifically, patients with certain anomalous wave characteristics in the ventricular contraction cycle—referred to as “late potentials”—were at high risk. The Stinson patent claim’s innovation was its ability to filter and isolate late potentials present in electrocardiograph readings and set off an alarm if the late potentials exceeded a specified threshold. In summary, the invention consisted of reading electrocardiograph signals, converting the reading to a data format readable by a generic computer, and analyzing the data using mathematical algorithms for the presence of late potentials by comparing the value of analyzed data to a predetermined level.

The Arrhythmia Research process claim was analogous to the alarms-rate process claim in Flook and the rubber-curing process claim in Diehr. In all three cases, the point of novelty resided entirely in the data-processing component of the claim. The electrocardiograph readings that indicate the presence of late potentials were well known, as were the threshold levels that warranted initiation of patient treatment. The novelty in the Stinson patent claim, then, was processing a noisy digital signal to produce a reading that provided a more accurate measure of the level of late potentials. Analogous to the Flook alarms-limit claim, the Stinson claim accepted input data, evaluated it using...
mathematical algorithms and provided a reading of the target variable. Unlike the rubber-curing claim in *Diehr*, the Stinson process was not part of a traditional industrial process that produced a physical end product. Thus, there was no significant difference between the Stinson process and algorithm-based process claims in *Benson* and *Flook* that the Supreme Court held were not statutory subject matter.

Nonetheless, the CAFC held that the Stinson claim constituted statutory subject matter. In applying the physical transformation test, the CAFC erroneously concluded that the Stinson process transformed matter from one state to another. In referring to the electrocardiograph signals that provided input data for the analysis, the CAFC mistakenly asserted that such signals were not abstractions, but were “related to the patient’s heart function.” But such signals are absolutely abstractions. Although related to the patient’s heart function, these signals are simply a measure of that function, just as a digital thermometer measures temperature. Once translated into digital form, the electrocardiograph signal is much more akin to financial data in a spreadsheet than it is to the function of a patient’s heart. Just as the information in a spreadsheet provides an abstraction of financial information, the electrocardiograph signal provides an abstract representation of the patient’s heart function. The CAFC also concluded that manipulation of electrical signals satisfied the physical transformation test.

Thus, according to the CAFC, the act of converting readings from a measuring device into a form that a generic digital computer could recognize also satisfies the physical transformation test.

Judge Rader’s concurrence in *Arrhythmia Research* provided a somewhat prescient insight into the direction that the CAFC was headed. Rather than accept the tortured attempts of the *Arrhythmia Research*
majority to stretch the physical transformation test to cover the Stinson process, Judge Rader advocated dropping all subject-matter tests that had no statutory basis in Section 101 of the Patent Act. Judge Rader noted that most of the tests relied on “vague and malleable terms” such as “law of nature,” “natural phenomena,” “formulae,” or “algorithm.” He noted that “[w]hen attempting to enforce a legal standard embodied in broad, vague, non statutory terms, the courts have floundered.” Judge Rader noted the specific difficulty courts had in interpreting and applying the prohibition against patenting mathematical algorithms. He identified two 1982 CCPA cases that had dramatically different interpretations of that term. In In re Pardo, the CCPA narrowly defined mathematical algorithm, whereas in In re Meyer, the CCPA broadly defined mathematical algorithm, “to include any mental process that can be represented by a mathematical algorithm.” In any event, Judge Rader argued that Diehr had already jettisoned the mathematical algorithm exclusion, leaving laws of nature, natural phenomenon, and abstract ideas as the only non-statutory subject-matter exclusions. Therefore, courts should look to the plain meaning of Section 101 to determine valid statutory subject matter. The CAFC soon followed Judge Rader’s exhortation.

In In re Alappat, the CAFC began moving towards the complete elimination of statutory subject-matter tests by jettisoning the Supreme Court’s physical-transformation test. The In re Alappat claim was a computer program that could run on any conventional digital computer, known as a rasterizer. A waveform data sequence provided the input, which the rasterizer processed and filtered to produce output data that was amenable to display on a cathode ray tube. Consistent with Judge Rader’s concurrence in Arrhythmia Research, the CAFC reiterated that Diehr recognized laws of nature, natural phenomena, and ab-

269 See Arrhythmia Research, 958 F.2d at 1066 (Rader, J., concurring).
270 Id. at 1062–63.
271 Id. at 1063.
272 Id.
273 Id.
274 In re Pardo, 684 F.2d 912, 916–17 (C.C.P.A. 1982).
275 In re Meyer, 688 F.2d 789, 796 (C.C.P.A. 1982).
276 Arrhythmia Research, 958 F.2d at 1063.
277 Id. at 1066.
278 See id.
280 See In re Alappat, 33 F.3d 1526, 1544 (Fed.Cir. 1994) (en banc).
281 Id. at 1537.
abstract ideas as the only exclusions from statutory subject matter. The CAFC then concluded that a claim that includes software is non-statutory only to the extent that the claim as a whole “represent[s] nothing more than abstract ideas.” Thus, unless a claim was unequivocally a mathematical algorithm, it qualified as statutory subject matter. To satisfy this standard for statutory subject matter, the inventor need only show that the claim produced “a useful, concrete, and tangible result.” The rasterizer in In re Alappat clearly satisfied this standard.

In re Alappat’s “useful, concrete and tangible” test removed all remaining ambiguity as to whether claims must satisfy the Cochrane physical transformation test to be considered statutory subject matter. Although the BCD claim in Benson would likely fail the useful, concrete and tangible test, the alarm-limits claim in Flook might satisfy the test based on In re Alappat’s broad definition of “concrete and tangible.” In addition, In re Alappat made it easier for clever claims drafters to frame software programs employed in conventional digital computers as machines. Drafting a claim as a machine rather than a process makes satisfying the concrete and tangible requirement trivial.

AT&T Corp. v. Excel Communications, Inc. and State Street Bank & Trust Co. v. Signature Financial Group, Inc. removed all remaining Section 101 impediments to software patents. The State Street Bank claim was a computerized accounting system used to allocate returns for mutual fund shareholders. The AT&T patent identified a method for recording certain information about long-distance telephone callers that was useful for billing purposes. These cases went a step beyond Alappat in that the subject patent claims consisted entirely of business applications. Prior to State Street Bank, it was generally accepted that meth-

282 Id. at 1542.
283 Id. at 1543 (emphasis omitted).
284 Id. at 1544 (“[T]he proper inquiry . . . is to see whether the claimed subject matter as a whole is a disembodied mathematical concept . . . which in essence represents nothing more than a ‘law of nature,’ ‘natural phenomenon,’ or ‘abstract idea.’”)(emphasis omitted).
285 Id.
286 In re Alappat, 33 F.3d at 1544.
287 Cochrane, 94 U.S. at 787–88; see In re Alappat, 33 F.3d at 1544.
288 See Flook, 437 U.S. at 585–86, 591; Benson, 409 U.S. at 72; In re Alappat, 33 F.3d at 1544–45.
289 See In re Alappat, 33 F.3d at 1545.
290 See AT&T, 172 F.3d at 1359–61; State St. Bank, 149 F.3d at 1375–76.
291 State St. Bank, 149 F.3d at 1370.
292 AT&T, 172 F.3d at 1353.
293 See id.; State St. Bank, 149 F.3d at 1370.
ods for conducting business were not valid subject matter for patents: this exclusion was known as the business method exception. *State Street Bank* rejected the business method exception, reasoning that reliance on Section 101 and patent law in general are sufficient to evaluate such claims.\(^{294}\) Instead, *State Street Bank* employed *In re Alappat*’s useful, concrete and tangible result test to evaluate the mutual fund accounting method as constituting statutory subject matter.\(^{295}\)

*AT&T* explicitly extinguished *Diehr*’s physical-transformation test. Excel Corporation based its defense to AT&T’s infringement claim on the patent’s failure to effect a physical transformation.\(^{296}\) Ignoring the physical transformation discussion in *Diehr*,\(^{297}\) the CAFC instead focused on the use of the term “e.g.” in *Diehr* to infer that physical transformation was only one of multiple ways software could satisfy Section 101 statutory subject-matter requirements.\(^{298}\) The court then proclaimed that the useful, concrete and tangible result test had supplanted the physical-transformation test, stating “[w]hatever may be left of the earlier test, if anything, this type of physical limitations analysis seems of little value.”\(^{299}\) Thus, software and business methods no longer receive special statutory subject-matter scrutiny. Any software claim drafted in terms of a process, machine, manufacture, or composition of matter may overcome the Section 101 hurdle.\(^{300}\)

Support for the CAFC’s expansion of statutory subject matter was not universal.\(^{301}\) Whereas Congress and the Supreme Court stayed on the sidelines, the United States Board of Patent Appeals (BPA) attempted to reign in patent coverage of business method patents. In 2001 in *Ex parte Bowman*, the BPA considered an invention that claimed a “method of evaluating an intangible asset of interest.”\(^{302}\) A representative claim consisted of choosing relevant variables, plotting the variables on a two-dimensional chart and then using the chart to ascertain

\(^{294}\) *State St. Bank*, 149 F.3d at 1375–76.

\(^{295}\) Id.

\(^{296}\) *AT&T*, 172 F.3d at 1358.

\(^{297}\) *Diehr*, 450 U.S. at 182–84.

\(^{298}\) *AT&T*, 172 F.3d at 1358–59.

\(^{299}\) Id. at 1359.

\(^{300}\) Id. at 1359–61.


the value of the asset.\textsuperscript{303} Although it might appear that such an invention would be subject to a Section 112 challenge for being vague, the BPA ruled that the Bowman claims met the Section 112 disclosure requirement.\textsuperscript{304} Instead, the BPA rejected the Bowman claims because the invention was not tied to any technological art and therefore “is nothing more than an abstract idea.”\textsuperscript{305} The BPA opined that the term “technological arts” was synonymous with the term “useful arts” that appears in the Patent Clause of the Constitution.\textsuperscript{306}

The technological arts requirement did not last. In 2005, in \textit{Ex parte Lundgren}, the BPA, in determining whether “a method of compensating a manager” is statutory subject matter, considered applying two tests: the technological arts test and the “useful, concrete and tangible” test.\textsuperscript{307} The BPA rejected the technological arts requirement, noting that \textit{Ex parte Bowman} lacked precedential value, and the Board found no valid support for the requirement.\textsuperscript{308} The BPA then concluded that the invention was statutory subject matter because it satisfied the three requirements of the useful, concrete, and practical test.\textsuperscript{309}

Two BPA judges disagreed with this decision. Judge Smith interpreted the constitutional mandate for patents as restricted to inventions associated with science or technology.\textsuperscript{310} Although he did not champion the technological arts test, Judge Smith noted that this test at least required an invention to be tied to a computer.\textsuperscript{311} He further noted that the method for compensating managers was a method that could have been executed at the time of the drafting of the Constitution.\textsuperscript{312} Judge Smith could not imagine that the drafters would have considered such an invention to be within the bounds of what they considered to be statutory subject matter.\textsuperscript{313} Judge Barrett, in agreeing with the majority that the technological arts test lacked validity, nevertheless disagreed that the \textit{Lundgren} invention was statutory subject mat-

\begin{itemize}
  \item \textsuperscript{303} Id.
  \item \textsuperscript{304} Id.; see \textsection 35 U.S.C. \textsection 112 (2006) (requiring patent claims to be sufficiently detailed in order to enable someone skilled in the art to replicate the invention).
  \item \textsuperscript{305} \textit{Ex parte Bowman}, 2001 WL 1646047 at *3.
  \item \textsuperscript{306} See id.
  \item \textsuperscript{308} Id. at 1387.
  \item \textsuperscript{309} Id. at 1386.
  \item \textsuperscript{310} Id. at 1388 (Smith, J., dissenting).
  \item \textsuperscript{311} Id.
  \item \textsuperscript{312} Id.
  \item \textsuperscript{313} \textit{Ex parte Lundgren}, 76 U.S.P.Q.2d at 1388 (Smith, J., dissenting).
\end{itemize}
In reaching this conclusion, Judge Barrett embraced the previously discarded physical transformation test. In essence, Judge Barrett agreed with Judge Smith that a disemboweled invention that could be executed separately and independently of any machine was not statutory subject matter. Thus, the two judges would not validate any business method patent claim that was not computer implemented. The judges’ position corresponds to the European approach under the EPC. In the parlance of the EPO, a business method not implemented on a computer would be a business method “as such,” and thus fail the technicality requirement of Article 52.

D. The Bilski Case

In 2008 the CAFC took up the statutory subject-matter question once again in In re Bilski. The CAFC’s Bilski decision was likely influenced by dicta in two 2006 Supreme Court decisions that expressed doubts about whether CAFC rulings with respect to business method patents were correct. In eBay, Inc. v. MercExchange, L.L.C., Justices Kennedy, Stevens, Souter, and Breyer voiced concern about granting permanent injunctive relief in business-method-patent infringement cases. Their rationale was that such patents are potentially vague and of “suspect validity.” In LabCorp v. Metabolite Laboratories, three justices questioned the validity of the State Street Bank holding. In a dissenting opinion, Justices Breyer, Stevens, and Souter flatly rejected the CAFC’s useful, concrete, and practical test because “[i]f taken literally, the statement would cover instances where this Court has held to the contrary.” Thus, the CAFC appeared to have a mandate to curtail patent coverage of business method patents.

The CAFC accepted this mandate in In re Bilski. The Bilski invention covered a method of hedging risk in commodity trading.

---

314 Id. at 1389 (Barrett, J., dissenting).
315 Id.
316 Id.
317 See EPC 2000, supra note 14, art. 52.
320 Id.
322 Id.
323 See In re Bilski, 545 F.3d at 961.
324 Id. at 949.
the *Bowman*\textsuperscript{325} and *Lundgren*\textsuperscript{326} claims, the *Bilski* invention did not require a computer for execution. It “merely manipulates [an] abstract idea and solves a purely mathematical problem.”\textsuperscript{327} The CAFC identified its task in examining claims that included mathematical algorithms to be that of determining whether the claimed use “would pre-empt all uses of that fundamental principle.”\textsuperscript{328} The CAFC considered this task daunting due to the complexity of modern claims.\textsuperscript{329} The CAFC, however, recognized that the machine or physical transformation test discussed in *Diehr* and earlier cases would be effective in rejecting over-reaching patent claims.\textsuperscript{330} Moreover, the CAFC determined that the machine or physical transformation test was valid law despite the Supreme Court’s refusal to explicitly adopt the test in *Benson*, *Flook*, and *Diehr*.\textsuperscript{331} Thus, *In re Bilski* wiped clean several decades of CAFC and CCPA jurisprudence while taking a step that the Supreme Court seemed reluctant to take. On the appeal for *In re Bilski*, the Supreme Court would respond to the CAFC’s reasoning.

On June 28, 2010, nearly eight months after hearing oral arguments in *Bilski v. Kappos*,\textsuperscript{332} the Supreme Court issued its long-awaited decision on the last day of its term. Although the justices were unanimous in finding Bilski’s invention not patentable, they divided on the legal reasoning and where to draw the line concerning the patentability of business methods. Five justices reasoned that Section 101 does not preclude the patentability of all methods of doing business, stating that “a business method is simply one kind of ‘method’ that is, at least in some circumstances, eligible for patenting.”\textsuperscript{333} Four justices concluded that “although a process is not patent-ineligible simply because it is useful for conducting business, a claim that merely describes a method of doing business” is not patentable.\textsuperscript{334}

The majority—while not rejecting the CAFC’s revitalized machine-or-physical-transformation test—refused to endorse the CAFC’s effort to resolve the subject-matter question, and ultimately failed to articulate

\textsuperscript{325}Ex parte Bowman, 2001 WL 1646047 at *2.
\textsuperscript{326}Ex parte Lundgren, 76 U.S.P.Q.2d at 1387.
\textsuperscript{327}In re Bilski, 545 F.3d at 950.
\textsuperscript{328}Id. at 954.
\textsuperscript{329}See id.
\textsuperscript{330}See id.
\textsuperscript{331}See Diehr, 450 U.S. at 189–90, 191–92; Flook, 437 U.S. at 589–90; Benson, 409 U.S. at 71.
\textsuperscript{332}Bilski v. Kappos, 130 S. Ct. 3218 (2010).
\textsuperscript{333}Id. at 3228.
\textsuperscript{334}Id. at 3232 (Stevens, J., concurring).
Rather, the Court returned responsibility to the CAFC to develop “less extreme means” than the machine or physical-transformation test to limit business method patentability. The Court advised the CAFC to try a variety of approaches “including (but not limited to) application of our decisions in Benson, Flook, and Diehr.”

Such approaches cannot rely solely on the machine or physical-transformation test, however. The Court admonished the CAFC that the test “is not the sole test for deciding whether an invention is a patent-eligible ‘process.’” Tracking earlier Supreme Court patent subject-matter jurisprudence, the Court asserted that the test “is a useful and important clue, an investigative tool, for determining whether some claimed inventions are processes under §101.” The Court went on to explain that while it may have been sufficient for evaluating processes “similar to those in the Industrial Age,” they doubted its usefulness for the “Information Age” in which “new technologies may call for new inquiries.” Therefore, rather than approving the CAFC’s gutsy attempt to curtail business method patent claims—or taking a similarly bold step to provide clarity—the Court cautiously relied on its Benson, Flook, and Diehr decisions to reject Bilski’s claims as unpatentable abstract ideas. Without articulating a clear test, the Court concluded that “[t]he concept of hedging, described in claim 1 and reduced to a mathematical formula in claim 4, is an unpatentable abstract idea, just like the algorithms at issue in Benson and Flook.”

Justice Stevens and the concurring justices would have drawn a clearer line in the sand. Relying on interpretation of “process” as a term of art anchored in historical practice, the concurrence concluded that “[a] business method is not a ‘process’” under Section 101, and therefore not patentable. Justice Breyer, while agreeing with Stevens that business methods are not patentable processes, wrote a separate

---

335 See id. at 3226–27, 3231 (majority opinion).
336 Id. at 3231.
337 Id. at 3227.
338 Bilski, 130 S. Ct. at 3227.
339 Id.
340 Id.
341 Id. at 3227–28.
342 See id. at 3229–30.
343 Id. at 3231.
344 Bilski, 130 S. Ct. at 3249–50 (Stevens, J., concurring).
concurrency to “highlight the substantial agreement among the many Members of the Court.”

Nevertheless, the case narrows the scope of patentable subject matter from the “useful, concrete and tangible result” test articulated in State Street Bank. Rather than completely rejecting the machine or physical transformation test, all three opinions agree that the test has been a “useful and important clue” but not the “sole test” for determining patentability. This previously discarded test has thus regained the stature it had under Benson, Flook, and Diehr. Although Justice Kennedy’s opinion does not expressly reject the useful, concrete, and tangible test, Kennedy’s opinion indirectly rejected past CAFC patent process subject-matter decisions. Indeed, five justices in the two concurring opinions went even further by explicitly rejecting the State Street Bank standard.

Following Bilski v. Kappos, U.S. patent process subject-matter law has thus returned to the 1980s, perhaps with a revitalized exclusion for abstract ideas. Any mixed process claim that does not solely claim a law of nature, physical transformation, or abstract idea meets the initial patent subject-matter threshold. Until the CAFC provides additional guidance, however, the USPTO must apply the amorphous standards of Benson, Flook, and Diehr. If the claim provides a physical transformation of matter, then the decisions agree that the claim is patentable. Otherwise, it is unclear whether and under what conditions process claims will escape the abstract idea exclusion. Regardless, a claim to a business method patent must also meet the other patent requirements of novelty and non-obviousness. Thus, as the majority opinion obliquely suggests, the CAFC may shift the critical patentability

---

345 Id. at 3257–58 (Breyer, J., concurring).
346 Compare id. at 3227 (majority opinion), with State St. Bank, 149 F.3d at 1373.
347 Bilski, 130 S. Ct. at 3227, 3232 (Stevens, J., concurring), 3258 (Breyer, J. concurring).
348 Id. at 3221 (majority opinion) (stating that the opinion should not “be read as endorsing interpretations of § 101 that the Court of Appeals for the Federal Circuit has used in the past”).
349 See id. at 3232 n.1 (Stevens, J., concurring) (“[I]t would be a grave mistake to assume that anything with a ‘useful, concrete and tangible result’ . . . is patentable.”) (citation omitted) (quoting State St. Bank, 149 F.3d at 1368), 3259 (Breyer, J., concurring) (“[I]f taken literally, the statement [that anything which produces a useful, concrete, and tangible result is patentable] would cover instances where this Court has held the contrary.”).
350 Id. at 3225 (majority opinion).
351 See Benson, 409 U.S. at 71–72.
352 See Flook, 437 U.S. at 594–95.
353 See Diehr, 450 U.S. at 191–92.
354 Id. at 192; Flook, 437 U.S. at 594; Benson, 409 U.S. at 71.
355 35 U.S.C. §§ 102, 103; see Diehr, 450 U.S. at 191.
standard to these tests rather than to the subject-matter requirement of Section 101.\textsuperscript{356}

III. THE EUROPEAN PATENT OFFICE TREATMENT OF SOFTWARE PATENTS

Direct comparison of U.S. and European patent law in the area of software and business method patents is difficult for both systemic and substantive reasons. The systematic reasons involve the lack of jurisprudential uniformity in Europe. Within Europe, the EPO as well as the national patent offices and courts of the contracting states interpret and apply the EPC.\textsuperscript{357} Each nation, in transposing EPC requirements into national legislation, altered its own substantive and procedural law to approximate its laws to the EPC.\textsuperscript{358} Although this has harmonized the law to some extent, national patent offices and courts can and do interpret the requirements of the EPC differently than the EPO and other national courts.\textsuperscript{359}

The substantive reasons for difficulty involve the “technical” requirement for patentability that exists at the heart of European patent law.\textsuperscript{360} U.S. patent law does not have a direct corollary. The exact nature of the technical contribution requirement as it relates to computer programs and business methods is chimerical. Not only have the tests for technicality been a moving target, often conflicting within and among jurisdictions, but the EPO has never provided a clear and workable legal definition of “technical” as it relates to software and business method patents.\textsuperscript{361}

\textsuperscript{356} See Bilski, 130 S. Ct. at 3229–31. This is similar to the approach taken by the EPO in determining whether business methods and software patent claims should be granted. See Keith Beresford, Patenting Software Under the European Patent Convention 116 (2000); European Patent Office, Patents for Software? European Law and Practice 3 (2009) [hereinafter Patents for Software?].

\textsuperscript{357} See infra notes 679–682 and accompanying text.

\textsuperscript{358} See, e.g., infra Part III.A.

\textsuperscript{359} See infra Part IV.C. Failure to separate questions of patentability from those of enforcement of granted patent presents a weakness in the current debate over software patents in Europe. Andreas Grosche, Software Patent—Boon or Bane for Europe? 14 Inter. J. L. & Info. Tech. 257, 269 (2006). Nevertheless, the possibility of differing interpretations by patent granting offices and by the courts tasked with determining the validity of European patents figure largely in the European problem. See infra Parts IV.C.3 and V.B.


\textsuperscript{361} See Greg Aharonian, Why All Business Methods Achieve a Technical Effect?, Internet Patent News Service (Oct. 2001), http://www.bustpatents.com/aharonian/bzmtdtch.htm (citing an EPO decision from September 2000, which recognizes that the meaning of the term “technical” is not very clear). Aharonian subsequently quotes a private communication between himself and an EPO official on the meaning of “technical”:.
In this Part, we begin with an introduction to the European patent system to highlight the systemic difficulties in creating a unified law on software and business method patents. We then examine the evolution and present state of patentability of computer programs and business methods in the European Patent Office. Part IV then addresses the United Kingdom. The United Kingdom serves as an apt European comparative jurisdiction because of its recent history of patent practice and EPC interpretation that is at odds with the EPO. Part IV also includes a brief analysis of Germany’s patent practice in this area due to Germany’s significance as the largest European Union state.

A. The European Patent System

In the United States, the USPTO examines applications and decides whether to grant a patent. Only federal courts decide issues of patent infringement and validity and only the CAFC hears patent appeals, with the Supreme Court as the final arbiter of patent law. As a result, patent law is arguably one of the more unified areas of law in the United States because there exists no possibility of conflicting interpretations of patent law among the circuits. The EU experience of patent law over the last half century is fundamentally different. Unlike copyright and trademark law, the European Union has no directive or other legal instrument harmonizing substantive patent law across the member states. Although pan-European agreements have developed for patents, these agreements, including the EPC, are not instruments of
the European Union. Although each EU Member State is a contracting party to the EPC, it is a *sui generis* European convention developed from efforts to harmonize both European and EU patent law. Within the European Union it has been characterized as a “stop-gap approach to a more far reaching harmonisation.”

Currently, a person wishing to protect an invention in a particular European Union member state has the option of filing an application for a national patent or for a European patent. For the former, the applicant files with the national patent office in the member state. For the latter, application is made to EPO under the terms of the EPC, which provides a mechanism for obtaining a “European patent” using one central application procedure. The applicant then designates the EPO contracting states in which he or she wants patent protection.

Once the EPO grants the European patent, the patent holder must register it in the appropriate contracting states, where it receives the same rights that would be conferred by a patent granted in that country. The European patent becomes a “bundle of national rights.”

---


373 EPC 2000, *supra* note 14, art. 75.

374 See Judge, *supra* note 371, at 28–1. The designation may include one or more of the 27 European Union contracting states, eight non-European Union EPO members and/or the three EPO extension states. See Member States of the European Patent Organization, EUROPEAN PATENT OFFICE, http://www.epo.org/about-us/epo/member-states.html (last updated Mar. 3, 2011). In addition to the 27 EU Member states, Albania, Croatia, Iceland, Lichtenstein, Macedonia, Monaco, Norway, San Marino, Switzerland and Turkey are contracting states and it is possible to extend patent protection to Bosnia & Herzegovina, Montenegro, and Serbia. *Id.*

375 EPC 2000, *supra* note 14, art. 64.

376 Grosche, *supra* note 359, at 269.
Theoretically, there is no difference between a European patent and one granted by a national patent office, with the exception that the European patent may be subjected to a post-grant opposition procedure at the EPO in Munich.\(^{377}\) Whether the patent is granted by the EPO or a national patent office, the patent holder must enforce it in the national courts in the countries in which the patent is registered.\(^{378}\) National courts have the power to revoke a European patent on a number of grounds, including invalidity, but the EPO does not have jurisdiction to revoke a nationally granted patent.\(^{379}\) National courts interpreting their substantive and procedural rules are not obligated to follow EPO cases as precedents.\(^{380}\) Likewise, national court decisions have no direct consequence on EPO decisions, although EPO cases have considered national decisions and vice versa.\(^{381}\) A major implication of this structure is that patent application, validity opposition, and litigation practice requires very sophisticated planning. It also opens the possibility of differing interpretations of the law and patentability standards among the various entities charged with interpreting and applying the law.

Technical Boards of Appeal within the EPO interpret the EPC through its case law.\(^{382}\) If an applicant is dissatisfied with a non-award, or if a third party wishes to oppose the grant of a patent, that party may appeal to the EPO Technical Board of Appeal.\(^{383}\) The Technical Boards of Appeal are highly centralized as to subject matter and expertise, with the same board deciding cases within a given technical area.\(^{384}\) Technical Board of Appeal 3.5.01 decides cases involving computer programs

\(^{377}\) Judge, supra note 371, at 28–1. The opposition must be commenced within nine months of the grant. Id. An opposition may last for five years or more and may result in the revocation of the patent or an amendment to the claims defining the scope of the patent. Id. There is no similar opposition procedure in U.S. patent practice. Cf. Hill & Cangolosi, supra note 366, § 5.2 (discussing the lack of role the USPTO plays in infringement litigation).

\(^{378}\) EPC 2000, supra note 14, art. 64(3) (“[I]nfringement . . . shall be dealt with by national law.”).

\(^{379}\) See Willem A. Hoyng, Global Patent Litigation (Strategy) § 1, in Global Patent Litigation, supra note 366; Leith, supra note 370, at 103–04.

\(^{380}\) See Leith, supra note 370, at 104–07.

\(^{381}\) See id. at 111–13, 167, 169–70, 177.

\(^{382}\) See id. at 112–14.

\(^{383}\) EPC 2000, supra note 14, arts. 107, 110.

and business methods as computer-implemented inventions. Technical Boards of Appeal have “interpretive supremacy” for the EPC. In the majority of cases, a particular Technical Board functions as the body of final appeal on matters of patent law. In narrowly defined circumstances, a legal question may go to the Enlarged Board of Appeal “to ensure uniform application of the law, or if a point of law of fundamental importance arises.” The President of the EPO or a Technical Board of Appeal on its own motion or in granting a party’s request, may refer questions on points of law to an Enlarged Board of Appeal. The EPC substantially constrains the nature of questions the Enlarged Board of Appeal may decide. It may only review Technical Boards of Appeal decisions in narrowly defined circumstances and only on narrowly defined legal issues. Therefore, the Enlarged Board of Appeal does not function in the same manner as the CAFC or the U.S. Supreme Court.

These limitations mean that the Enlarged Board has provided little guidance for subject-matter questions. To date, despite a recommendation from an English judge, a request for a referral from a party to a Technical Board of Appeal case, and a referral from the EPO Presi-

---

385 See id.; see also G06: Computing; Calculating; Counting World Intell. Prop. Org. 8, 14–15 (Jan. 2010), http://www.wipo.int/itpc/itpc/TSupport_and_download_area/20100101/pdf/scheme/advanced/en/g06.pdf (Technical Board of Appeal 3.5.01 has jurisdiction over patents with main international patent classifications G06F17 and G06Q, which cover digital computing methods and data processing and other methods adapted for business purposes). The EPO uses the terminology “computer-implemented invention” (CII), defined as “an invention whose implementation involves the use of a computer, computer network or other programmable apparatus, the invention having one or more features which are realised wholly or partly by means of a computer program.” PATENTS FOR SOFTWARE?, supra note 356, at 3. The EPO explains the term “software” is too ambiguous. Id. Recently, however, CII has become a euphemism for the more controversial “software patent.” Andrés Guadamuz González, The Software Patent Debate, 1 J. Intell. Prop. L. & Prac. 196, 198–99 (2006).


387 See id. at 23.

388 EPC 2000, supra note 14, art. 112(1).

389 Id. The Enlarged Board may also initiate a referral. See Leith, supra note 370, at 113–114.

390 See EPC 2000, supra note 14, art. 112(1).

391 See Aerotel Ltd. v. Telco Holdings Ltd., [2006] EWCA (Civ) 1371, [75–76], [2007] 1 All E.R. 225 (A.C.) at 236 (Eng.).

the Enlarged Board of Appeal has not directly addressed questions on law relating to computer software. On May 12, 2010, after eighteen months of deliberation, the Enlarged Board of Appeal declined to rule on the questions referred by the EPO President. The Enlarged Board rejected the referral on grounds that the questions were inadmissible or not subject to review because they did not meet the narrow requirements for review under the EPC. The Enlarged Board concluded that the EPC allows a review only when it is necessary to ensure uniform application of the law and when two Technical Boards of Appeal have given “different decisions” on the question referred.

After reviewing the relevant case law, the Enlarged Board determined that the cases identified as “different” did not meet EPC requirements. The Enlarged Board interpreted the EPC as requiring a “conflict in the case law making it difficult if not impossible for the Office to bring its patent granting practice into line with the case law of the Boards of Appeal.” The implications of this “non-ruling” are immense. The Technical Board of Appeal 3.5.01 remains the body of last resort within the EPO for decisions concerning computer programs and business methods and determines examining practice within the EPO. Once a Technical Board of Appeal or Enlarged Board has decided a case, there is no recourse for review to a national or multinational court. Thus, it is unlikely that any judicial or quasi-judicial European entity will soon provide any software subject-matter guidance.

One objective of the EPC is for patent examination and patentability decisions to produce the same result whether the patent is examined or opposed in the EPO, a national patent office in France or Germany, or in an invalidity action in the United Kingdom or the Netherlands.
To facilitate harmonization of patent law through EPC implementation, national examining offices and courts do not possess broad discretion to interpret claims in accordance with what had been national practice prior to the EPC. Instead, they should interpret according to EPO practice.\textsuperscript{401} Nevertheless, the structure of the system has precluded complete harmonization. Variances between approaches to EPC substantive law in national patent offices, which generally rule in accordance with national decisions, mean that substantive differences remain.\textsuperscript{402} Under the current system, it is possible that an application for the same invention may be granted by the EPO or the German patent office but denied by the U.K. Intellectual Property Office, and a claim for invalidity of the same European patent may be interpreted differently in different countries.\textsuperscript{403} These factors contribute to a lack of uniformity in the substantive law.

Furthermore, the absence of a multinational reviewing court to rule on inconsistencies between national courts and the EPO interpretations has impeded harmonization. Because national courts decide cases within their own legal traditions, they interpret provisions of the EPC differently, as implemented into their national law. Thus a patent may be revoked in one country that remains valid in another.\textsuperscript{404} This has led to particular problems in the area of computer software and business methods. In difficult cases, national courts applying national law do not decide consistently.\textsuperscript{405} Furthermore, because the courts in the United Kingdom are required to follow their precedents,\textsuperscript{406} and the EPO does not recognize \textit{stare decisis}, the possibility of differing results is magnified.

\textbf{B. Origins of the “Technical” Requirement}

The U.S. patent statute and EPC take differing approaches to patentability. U.S. law defines patentable inventions \textit{positively}, as “any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof,”\textsuperscript{407} and leaves exclusions...
to judicial interpretation. The EPC—while setting out the general requirements of novelty, inventive step, and industrial application—defines invention negatively, according to what is excluded. Article 52 presents a non-exclusive list of things not considered inventions, including: “discoveries, scientific theories and mathematical methods”; “aesthetic creations”; and “schemes, rules and methods for performing mental acts, playing games or doing business, and programs for computers.” The EPO’s “Guidelines for Examination” indicate that the items listed in Article 52(2) are those that are either abstract or do not have a technical character. The list of excluded non-inventions, however, must be read in conjunction with Article 52(3), which narrows the exclusion of these items and activities to the extent that the patent application relates to excluded subject-matter “as such.” This means that a claim to a business method or computer program is not allowed, but “claims to physical entities or processes relating to such items may be allowable.” The exclusions and their limitations are understood to reinforce the requirement that “invention” means technical invention.

“Technical character” does not present a problem in traditional fields such as chemistry or engineering. The requirement resides deep in the origins of European patent law and has been “part of the European legal tradition since the early days of the patent system.”

According to the EPO, in order to be patentable, the subject matter

---

408 EPC 2000, supra note 14, arts. 52(1), 54.
409 Id. arts. 52(1), 56. Under U.S. law, this requirement is known as the “non-obvious” condition for patentability. 35 U.S.C. § 103.
410 EPC 2000, supra note 14, arts. 52(1), 57.
411 Id. art. 52(2)(a).
412 Id. art. 52(2)(b).
413 Id. art. 52(2)(c).
415 EPC 2000, supra note 14, art. 52(3).
416 CIPA, supra note 400, § 3.4.
417 Tritton et al., supra note 23, at 91.
418 Aloys Hüttermann & Ulrich Storz, A Comparison Between Biotech and Software Related Patents, 31 Eur. Intell. Prop. REV. 589, 589 (2009) (explaining that circumstances in which chemical compounds may be patented were resolved decades ago, but the technicality of biotech and software-related inventions are still in flux); see also Beresford, supra note 356, at v–vi (explaining that Boards of Appeal cases reveal that some patents are rejected based on the steps performed by an operator regardless of technical character); Grosche, supra note 359, at 271 (explaining that whether software makes a “technical contribution” turns on how this term of art is defined).
must have “a ‘technical character’ or, to be more precise, involve a ‘technical teaching.’”

Discussions of the technical requirement were part of the earliest European harmonization efforts. The Strasbourg Convention, the Council of Europe instrument that began the process of harmonization of European patent law, did not mention a technical requirement or expressly define the term “invention.” It required signatory countries to grant patents “for any inventions which are susceptible of industrial application, which are new and which involve an inventive step.” When the Strasbourg Convention was adopted in 1963, the six members of the European Economic Community—the precursor to the European Union—were working toward both a single patent legal system for the common market and a single system for granting patents. During the first phase of the negotiations from 1961 to 1964, the delegations debated whether to follow the language of the Strasbourg Convention or to adopt a new, positive definition of inventions. The delegations rejected a proposal to make “technical progress” an explicit requirement of patentability, reasoning that it was unnecessary because the term “inventions” implies a contribution to technical knowledge and technical progress. Therefore, including “technical” in the language of the convention would be redundant. As a result, no positive legal definition of “technical” emerged from the

---

420 Id.
421 See id. As early as 1959, differences in national patent regimes acted as a barrier to trade and movement of goods, so the members of the European Economic Community convened a working group to discuss a unified patent system. See Cornish & Llewelyn, supra note 18, at 127 (noting that efforts were set aside when Britain failed to enter the EEC, because member states desired having Britain’s expertise in the matter). Such a system would include not only uniform substantive law, but a judicial system through the European Union courts, as arbiter of European patent disputes. See id. at 127–29. Neither a Community patent nor a patent court has been established within the European Union. See id. at 128–29.
422 See Strasbourg Convention, supra note 369, arts. 1–6. The Convention was not limited to members of the Council of Europe. See id. at pmbl n.1. The Convention’s purpose was to unify certain points of substantive patent law to assist European industry and, more ambitiously, to contribute “to the creation of an international patent.” Id. at pmbl.
424 See id. at 758–59.
425 See Beresford, supra note 356, at 13; Pila, supra note 424, at 759.
426 See Beresford, supra note 356, at 13; Pila, supra note 424, at 759.
conference and the *travaux preparatoires* provide no assistance in determining the framers’ intent on the issue.\(^{428}\)

At the time the EPC was negotiated, computer programs existed, but the software industry did not.\(^ {429}\) The delegates feared that fixing definitions of “invention” and “technicality” would impair the flexibility of the EPC to accommodate developing technology.\(^ {430}\) The EPC also does not define “programs for computers” or the other exclusions.\(^ {431}\) Although some delegations strongly opposed including “programs for computers” in the list of Article 52(2) exclusions, the EPC ultimately included it.\(^ {432}\) The delegates concluded that any attempt to define this term would be futile, and that interpretation should be left to the EPO.\(^ {433}\) Leaving interpretive decisions to the EPO has far from clarified the legal definitions of “technical character,” “computer programs,” and the “as such” limitation on exclusions.

In 2000, the contracting parties adopted a new version of the EPC. Currently, Article 52(1) of the EPC mandates that European patents shall be granted “for any inventions, *in all fields of technology*, provided that they are new, involve an inventive step, and are susceptible of industrial application.”\(^ {434}\) It has been suggested that the addition of language “in all fields of technology,” which did not appear in the original EPC, codifies the technology requirement.\(^ {435}\) Although commentators characterize this change to Article 52 as one of the most significant

\(^{428}\) Pila, *supra* note 424, at 760. In examining Professor Pila’s work, Justice Jacob found that her analysis of the *Travaux Preparatoires* led to this result. *Aerotel*, [2007] 1 All E.R. at 232.


\(^{431}\) See EPC 2000, *supra* note 14, art. 52(2). In addition to “programs for computers,” the EPC excludes the following from patentability: “discoveries, scientific theories and mathematical methods”; “aesthetic creations”; “presentations of information”; and “schemes, rules, and methods for performing mental acts, playing games, or doing business.” *Id.*

\(^{432}\) See Pila, *supra* note 424, at 769.

\(^{433}\) See *id*.


\(^{435}\) See Bakels & Hugenholtz, *supra* note 6, at 54. EPC 2000 was adopted by decision on June 28, 2001, in part to incorporate obligations imposed by the Agreement on Trade Related Aspects of Intellectual Property (TRIPS) and the Patent Cooperation Treaty (PCT). *Id.* at 54–55 (citing explanatory document from the Munich conference).
changes in the EPC, they do not expect it to impact substantive law as it relates to patentability. As discussed, under EPC 1973, technical character acted as an implicit requirement of patentability. Moreover, although a technical requirement is not expressly mentioned in the EPC 1973, it is at the heart of Article 52 jurisprudence.

Commentators have also suggested that the technical requirement originates from the Article 57 requirement that inventions be "susceptible of industrial application." The meaning of this phrase, however, differs in various European translations. For example, the French and Dutch languages use the term "industrie" only when referring to manufacturing businesses. In comparison, the English language uses "industry" more broadly. The German requirement of "gewerblich anwendbar" (commercially applicable) is broader than technical manufacturing, but not as broad as the English meaning of industry.

C. The EPO Technical Board of Appeal Decisions

Originally EPO examiners routinely denied any applications related to software inventions and programs under EPO examination guidelines in place at the time. Since its first decision in 1986, however, the Technical Board of Appeal 3.5.01 has taken what could be characterized as an expansive view of these inventions' patent eligibility. Once the board took the position that examiners should not deny an application simply because it involved a computer program, the EPO started down a slippery slope. The Technical Board of Appeal has articulated various tests to delimit the contours of the technical charac-

---

436 CIPA supra note 400, § 2.7.2.
438 See supra text accompanying notes 415–417.
439 See Steinbrenner, supra note 437, at 31–39 (discussing the development of Boards of Appeal decisions relating to Article 52: each decision requires a computer program to have a technical feature to be patentable).
440 Bakels & Hugenholtz, supra note 6, at 4.
441 Id.
442 Id. at 4–5.
443 Id. at 5.
444 See Leith, supra note 6, at 11.
445 Id. at 8–11. Because "any good patent attorney . . . could transmute a software invention into a hardware one," once the decision was made that the inventions were not automatically excludable, "any attempt to hold the line becomes untenable because the definition of protectable technology changes under the continual assault of perceptive patent attorneys who locate logical contradiction and push the examiners towards removing that logical weakness." Id. at 9.
The concept of “technical,” although central to European patent subject-matter determinations, is a difficult legal concept to define. Scholars, judges, and more recently, the EPO president have noted that the Technical Board of Appeal tests, arguments, and justifications vary from case to case. Moreover, the meaning of “as such” has been “anyone’s guess during the past two decades.” The scope of what is excluded from patent subject matter based on 52(2) and (3) has progressively narrowed, leading to a more liberal granting of software-related patents, and moving toward an approximation of U.S. practice.

In its 2009 publication, *Patents for Software? European Law and Practice*, the EPO stated that it “does not grant patents for computer programs (‘software patents’) or computer-implemented business methods that make no such technical contribution.” The EPO uses the term “computer-implemented inventions” to describe inventions that involve the use of a computer, a computer network, or a programmable apparatus, with features that are realized by a computer program. Despite this assertion, a 2000 study indicates that the vast majority of software patent applications up to that time proceeded through the EPO without objection and the vast majority of appeals were granted, provided that the claims were appropriately drafted. According to the EPO, computer-implemented inventions—whether claimed to a physical product or apparatus or to a process or method—are patentable so long as they involve an “inventive technical contribution to the prior art.”
tion, the EPO will grant claims to “computer program” products, such as those stored on some kind of a carrier like a CD or DVD, provided they cause a “further technical effect” beyond the “normal physical effects,” such as the flow of electric current through a computer.\footnote{457}{Id. at 11.}

EPO computer software patent subject-matter decisions have shaped European law concerning business methods.\footnote{458}{See Cornish & Llewelyn, \textit{supra} note 18, at 829.} Nevertheless, business method examinations produce results that are markedly different from similar examinations of non-business related computer programs.\footnote{459}{See Beresford, \textit{supra} note 356, at 183.} Very few computer-implemented business methods are successfully prosecuted in Europe.\footnote{460}{Nicholas Fox & Alex Rees, \textit{A European Perspective on Business Method Patents}, Landslide, July/August 2010 at 30, 30 (confirming that the examining division dealing with business methods currently rejects approximately 95–97\% of the applications).}
The EPO has recently stated that computer programs that “implement business, mathematical or similar methods and do not produce technical effects (e.g. because they solve a business problem rather than a technical one) are not patentable.”\footnote{461}{Patents for Software?, \textit{supra} note 356, at 12.}

Although the EPO considers technical innovations associated with business methods, it does not conduct searches of business method art.\footnote{462}{Gregory A. Stobbs, \textit{Business Method Patents} §14.03[A] (Supp. 2004). Fox & Rees assert that the applications for the three to five percent of business method patents granted have avoided classification as a business method, emphasizing technical advantages and minimizing business benefits. Fox & Rees, \textit{supra} note 460, at 35–37.}

1. The Technical Contributions Approach

The \textit{Vicom} Board considered whether a CAD program—a mathematical method\footnote{465}{VICOM, [1987] O.J.E.P.O. at 14. The EPO cases are more likely to use the term “method,” whereas U. S. cases use “process,” as a result of differing statutory language. See EPC 2000, \textit{supra} note 14, art. 52(2)(a); 35 U.S.C. § 100(b).} for improving digital images by increasing processing speed—or a machine for carrying out the method were excluded from patentability “as such.”\footnote{466}{VICOM, [1987] O.J.E.P.O. at 14.} Following EPO Ex-
amination Guidelines in place at the time, the examiners had rejected
the claim as a mathematical method “as such.”467 The Technical Board
of Appeal, which was not bound by the Examination Guidelines, found
the method to qualify as patent subject matter and not be excluded “as
such.”468

In delineating the difference between an excluded mathematical
method or algorithm and a technical process, the Technical Board of
Appeal stated that a mathematical method produced no direct technical
result, being “an abstract concept prescribing how to operate on the
numbers.”469 But, the Board went on to note:

[I]f a mathematical method is used in a technical process,
that process is carried out on a physical entity (which may be a
material object but equally an image stored as an electric sig-
nal) by some technical means implementing the method and
provides as its result a certain change in that entity.470

Further, the Board of Appeal noted, “[t]he technical means might in-
clude a computer comprising suitable hardware or an appropriately
programmed general purpose computer.”471 As one scholar observed,
 “[t]he Board held that where the claims relate to a technical process,
patentability may arise from novelty in the mathematical algorithm
employed for a technical benefit, and it was immaterial whether the
algorithm was to be implemented in hardware or software.”472

The Board’s reasoning failed to provide clarity. The Technical
Board of Appeal explained that even if the idea for the invention resides
in the non-patentable mathematical method, the applicant will not be
considered to seek protection for the mathematical method “as such” so
long as the claim is directed to a technical process in which the method
is used.473 The most oft-quoted Reason for the Decision provides:

Generally speaking, an invention which would be patentable
in accordance with conventional patentability criteria should
not be excluded from protection by the mere fact that for its
implementation modern technical means in the form of a
computer program are used. Decisive is what technical contribu-

467 Id.
468 See id.
469 See id.
470 Id.
471 Id.
tion the invention as defined in the claim when considered as a whole makes to the known art.474

Vicom established the importance of the overall technical contribution of the invention as the baseline for examining computer-related inventions. This remains European law.475 Under Vicom, the invention as a whole must provide some technical contribution over the state of the art to be considered an invention within the meaning of Article 52.476 The overall reasoning is similar to the holistic approach taken in the 1981 U.S. Supreme Court decision Diamond v. Diehr, which produced a similar result.477

The Technical Board of Appeal remitted the claim back to the examiners, holding that “[a] claim directed to a technical process which process is carried out under the control of a program (whether by means of hardware or software), cannot be regarded as relating to a computer program as such.”478 Before the Board, Vicom had agreed to amend its original patent claims to a multipurpose algorithm and a method of digital filtering, and to direct the claims more specifically to the processing of digital images.479 Applying Vicom’s technical contribution approach, the examiner allowed the amended and limited claims because they were amended for “the general functioning of the computer, rather than to an application designed to execute particular tasks.”480 In taking this holistic approach, the form of the patent claim is not important; the examiner should “disregard the form or kind of claim and concentrate on its content in order to identify the contribution which the subject-matter claimed, considered as a whole, adds to the known art.”481 As one scholar has noted, “[n]o distinction should be drawn between implementation of a computational task in software or in hardware since this does not affect the inventive concept.”482 Another scholar adds that if the “contribution is not of a technical character (i.e. if it falls exclusively within one of the excluded areas), then there is no

474 Id. at 21–22 (emphasis added).
475 Leith, supra note 6, at 28–29.
476 CIPA supra note 400, § 18A.4.1.
477 Diamond v. Diehr, 450 U.S. 175, 188 (1981) (reasoning that an industrial claim that employed a well-known mathematical algorithm was patentable even if the non-computer-related aspects of the process were well known in the art).
479 Kretschmer, supra note 429, at 8.
480 Cornish & Llewelyn, supra note 18, at 824; see Kretschmer, supra note 429, at pt. 2.5.
481 Shemtov, supra note 7, at 507.
482 de Mauny, supra note 7, at 147.
invention.” 483 Vicom opened the door to computer-related inventions being patentable “within the realms of computer science.” 484

Vicom stands for the proposition that subject matter for controlling or carrying out a technical process is patentable, regardless of whether it is implemented on hardware or software. 485 The involvement of a computer program is not sufficient grounds to deny patentability. 486 Although the Technical Board of Appeal drew a line concerning “technical contributions” in Vicom, its reasoning is not compelling. The Board used the term “technical” sixteen times without defining it; likewise, it used the term technical “features” twice, technical “process” on six occasions, technical “means” on four occasions, and technical “subject matter,” “result,” “considerations,” and “contribution” once each. 487 The Technical Board of Appeal neither identified the particular technical contribution made by the invention in Vicom, nor did it discuss the exact nature of that technical contribution. Thus, after Vicom, the meaning of technical contribution remained “a little elusive.” 491

In 1987, the same year as Vicom, the second leading Technical Board of Appeal decision, Koch & Sterzel, appeared. Koch & Sterzel concerned a patent claim to the use of a computer program to control an X-ray machine to ensure optimal performance without overloading the X-ray tube. 492 In the opposition to this claim, Siemens and Philips argued that the only difference between the claim’s subject matter and the state of the art was the use of a new program for a known computer. 493 The essence of the invention was a computer program that only produced a technical effect at the very end of the computing operation. 494 Because the invention involved both a mathematical method in the form of a computer program and a technical apparatus, the op-

483 Shemtov, supra note 7, at 507.
484 Kretschmer, supra note 429 at pt. 2.5. For example, procedures at the operating system level to improve machine functionality or generic algorithms at the application level would normally be patentable. Id.
486 PATENTS FOR SOFTWARE?, supra note 356, at 11.
488 Id. at 14, 19.
489 Id. at 19–21.
490 Id. at 18–19, 21.
491 Aerotel, [2007] 1 All E.R. at 249.
493 Id.
494 Id.
ponents argued that the X-ray apparatus and the computer program had to be considered separately for purposes of determining whether the claim involved an invention. The opposition cited the German Federal Court of Justice, which ruled as follows:

[A teaching in a claim is not technical if] in its essence it states a rule that can be carried out without employing controllable natural forces other than human brainpower, even if the use of a technical means appears expedient or indeed the only sensible and hence the necessary procedure, and even if reference is made to these technical means in the claims or description.

The Technical Board of Appeal rejected this approach and drew a line of distinction between unpatentable computer programs on general purpose computers, in which the electrical signals produced amounted to no more than a reproduction of the information, and those that technically altered the functioning of the unit, which may be patentable. Affirming *Vicom*, the Technical Board of Appeal held that the invention must be assessed as a whole, because the EPC does not prohibit patenting of inventions consisting of a mix of both technical and non-technical means. Further, the Board held that there need not be a constant interaction between the program and the apparatus: “[w]hen the technical effect occurs is irrelevant to the question of whether the subject matter claimed constitutes an invention under Article 52(1) EPC. The only fact of importance is that it occurs at all.”

Therefore, under the technical contribution approach, it does not matter that the contribution to the prior art is a mathematical algorithm so long as there is a technical effect in the apparatus beyond the normal functioning of the computer. Neither the magnitude of the effect, nor the point in the process at which the effect occurred, are determinative.

This reasoning is analogous to the line of reasoning used in *Diehr*, which concluded that it was erroneous to take a point of novelty approach. Rather, if the invention as a whole is different

---

495 Id. at 21.
496 Id. at 22–23.
497 Id. at 23.
499 Id. at 22–23.
500 Id. at 24.
501 See id. at 22–23, 24.
502 See 450 U.S. at 189.
from the prior art, it is irrelevant that the source of novelty resides entirely in a mathematical algorithm or computer program.\textsuperscript{503} The Diehr dissent had predicted that this approach would open the floodgates to the patentability of computer software.\textsuperscript{504} Indeed, patenting strategy in the United States shifted towards “indirect drafting” of software claims following Diehr, subverting the algorithm exclusion to practical nullity.\textsuperscript{505} One commentator has observed that the USPTO Guidelines acknowledged the practice, providing that “the utility of an invention must be within the ‘technological’ arts. A computer-related invention is within the technological arts.”\textsuperscript{506} Thus, the novelty of a patent claim may arise in the software itself, not only through some physical transformation brought about by software.\textsuperscript{507} The EPO developed a similar approach post-Vicom.\textsuperscript{508}

2. The Further Technical Effect Approach

Toward the end of the 1990s, the technical contribution approach met criticism.\textsuperscript{509} Applicants claimed that when software is run on a computer, there is always a machine involved and the invention was, therefore, automatically technical.\textsuperscript{510} Another problem emerged. Vicom and Koch & Sterzel dealt with apparatus claims and process or methods claims.\textsuperscript{511} The EPO still disallowed direct patent claims to “computer program products,” or software.\textsuperscript{512} This created an enforcement issue. The only way to infringe a process or method patent on a computer system solution in Europe was to run the patented program directly.\textsuperscript{513}

\textsuperscript{503} See id. at 183.
\textsuperscript{504} See id. at 218 (Stevens, J., dissenting).
\textsuperscript{505} Kretschmer, supra note 429, at pt. 2.5.
\textsuperscript{506} Id. (citing Examination Guidelines for Computer-Related Inventions, 61 Fed. Reg. 7478, 7479 (Feb. 28, 1996)).
\textsuperscript{507} See id.
\textsuperscript{508} Id.
\textsuperscript{509} Steinbrenner, supra note 437, at 37.
\textsuperscript{510} Laub, supra note 7, at 348.
\textsuperscript{511} KOCH & STERZEL, [1988] O.J.E.P.O. at 19 (involving the patenting of inventions consisting of a mix of technical and non-technical features); VICOM, [1987] O.J.E.P.O. at 14 (involving a claim by which a technical process is considered to reside in a mathematical method).
\textsuperscript{512} See EPC 2000, supra note 14, art. 52(2)(c) (barring computer programs from consideration as inventions eligible for patents).
Unlike the patent on the hardware component of the invention, putative infringers copying and distributing the process could only be liable for indirect infringement, a much more difficult claim.\textsuperscript{514}

In 1997, an IBM appeal clarified the issue concerning direct claims to computer software and articulated what became known as the “further technical effect” approach.\textsuperscript{515} \textit{IBM/Computer Program Product} involved a method claim and claims directed to “computer program products directly loadable into the memory of the computer” and to a “computer program product stored on a computer usable medium.”\textsuperscript{516}

In a second unreported case decided by the Technical Board of Appeal at the same time, IBM appealed an EPO examining division decision refusing a patent application for a computer program.\textsuperscript{517} Following EPO Examination Guidelines in place at the time, the examining division refused the computer program products claims, drawing a clear line between patentable and unpatentable subject matter by disallowing the computer program products claim to preclude the possibility of a program written on a sheet of paper from patentability.\textsuperscript{518} The Technical Board of Appeal did not agree and considered whether and under which circumstances a computer program product could be valid subject matter.\textsuperscript{519}

The Board in \textit{IBM I} interpreted the language of the EPC to mean that the drafters had not intended to exclude all computer programs from patentability, but only computer programs “as such.”\textsuperscript{520} Because computer programs must be patentable when they have a technical character, not all computer programs are \textit{prima facie} excluded from patentability.\textsuperscript{521} In delineating between patentable and non-patentable programs, the Technical Board of Appeal again excluded from patentability programs causing common physical modifications to the

\textsuperscript{514} Steinbrenner, supra note 437, at 62.
\textsuperscript{516} Id. at 610, 611.
\textsuperscript{518} See id.
\textsuperscript{519} See id. at 309. One commentator even suggests that IBM colluded with the EPO by drafting its claims for direct patentability of the computer program, and not indirectly to a system or method for inventions that would have “clearly” been patentable as method claims in line with previous EPO decisions. See Kretschmer, supra note 429, § 2.5.
\textsuperscript{521} See id. at 619.
computer, such as electrical currents carrying out program instructions. The Board of Appeal held that a computer program may not be excluded from patentability “[i]f the program, when running on a computer or loaded into a computer, brings about, or is capable of bringing about, a technical effect which goes beyond the ‘normal’ physical interactions between the program (software) and the computer (hardware on which it is run).” The Board found that the technical character in the “further effects” derived from the hardware’s execution of the instructions given by the computer program. This further technical effect beyond the normal functioning of the computer may occur when the software manages an industrial process or the working of a machine. A “further technical effect” also occurs in cases where the computer is a necessary means to obtain the further technical effect, although the Board neither defined nor provided examples of when a computer might provide the necessary means. In reaching its decision, the Technical Board of Appeal cited Vicom’s reasoning that it was illogical to grant a patent for the process, but not for the apparatus for carrying out the method. The IBM I Board utilized the reasoning of Vicom in concluding as follows:

[It would be illogical to] grant a patent for both a method and the apparatus adapted for carrying out the same method, but not for the computer program product, which comprises all the features enabling the implementation of the method and which, when loaded in a computer, is indeed able to carry out that method.

The Board distinguished its holding from the practice in U.S. and Japanese patent offices, which allow patent claims to computer

---

522 See id. at 620.
523 Id. at 632 (emphasis added). The Technical Board of Appeal defined “running on a computer” to mean that “the system comprising the computer program plus the computer carries out a method or process.” Id. “Loaded into a computer,” according to the Board, means that the programmed computer is capable of carrying out a method that constitutes a system, device, or apparatus. Id. Professor Kretschmer characterizes this reasoning concerning further technical effect as “sailing close to the wind” and puts the reasoning in plainer language: “software is not a computer program ‘as such’ if it is innovative and it works.” See Kretschmer, supra note 429, § 2.5.
525 Id.
526 Id. at 620–21.
527 See id.
programs, acknowledging that while these foreign practices represented a “useful indication of modern trends,” those legal systems did not contain exclusions to patentability similar to those in Articles 52(2) and (3).\footnote{530}{See id. at 617.}

Prior to \textit{IBM I}, the EPO had focused on putting computer-implemented inventions into the framework of tangible, physical machines.\footnote{531}{See Leith, supra note 6, at 29–31.} After \textit{IBM I}, the focus shifted to the nebulous “further technical effect” caused by the computer program.\footnote{532}{See id. at 33.} One scholar suggests that this transition provides evidence that the Board changed from its previous machine metaphor to an analogy that the software itself is “machine-like” and that the Board thus “dropped the fiction that a patentable invention was in the machine which was part hardware and part software.”\footnote{533}{Id. at 32–33.} Nevertheless, the \textit{IBM I} case narrowed the scope of the exclusion so that more inventions achieved patentability. The requirement of a “further technical effect” is not a bar to patentability: “[a]ny computer program that works is not a computer program” as such if a carefully drafted claim refers to technical considerations.\footnote{534}{Grosche, supra note 359, at 273.}

Despite the Board’s statement to the contrary, it is difficult to avoid speculating that the U.S. treatment of software influenced the EPO to liberalize its treatment of computer program patent validity questions. By the time of the \textit{IBM I} decision, the CAFC had largely removed patent subject-matter restrictions with respect to computer software and business methods.\footnote{535}{See Diehr, 450 U.S. at 177–78; State St. Bank & Trust Co. v. Signature Fin. Grp. Inc., 149 F.3d 1368, 1373 (Fed. Cir. 1998).} In light of this real or apparent pressure, it is not surprising that the EPO may have been motivated to loosen constraints in order to allow European inventors to compete more evenly with U.S. inventors.

Ultimately, the reasoning in \textit{IBM I} is unsatisfying and does little to advance the meaning of either technical contribution or the boundaries of the “as such” exception. The circular reasoning that computer programs “as such” are not patentable because they are not technical, and that programs with technical character are valid patent subject matter because they are not programs “as such,” is not illuminating. Without further articulating the nature of the “further technical effect” test, the \textit{IBM I} decision failed to delineate the exact nature and scope
of the computer program exclusion.\textsuperscript{536} To many, \textit{IBM I} formally ended the computer program exclusion in the EPC and the fiction of a limitation based on technical character, technical contribution, or technical effect.\textsuperscript{537}

3. The Any Hardware Approach

Until 2000, the EPO applied three concepts in evaluating the patent eligibility of computer-implemented inventions: first, there must be a technical contribution (\textit{Vicom}); second, in determining technicality, the invention must be evaluated as a whole (\textit{Koch & Sterzel}); and third, the contribution must cause a further technical effect (\textit{IBM}).\textsuperscript{538} In 2000, however, the EPO clearly departed from requiring a technical contribution as part of the Article 52 analysis in a series of cases beginning with \textit{PBS Partnership}.\textsuperscript{539}

\textit{PBS}, decided one year after \textit{IBM I}, did not involve a patent claim to a computer program, but claimed a method and an apparatus, where the apparatus was a computer programmed to run the method.\textsuperscript{540} The method used data processing to control pension benefit programs for subscriber employers by using standard factors, such as actuarial life spans, for calculating pensions.\textsuperscript{541} The Technical Board of Appeal denied the method claim as a method for doing business “as such,” and accordingly, relied on Article 52(2) and (3) to preclude the claim.\textsuperscript{542} The Board acknowledged the technical effects test and found that all of the steps involved in processing and producing the information had a “purely administrative, actuarial and/or financial character.”\textsuperscript{543} It rejected the appellant’s argument that referring to data processing and computing means in the claims conferred technical character on the

\textsuperscript{536} See \textit{Leith}, supra note 6, at 33.

\textsuperscript{537} See Kretschmer, supra note 429, § 2.5. Nevertheless, Beresford’s work demonstrates that the exclusions on patentability under the EPC were almost equivalent to United States practice by 2000, and that the “technical effect” criterion, which applies to all EPO examinations, could be met with careful drafting of claims. \textit{Id.} Under EPO practice, most innovative software includes a technical effect, such as higher speed, more economical use of memory, or an improved user interface. \textit{Id.} § 3.1


\textsuperscript{540} See \textit{id.} at 441–42.

\textsuperscript{541} See \textit{id.} at 445.

\textsuperscript{542} \textit{Id.} at 449.

\textsuperscript{543} \textit{Id.}
method, and concluded that “[a] feature of a method which concerns the use of technical means for a purely nontechnical purpose and/or for processing purely non-technical information does not necessarily confer a technical character to such a method.” 544 There is nothing in PBS’s holding regarding the method claim that departed from previous EPO cases. 545

By contrast, the Board treated the apparatus claim differently, signaling a “substantial departure from previous case law.” 546 The Technical Board of Appeal found that an apparatus programmed for use in a particular field was an invention within the meaning of Article 52(1), even if the field was business, because it involved a physical entity. 547 If the patent claim was directed to the apparatus, the formal category of the claim implied physical features which could qualify as technical features of the invention. 548 Therefore, use of any physical entity would bring the claim outside the Article 52(2) and (3) exclusions. 549 Lord Justice Jacob in Aerotel coined the descriptive nomenclature “any hardware” to refer to this approach. 550

The PBS reasoning appears formalistic rather than substantive. Article 52(2)(c) excludes “methods” of doing business, but not “apparatuses” or “products.” 551 Although the claim proceeded as an Article 52 invention, the Technical Board of Appeal ultimately found that the apparatus was nonetheless not patent subject matter because it did not meet the requirements of “inventive step” under Article 56. 552 In making this determination, the Board found no non-obvious improvement over the identified prior art of “existing private pension plans,” because the improvement was essentially economic and not technical. 553 The

544 Id. at 450.
545 See Aerotel, [2007] 1 All E.R. at 264–65 (noting an English judge’s observation that “it [would be] difficult to think of a concrete case where the approaches [in VICOM and PBS] would lead to a different result, but . . . they could.”).
546 See id. at 264 (“Thus far there is nothing remarkable about [PBS]. It is the reasoning in relation to the apparatus claim which represented a substantial departure from previous case law.”).
548 See id.
549 See id.
552 See EPC 2000, supra note 14, art. 56 (“An invention shall be considered as involving an inventive step if, having regard to the state of the art, it is not obvious to a person skilled in the art.”).
claimed invention did not provide a technical solution to a technical problem.\footnote{See id. at 456–57.}

*Hitachi/Auction Method*, the next EPO case decided using the “any hardware” approach, is the more important of the two cases in that it set the standard for EPO examination of business method and business systems patents.\footnote{Case T-258/03, Auction Method/HITACHI, [2004] O.J.E.P.O. 575, 587 (Technical Bd. Appeal 3.5.01, Apr. 21, 2004), available at http://www.epo.org/law-practice/case-law-appeals/pdft030258ex1.pdf (“Method steps consisting of modifications to a business scheme and aimed at circumventing a technical problem rather than solving it by technical means cannot contribute to the technical character of the subject-matter claimed.”).} *Hitachi* involved a method claim for the automated auction method, an apparatus claim for running the auction via a network, and a computer program claim.\footnote{Id. at 579.} The *Hitachi* Board, following *PBS*, explained that the technical contribution approach was incorrect in that an invention’s technical contribution was more appropriately considered for determining novelty and inventive step rather than subject matter.\footnote{Id. at 581–82.} The *Hitachi* board instructed that “[a] mix of technical and non-technical features may be regarded as an invention within the meaning of 52(1) EPC and that the prior art should not be considered when deciding whether the claimed subject matter is such an invention.”\footnote{Id. at 582.}

The *Hitachi* Board found that the apparatus claim was outside of the Article 52 exclusions because it possessed technical features, such as a “service computer,” “client computer,” and “network.”\footnote{Id. at 583.} The claim thus met the requirements of technicality.\footnote{See id.} This conclusion was consistent with the *PBS* finding that “[a]n apparatus constituting a physical entity or concrete product, suitable for performing or supporting an economic activity is an invention within the meaning of Article 52(1).”\footnote{*PBS*, [2001] O.J.E.P.O. at 451.} In this reasoning, both *Hitachi* and *PBS* depart dramatically from the technical contribution requirements of *Vicom* and *Koch & Sterzel*. Rather than achieving technical character by making a computer run better or faster, the claimed inventions in *Hitachi* and *PBS* achieved technical character by virtue of being loaded onto the ma-
chine. Therefore, the claimed inventions escaped classification as a computer program or a business method “as such.”

Hitachi also changed course from PBS in finding that the method claim was not excludable as a business method “as such.” The PBS Board had found that use of a technical means for purely non-technical purposes did not confer technical character on the method. A different Technical Board of Appeal panel, roughly forty months after PBS, found that it was inappropriate to quantify or weigh the technical aspects of the method claim as part of the Article 52 analysis. Doing so would require consideration of possible novel or inventive contributions to the prior art, which the Article 52 analysis of invention does not allow. The Hitachi Board mandated that method claims and apparatus claims be treated the same for purposes of Article 52 analysis. In both instances a physical feature of the entity, or the nature of the activity, could imply technical character. Therefore, what examiners and practitioners had previously considered a non-technical activity now achieved technical character and could not be dismissed as a non-invention “as such.” This interpretation of Articles 52(2) and (3) substantially broadened the concept of invention. The Technical Board of Appeal acknowledged this expansion, noting that technical acts as familiar as writing with pen and paper would meet its Article 52 technical criteria. Under the Hitachi analysis, then, the first step in determining patentability was whether the claimed subject matter has a prima facie technical effect. A business method passes the Article 52 test as long as it is attached to “any hardware.” But such inventions must also pass the Article 56 inventive step test.

562 See HITACHI, [2004] O.J.E.P.O. at 583; PBS, [2001] O.J.E.P.O. at 453 (“An apparatus constituting a physical entity . . . suitable for performing or supporting an economic activity, is an invention.”).
566 Id.
567 See id. at 585 (“[T]he Board . . . is not convinced that the wording of Article 52(2)(c) EPC . . . imposes a different treatment of claims directed to activities and claims directed to entities for carrying out these activities.”).
568 Id.
569 Id.
570 Id.
571 See Fox & Rees, supra note 460, at 32 (“[F]ollowing Hitachi[,] this initial hurdle is not hard to overcome as the presence of involvement of any physical apparatus is sufficient.”).
Article 56 provides that “[a]n invention shall be considered as involving an inventive step if, having regard to the state of the art, it is not obvious to a person skilled in the art.” The EPC follows a “problem-and-solution” approach to the inventive step analysis. For there to be an inventive step, there must be an objective technical problem with a technical solution, and the EPC considers whether the claimed invention would have been obvious to a skilled person starting from the closest prior art. Patent examiners often break this approach into four steps: first, determine the closest prior art; second, determine the distinguishing technical feature and its technical effect; third, formulate the objective technical problem; and fourth, determine whether a skilled person would have solved the technical problem by the solution specified in the patent claim. The problem-and-solution approach requires analysis of the prior art, but only takes into account features of the invention that contribute to the technical character of the invention. Thus, it is the second step, technical effect, at which most business methods fail. The central feature of Hitachi’s auction method automatically increased the auction price if more than one bidder offered the same “desired price.” Although Hitachi argued that the technical effect resided in the claimed invention’s ability to overcome delay between bidders and the server, the Board found the invention claimed was “a mere automation of the non-technical activity of performing a Dutch auction.” The Board also acknowledged that the invention might have contained a technical feature that went beyond how a human auctioneer would perform the auction without technical support. Because the programming measure required to rank such bids “would have been obvious” to anyone skilled in data

573 EPC 2000, supra note 14, art. 56. The EPC’s inventive step is similar to non-obviousness in the United States, although the approaches are different. See Gwilym V. Roberts et al., Transatlantic Patenting, Landslide, Nov./Dec. 2009, at 32, 33.
574 See Laub, supra note 7, at 349, 360 (comparing the EPO and USPTO approaches to the examination of the patentability of computer-implemented inventions).
575 Guidelines for Examination in the EPO, supra note 414, Part C, ch. IV, § 11.5.
576 Fox & Rees, supra note 460, at 32.
577 Guidelines for Examination in the EPO, supra note 414, Part C, ch. IV, § 11.5.
578 Id.
580 Fox & Rees, supra note 460, at 32.
582 Id. at 587.
583 Id. at 588.
584 Id. at 589.
processing, however, the invention did not meet the requirements of an inventive step.\footnote{585}{Id.}

EPO boards have continued to refine their approach. Less than two years after deciding Hitachi, the Technical Board of Appeal decided a third case. Microsoft/Clipboard Formats involved an appeal from the examining division’s refusal of claims to a method that improved the functionality of Windows 3.1, including a computer program to execute the method.\footnote{586}{See Case T-424/03, MICROSOFT/Clipboard Formats I, [2006] E.P.O.R. 414, 417 (Technical Bd. Appeal 3.5.01, Feb. 23, 2006), available at http://www.epo.org/law-practice/case-law-appeals/pdf/t030424eu1.pdf.} The Technical Board of Appeal set aside the appeal, remitting the claim to the examining division to grant the patent.\footnote{587}{Id. at 422.} Following Hitachi’s approach that a method using technical means is an invention within the meaning of Article 52, the Board found the clipboard method claim eligible for patent protection because a “computer system including a memory is a technical means.”\footnote{588}{Id. at 419–20.} Rather than moving to the Article 56 analysis, however, the Board took the opportunity to distinguish a patentable method implemented in a computer system from non-patentable programs. The former “represents a sequence of steps actually performed and achieving an effect”\footnote{589}{Id. at 420.} while the latter “just have the potential of achieving such an effect when loaded into, and run on, a computer.”\footnote{590}{Id.} The Board further explained that even though a method of operating a computer may use a computer program, a claim to the method is not a claim to the computer program “as such.”\footnote{591}{Id.}

In applying their reasoning to the claim in question, the Microsoft Board delineated their divergent approach. Claim 5 covered a “computer-readable medium having computer-executable instructions (i.e. computer program)” to perform the method.\footnote{592}{MICROSOFT, [2006] E.P.O.R. at 420.} Citing Hitachi, the Microsoft Board found that the computer software passed the Article 52 hurdle because it “relates to a computer-readable medium, i.e. a technical product involving a carrier.”\footnote{593}{Id. (finding that the software possessed the technical character necessary under Article 52).} The Board cited IBM I to support its finding that:
[The] computer-executable instructions have the potential of achieving the . . . further technical effect of enhancing the internal operation of the computer, which goes beyond the elementary interaction of any hardware and software data processing . . . The computer program recorded on the medium is therefore not to be considered a computer program as such.\textsuperscript{594}

This conclusion does not follow from \textit{IBM I}, however, because \textit{IBM I} determined that a computer program could not be considered a technical means unless it produced a “further technical effect.”\textsuperscript{595} Thus, \textit{Microsoft} seemed to carve out a \textit{sui generis} category for computer programs.\textsuperscript{596} As one scholar summarized the \textit{Microsoft} holding, “a computer-reusable medium, including a program stored on it, has technical character because the computer-readable medium is a technical product.”\textsuperscript{597}

The Technical Board of Appeal subsequently determined that the method claim met the Article 56 requirement of inventive step. Referring to Windows 3.1, the closest prior art, the Board found the method “solves the problem of how to facilitate a data exchange across different data formats, in particular when transferring non-file data.”\textsuperscript{598} The Board never identified the problem-solution approach, but merely concluded that there is a problem, there is a solution, and that the invention “does not derive in an obvious manner from the pre-existing operating system.”\textsuperscript{599} The \textit{Microsoft} Board found that the method thus met the requirement of inventive step because the method “solves the problem of how to facilitate a data exchange across different data formats.”\textsuperscript{600} The Board also found that the method met the novelty requirement.\textsuperscript{601} The Board noted that the method “solve[d] a technical problem by technical means . . . in order to enhance the internal operation of a computer.”\textsuperscript{602}

Although \textit{Microsoft} followed \textit{PBS} and \textit{Hitachi} in the “any hardware” approach to Article 52 analysis, it departed in its Article 56 analysis of inventive step. The \textit{Microsoft} Board did not treat the computer program

\textsuperscript{594} \textit{Id.}
\textsuperscript{596} Steinbrenner, \textit{supra} note 437, at 66.
\textsuperscript{597} \textit{Id.}
\textsuperscript{598} \textit{MICROSOFT}, [2006] E.P.O.R. at 421.
\textsuperscript{599} \textit{Id.} at 421–22.
\textsuperscript{600} \textit{Id.}
\textsuperscript{601} \textit{Id.} at 421.
\textsuperscript{602} \textit{Id.} at 420.
as excludable prior art, as the business methods had been in the previous cases.603 The Technical Board of Appeal examined the computer program along more conventional lines without explaining why the approach was different.604 Without explaining what made the method and computer claims different from those in PBS and Hitachi, the Technical Board of Appeal directed the Examining Board to grant the patent.605 This order “opens the way to the patentability in principle of any computer program in Europe,”606 and indeed, the Technical Board of Appeal and national courts have recently moved in this direction.607 The difference in analysis and results in the recent “any hardware” trilogy of Technical Board decisions depends on whether the prior art is a “business method,” as in PBS and Hitachi,608 or a computer program, as in Microsoft.609 Although neither is excluded from patentability “as such” if they are present on any hardware, there appears less hostility to the patentability of computer programs than to business methods.610 Thus, the Technical Board of Appeal provided a much higher exclusionary bar for business methods under Article 56.611 Nonetheless, even with this refinement, subject-matter boundaries remained unclear.

Confusion persisted after Microsoft. Duns Licensing, decided months after Microsoft, is significant not only for applying the any hardware test, but for addressing differences between the EPO and English Court of Appeal approaches to the “technical” requirement.612 Duns Licensing claimed a research method of estimating sales activity by correlating sales activities at reporting outlets according to certain criteria, and claimed an apparatus for maintaining inventory based on the method’s results.613 The Technical Board of Appeal articulated the EPO’s pro-

---

603 See id. at 420–21.
604 Ballardini, supra note 7, at 567.
606 Ballardini, supra note 7, at 567.
607 See Steinbrenner, supra note 437, at 66 (citing a number of cases in which computer programs have been patented, including software to control a car radio module—the only case to meet the prior art hurdle—a garbage collection in a computer memory, and a data retrieval method).
608 See Ballardini, supra note 7, at 566–67.
609 See id. at 567.
610 See id. (claiming that the scope of patent protection for computer programs has narrowed to the requirement of technicality).
611 See EPC 2000, supra note 14, art. 56; Fox & Rees, supra note 460, at 32.
612 DUNS LICENSING, [2008] O.J.E.P.O. at 70–71 (showing EPO characterizing the divergent U.K. software subject matter approach as “not consistent with a good-faith interpretation of the European Patent Convention” in part because the U.K. approach relied on the technical contribution approach that the EPO had previously abandoned).
613 See id. at 51–52, 54.
patent approach, stating that Article 52(1) presents the “fundamental maxim of the general entitlement to patent protection for any invention in all technical fields.” 614 The Board continued on to explain that EPC 52(2) and (3) embody the technical character requirement. Despite acknowledging that Article 52 presented interpretive problems because there was no legal or commonly accepted definition of “invention,” 615 the Board insisted that by not defining “invention” the EPO had allowed new technologies to develop. 616

Reviewing the legislative history of EPC 52(2), the Duns Licensing Technical Board of Appeal found that the EPC introduced Article 52(3) to prevent a broad interpretation of Article 52(2) excluded matter. 617 The Board considered the revised language in Article 52 in EPC 2000—requiring examiners to grant patents “in all fields of technology”—as expressly confirming technical character as a legal requirement for an invention. 618 Citing the Basic Proposal for EPC 2000, the Board asserted that claimed subject matter is reserved for inventions “with ‘technical character’ or to be more precise—[inventions that] involve a ‘technical teaching’, i.e. an instruction addressed to a skilled person as to how to solve a particular technical problem using particular technical means.” 619 In the same paragraph, the Board emphasized that creations in engineering and technology are entitled to protection under the EPC. 620

The Duns Licensing Board then explained the relationship between the Article 52 and Article 56 tests. The first question is whether the claimed subject matter meets the Article 52 requirements for invention. 621 This analysis should be “strictly separated from and not mixed up with” the other patentability requirements, including inventive step. 622 When evaluating claims with a mix of technical and non-technical features, Article 56—the inventive step requirement—is key to distinguishing between valid patent subject matter and claims not

---

614 Id. at 62.
615 Id. at 60, 62.
616 See id. at 62 (“[T]he EPO has not developed any such explicit definition . . . for good reasons.”).
617 Id. at 63.
619 Id. at 65.
620 Id at 64.
621 See id. at 67 (quoting an earlier decision where the Technical Board of Appeal first determined whether the claim constituted an invention within the meaning of Article 52(1)).
622 Id. at 68.
entitled to patent protection. Only technical features of a claimed invention are relevant when assessing inventive step because the innovation must be in a technical field, not in an unpatentable field.

Based on this framework, the Duns Licensing method claim failed the Article 52 analysis. The Board determined that “gathering and evaluating data as part of a business research method do not convey technical character to the business research method if such steps do not contribute to the technical solution of a technical problem.” The method claim had referred to a database, which did not confer technical character because it was not a technical system. Therefore, it did not solve a technical problem and the claim was not eligible for patent protection. The Technical Board of Appeal was explicit in its conclusion that business research activities do not solve a technical problem related to a technical field: “interaction with and exploiting information about the physical world belongs to the very nature of any business” and accepting those features as technical would “render the exclusion for business methods under Article 52(2)(c) EPC meaningless.”

The Duns Licensing analysis differed in addressing the claim to a central processor to perform the individual steps of the method. The Board concluded, without analysis, that under Hitachi the claim to the technical apparatus qualified it as an Article 52 invention. In line with Hitachi, however, the claim failed on the inventive step analysis. The Board found that the new algorithm used and the method of estimating sales activity on a known system were “part of a business research method and do not contribute to the solution of any technical problem.” Therefore the examiner should not consider such methods in assessing inventive step because they were inherently non-technical, as well as being known.

These cases illustrate both the difficulty in drawing a line in the patent subject-matter sands, and the shifting of the line when applying

623 See id. at 61 (stating principles of patentability, one of which is that an “inventive step” can only be based on technical features and that non-technical features “as such” do not provide a technical contribution and are ignored when assessing the “inventive step” element of the analysis).


625 Id. at 46.

626 Id. at 75–76.

627 Id.

628 Id. at 75.

629 Id. at 76.

630 DUNS LICENSING, [2008] O.J.E.P.O. at 76.

631 Id. at 78.

632 See id. at 77–78.
the “technical” requirements test. More than two decades of patent case law demonstrate that the technical character tests draw a somewhat arbitrary line.633 The primary reason to stick to the requirement of technicality appears to be that it has “always existed in Europe.”634 Nevertheless, considerable consensus exists that the rule in Europe is nebulous and that clarification is needed.635

The EPO, in seeking to define the relevant criteria for determining patent subject matter, began by focusing its Article 52 analysis on whether there is an invention and whether that invention is technical in nature and makes a technical contribution.636 The EPO then changed course, with claimed inventions perfunctorily passing the technicality requirement under Article 52 if the claim explicitly included any hardware.637 Evaluation of technicality shifted to the problem-and-solution approach under the Article 56 inventive step analysis.638 Difficulty in defining the terms “technical contribution” and “as such” has resulted in various and inconsistent approaches which have undermined predictability in the field, as the “any hardware” approach has shown.639 Furthermore, the EPO draws the line differently when the excluded category is a “business method” as opposed to a “computer program.”640 Although Article 52(2) expressly excludes both claim types,641 post-Microsoft it appears that a computer program is more likely to clear the inventive step hurdle, whereas computer-implemented pension benefit systems, auctions, and sales estimating activities will not.642 Regardless of which “technical” test reviewing bodies employ, or whether they conduct the analysis under Article 52 or 56, such reviews do not yield clear and satisfying results.643

633 Bakels & Hugenholtz, supra note 6, at 33.
634 Id.
635 See id.
636 See Ballardini, supra note 7, at 567.
637 See id.
638 See Guidelines for Examination in the EPO, supra note 414, Part C, ch. IV, § 11.5.
639 See Ballardini, supra note 7, at 567, 570.
640 See id. at 566–67.
641 See EPC 2000, supra note 14, art. 52.
642 See Ballardini, supra note 7, at 567 (observing that, in Microsoft, the Board was able to avoid the Article 52 “as such” exclusion and set the stage for future computer program patentability in Europe, although the Board had excluded, for example, pension benefit systems in Pension Benefits Systems and auctions in Hitachi).
643 See id. (“[T]he difficulty in pinpointing . . . a criterion [for assessing the patentability of computer programs] has caused the Boards to embrace various and inconsistent approaches, leading to a general lack of legal coherency in the field.”).
D. Decision of the Enlarged Board of Appeal

On October 22, 2008, Alison Brimlow—then President of the EPO—referred questions to the Enlarged Board of Appeal under Article 112(1)(b), asking for clarification on a number of issues concerning the patentability of computer-implemented inventions. Her referral was not the first request for clarification. Although national court justices have no standing to refer cases to the EPO, Lord Justice Jacob of the English Court of Appeal had suggested in [*Aerotel*](#) that certain issues needed clarification through a referral to the Enlarged Board. The EPO did not agree. In an informal letter dated February 22, 2007, the EPO President, Alain Pompidou, found the request for referral unnecessary. During oral arguments, the appellant in [*Duns Licensing*](#) submitted Justice Jacob’s questions for referral to the Enlarged Board of Appeal. The [*Duns Licensing*](#) Board rejected the request, explaining that diverging decisions were allowable in the EPC legal system as part of the “evolution of the jurisprudence,” unlike “‘case law’ in the strict Anglo-Saxon meaning of the term.”

Eighteen months after President Brimlow’s referral, the Enlarged Board of Appeal issued what amounted to a non-decision and did not reach the merits of the questions referred, finding the referral did not meet the requirements of EPC 112(1)(b). When an EPO President refers questions to the Enlarged Board of Appeal, the EPC has two requirements for admissibility: first, either the questions must be answered to ensure uniform application of the law, or the questions concern points of law of fundamental importance; and second, two Boards of Appeal must have promulgated different decisions on the question referred. Addressing the first requirement, the Enlarged Board observed that the general subject of the patentability of computer-implemented inventions is of fundamental importance, noting the “heated debate in administrative and judicial practice” and the prob-

---

645 [*Aerotel*, [2007] 1 All E.R. at 236. The questions were different from those ultimately submitted by the EPO president to the Enlarged Board of Appeal. Justice Jacob specifically asked for clarification on the key characteristics of the method of doing business exclusion. See *id.* at 236, 241. This was a question not addressed by the Enlarged Board. See Reiner B. Bakels, [*Software Patentability: What Are the Right Questions?*] 31 Eur. Intell. Prop. Rev. 514, 520 (2009).
648 *Id.* at 59.
650 EPC 2000, *supra* note 14, art. 112.
lem of EPC contracting states applying different reasoning. The Enlarged Board noted internationally “increasingly convergent decisions,” which included Duns Licensing, the 2008 English Court of Appeal decision Symbian Ltd v. Comptroller-General of Patents, and the CAFC case In re Bilski. The Enlarged Board also noted that the failure of the European Union to harmonize EU patent law for computer-implemented inventions was evidence that where to “draw the dividing line between applications relating to programs for computers as such” and “applications related to patentable technical solutions, in the form of [computer-implemented inventions], still cannot be assumed.” Nonetheless, despite worldwide disharmony, the Enlarged Board did not consider the worldwide debate on the patentability of computer-implemented inventions, and diverging national decisions, relevant to their resolution of the referral.

In addressing the second part of the test, the Enlarged Board concluded—without deciding the issue—that existing EPO case law provided for a consistent approach. In construing the meaning of “different decision” in the context of the EPC second requirement, the Enlarged Board stressed the “interpretative supremacy” of the Technical Boards of Appeal and noted that the Enlarged Board cannot develop law as do the Technical Boards of Appeal. The Enlarged Board found “different decisions” to mean a “conflict in case law making it difficult if not impossible for the Office to bring its patent granting practice into line with the case law of the Boards of Appeal.” In making this determination, the Board stated:

---

652 Id. at 16.
653 See DUNS LICENSING, [2008] O.J.E.P.O. at 78.
657 Id. at 17.
659 Programs for Computers, [2010] O.J.E.P.O. at 20–21 (finding that the provision for “different decisions” was ambiguous in the English, French, and German versions of the EPC, and resorting to the guiding principles of the Vienna Convention on the Law of Treaties to determine the intent of the EPC framers).
660 Id. at 22.
661 Id. at 25.
[T]he Enlarged Board must also consider whether the divergent decisions might not be part of a constant development, possibility still ongoing, in jurisprudence on recent patent law issues, in the course of which older decisions have lost their significance and so can no longer be considered in connection with newer decisions. Such putative differences do not justify presidential referrals, legal development being one of the principal duties of the Boards of Appeal, in particular in new territory.662

The Enlarged Board then proceeded to review decisions in light of the referred questions to determine whether a conflict existed. The only inconsistencies the Enlarged Board found related to the referred question, “[c]an a computer program only be excluded as a computer program as such if it is explicitly claimed as a computer program?”663 The only “divergence” in case law the Enlarged Board identified was between the IBM I and Microsoft cases. In IBM I, the Technical Board of Appeal had determined that a claim to a computer program itself is patentable if it produces a “further technical effect” while it runs.664 The definition of “further technical effect” did not mention the state of the art, so that the Article 52(2) and (3) determination did not consider the prior art.665 The further technical effect does not have to be new.666 The Enlarged Board affirmed that IBM I consciously abandoned the “contribution approach,” and observed that the Technical Board of Appeal has not contested this shift in any decision since.667

The Enlarged Board subsequently discussed PBS and Hitachi, noting that neither case addressed whether a claim to a program on a computer readable medium avoided exclusion.668 The Enlarged Board found that Microsoft had extended the reasoning of Hitachi to decide that such claims have technical character because they relate to a computer readable medium.669 Although the positions taken in IBM I and Microsoft were different, the Enlarged Board found the differences reflected development in the case law over seven years, not different opin-

662 Id. at 30.
663 Id. at 32.
664 Id. at 35, 37.
666 Id.
667 Id.
668 See id. at 39.
669 See id.
ions meriting a referral. To support this conclusion, the Enlarged Board indicated that although IBM I remained seminal on the further technical effect requirement for claims directed to computer software, no Technical Board of Appeal had followed IBM I on its technical contribution approach, and no Technical Board of Appeal had challenged the Microsoft approach.

In addition to acknowledging that the law in the EPO is in a state of development, the Enlarged Board expressly declined to define the term “technical.” Addressing the question of whether the activity of programming a computer includes technical considerations, the Enlarged Board conceded that computer algorithms can genuinely be viewed as either a pure mathematical-logical exercise or as defining a procedure to make a machine carry out a certain task. The Enlarged Board found that the EPC takes the former view: abstract formulations of algorithms do not belong to a technical field, but require “further” technical effects to be patentable.

The Enlarged Board’s decision neither advances the law nor clarifies where to draw the line in these cases. Although it did not expressly address the issue of business method patents, the Enlarged Board cited Duns Licensing as laying out the “elaborate system” developed by the EPO for taking the list of excluded subject matter in Article 52(2) into account in assessing the inventive step. The Enlarged Board did not “judge whether this system is correct.” Rather, it found that “it is evident from its frequent use in decisions of the Boards of Appeal that the list of ‘non-inventions’ in Article 52(2) EPC can play a very important role in determining whether claimed subject-matter is inventive.” Therefore, the problem-and-solution approach of Article 56 now determines whether a computer-implemented invention achieves the technical character required for patentability, and the Article 56

670 Id. at 45.
672 Id. at 31.
673 See id. at 58.
674 See id. at 58–59.
675 See id. at 31, 45. Nevertheless, the decision’s impact may be persuasive in shifting the practice in the U.K. Patent Office to a more favorable approach to computer-implemented inventions. See Chris Benson, United Kingdom: Business as Usual for Software Patents at the EPO, MONDAQ BUSINESS BRIEFING (May 18, 2010), http://www.mondaq.com/article.asp?articleid=100752.
676 See Programs for Computers, [2010] O.J.E.P.O. at 47.
677 Id.
678 Id.
hurdle is higher for computer-implemented business methods than for computer programs.

IV. The U.K. and German Judicial Treatment of Software Patents

In this Part, we consider national treatment of software and business method patent subject-matter questions in the United Kingdom and Germany. Not only are the United Kingdom and Germany two of the most influential countries in Europe, the two countries have the most developed patent law jurisprudence in Europe. The U.K. patent jurisprudence in particular is well-developed, and its patent subject-matter approach to software and business method claims has clashed with EPO rulings. German practice, while not as divergent as U.K. practice, is likely to differ from EPO approaches in at least some respects. We start our national analysis with a discussion of U.K. patent subject-matter practice.

A. U.K. Patent Subject-Matter Treatment of Software and Business Method Claims

In 1977, the United Kingdom transposed the EPC into law. The Patents Act altered both substantive and procedural law so extensively that it caused the “largest culture shock in [U.K. patent] history.” Although one purpose of the EPC is harmonization of substantive patent law among the contracting states, the United Kingdom chose statutory language different from EPC Article 52. Differences in the wording in the U.K. statute may contribute to differing interpretations of

---

679 See Ballardini, supra note 7, at 567 (referring to the United Kingdom and Germany as European Patent Convention “major players”).
680 See infra text accompanying notes 679–690.
681 See infra text accompanying notes 843–860.
683 Cornish & Llewelyn, supra note 18, at 113–14, 124 (explaining that the culture shock was due in part to subjecting national and EPO patents to the same substantive regimes, and also because there was no international patent system prior to 1977). In addition to the EPC, the Patents Act implemented the Patent Cooperation Treaty and the Community Patent Convention. Patents Act, c. 37, §130(7); Jacob, supra note 682, at 312.
684 Compare Patents Act, c. 37, §1(2) (excluding, among other things: discoveries; aesthetic creations; methods for performing mental acts, doing business, or a computer program; and the presentation of information), with EPC supra note 14, art. 52(2) (excluding: discoveries; aesthetic creations; methods for performing mental acts, doing business, and computer programs; and presentation of information).
what constitutes an invention. English courts have resorted to bypassing the Patents Act, instead directly interpreting the EPC. The English courts have also demonstrated a strong preference for construing the substance of a claim over its form. As a result, the United Kingdom’s approach to excluded subject matter is much stricter than that of the EPO. One scholar noted that a “UK-based applicant for a computer-implemented invention, if he is interested solely in the British market, would be well advised to apply for such a patent at the EPO level, designating the UK as the relevant jurisdiction.” This divergence in practice may be partially due to the nature of the legal system in the United Kingdom where stare decisis constrains judicial decision making. Although English courts make an effort to defer to EPO decisions, Technical Board of Appeal decisions are not binding. Rather, British patent examiners must follow English court decisions. Thus, unless the courts have expressly approved a Technical Board of Appeal decision, the decision will have only persuasive effect before the U.K. Intellectual Property Office (UKIPO). The differences in policy regarding software patent protection have resulted in a “deep rift” between EPO and British practice. The following section traces the development of the differing tests and standards in one of the most important patent-granting jurisdictions in Europe, and its convergence with and divergence from EPO standards.

---

686 Cornish & Llewelyn, supra note 18, at 144 (“When an issue arises which is governed by the UK counterpart of an EPC . . . provision within s.130(7), an English court will . . . treat the Convention provision directly.”).
687 See Jacob, supra note 682, at 312 (“There was, among the framers of the Convention, the notion that the British . . . read a claim in the abstract and ignored the specification.”).
688 Shemtov, supra note 7, at 514.
689 See id. at 510.
690 See id. (arguing that the Aerotel court decided to follow the 1986 EPO decision in Viacom, rather than the newer “trilogy” of PBS, Hitachi, and Microsoft, both to uphold long-standing English precedent and because they considered the approaches in the “trilogy” cases to be incorrect with regards to the language of the EPC).
691 See Richard Willoughby, United Kingdom, in GLOBAL PATENT LITIGATION: STRATEGY AND PRACTICE 1, 19 (Willem A. Hoyng & Frank W.E. Eijsvogels eds. 2006).
692 See Robert B. Franks, United Kingdom, at 18 (Supp. 2 Nov. 2008) in 1 SOFTWARE PATENT WORLDWIDE, supra note 437. The United Kingdom is composed of three jurisdictions: England and Wales, Scotland, and Northern Ireland. Willoughby, supra note 691, at 19. Most patent litigation takes place in the London-based Patents Court (High Court) and the Patents County Court. Id.
693 See Ballardini, supra note 7, at 569.
B. Early English Cases

Before implementing the EPC, English courts had shown a “distinct readiness” to allow patent claims for computer programs even if such claims did not affect the production of a distinct product. After the United Kingdom became a signatory of the EPC, its patent Court of Appeal cases initially tracked Technical Board of Appeals law. Nevertheless, the approaches have diverged.

Merrill Lynch’s Application decided a few years after Vicom, was the first English Court of Appeal case to address excluded subject matter under the Patents Act Section 1(2). Merrill Lynch claimed “an improved data processing based system for implementing an automated trading market for one or more securities.” The program automatically executed stock transactions against a customer’s orders using known data-processing equipment. Both the Patents Office and the Patents Court (High Court) rejected Merrill Lynch’s claim. The Court of Appeal agreed and by taking judicial notice of the EPO’s Vicom decision, the technical contribution approach part of U.K. precedent. The court interpreted the approach as requiring a technical advance over the prior art in the form of a new result. The court explained: “it cannot be permissible to patent an item excluded by Section 1(2) under the guise of an article which contains that item—that is to say, in the case of a computer program, the patenting of a conventional computer containing that program. Something further is necessary.” The court did not find the “something further,” like the substantially increased processing speed of the CAD program in Vicom.

The approach taken in Merrill Lynch indicated a preference for the substance of the patent claim over the claim’s form. In determining whether the claim was eligible for patent protection, the court considered both the nature of the invention and the nature of the result. Although a data processing system may be valid patent subject matter, if

---

694 Corinsh & Llewelyn, supra note 18, at 823 n.94.
696 Id. at 562.
697 Id. at 561.
698 Id. at 567 (stating that “[t]he decision of the board is a matter of which we are required, by section 91(1) of the Patents Act 1977, to take ‘judicial notice’”).
699 See id. at 569.
700 Id.
702 See id.
Patents Act Section 1(2) (c) excludes what it produces, the invention is not valid subject matter. Even if a computer program itself is non-obvious, Section 1(2) (c) excludes the claim as a whole unless it contains non-excluded subject matter which is also non-obvious and contains an inventive step. The excluded subject matter is not considered for purposes of establishing inventive step. The court found the Merrill Lynch claim to a “data-processing system . . . making a trading market in securities” to be an excluded business method. Commentators have criticized the English approach as “applying a convoluted decision process” of determining whether a purported invention is excluded subject matter, rather than applying section 1(2) as a self-contained filter. The methodology adopted in Merrill Lynch has led to what some commentators describe as an “undue level of exclusion” in the U.K. patent system.

The English courts consistently employ the approach of disallowing software inventions cloaked as machines or technical inventions. In 1991, the English Court of Appeal in Gale’s Application addressed whether a ROM containing a computer program is valid subject matter. The claim described the invention as an improved iterative algorithm for computing a square root stored on the Read-Only-Memory (ROM) of a computer. The applicant argued that the characteristic distinguishing the claim from a pure abstract idea was its storage in the electronic circuitry of the computer. The examiner rejected the application as excluded subject matter because it was a computer program “as such.” The Patents Court reversed, however, on the grounds that the claim was not disqualified under Section 1(2) because the claim related to a new technical product. Distinguishing between a non-patentable program loaded on a disk from a program loaded on the ROM, the court reasoned that “[t]here is a difference between a claim which relates to a disc containing a program and a ROM with

---

703 See id.
704 See Franks, supra note 692, at 41.
705 See id.
707 Franks, supra note 692, at 40.
708 Cornish & Llewellyn, supra note 18, at 826.
711 See id. at 308.
712 See id. at 308–09.
713 Id. at 306.
particular circuitry.”714 According to the court, the key difference was that the disk carried the program, whereas the programmed ROM’s structure was altered by the program such that it became a “dedicated piece of apparatus.”715

The Court of Appeal did not agree. It characterized a “program for a computer” as “essentially a series of instructions capable of being followed by a cpu to produce a desired result.”716 The Court of Appeal accurately recognized that the disk and ROM were merely different kinds of artifacts on which a program may be carried.717 Comparing programs on these media to different pieces of music loaded onto compact disks, the court found the differences in storage media immaterial for purposes of determining patent eligibility.718 The music was the same regardless of the chosen storage media.719 Similarly, the instructions stored on a disk or ROM were also the same.720 The court noted that deciding otherwise would exalt “form over substance.”721

Following the reasoning of Vicom, the English Court of Appeal questioned whether the instructions contained on the ROM include more than disqualified subject matter.722 The court found Gale’s claim not eligible subject matter because “the claim is in substance a claim to a computer program, being the particular instructions embodied in a conventional type of ROM circuitry, and those instructions do not represent a technical process outside the computer or a solution to a technical problem within the computer.”723 Even though the program arguably made the computer more efficient, the Court of Appeal rejected the claim because it only provided the CPU with a different set of calculations for determining a square root.724 The program was not valid patent subject matter because it did not define a new way of operating the computer.725 Therefore, the claim was to the instructions, a computer program “as such.”726 Both Merrill Lynch and Gale’s Application illustrate the English courts’ early attempts to avoid the problem of clever draft-

714 Id. at 316–17.
715 Id. 317.
717 Id. at 325.
718 Id.
719 See id.
720 See id.
721 Id.
723 Id. at 328.
724 See id. at 327–28.
725 See id.
726 See id.
ing to avoid exclusion, by directly addressing the issue of substance over form in determining the nature of the invention claimed.\textsuperscript{727}

In 1997, the English Court of Appeal heard a case involving a patent granted in Japan and submitted to the UKIPO as a matter of priority.\textsuperscript{728} The invention in \textit{Fujitsu} involved both a “method and apparatus” for modeling synthetic crystal structures used for designing semiconductors and superconductors.\textsuperscript{729} The court noted that ordinarily a person would assemble plastic models of the structures by hand to model the new crystal structure.\textsuperscript{730} The claimed invention used a computer programmed to allow a human operator to “select an atom, a lattice vector and a crystal face in each of two crystal structures.”\textsuperscript{731} The program converted the data representing the two crystal structures into data representing the physical layout of the combined structure and a pictorial display of the new structure.\textsuperscript{732} Both the UKIPO and the trial court rejected the application as a computer program and “a method for performing a mental act” excluded under Section 1(2) of the 1977 Patents Act.\textsuperscript{733} The Court of Appeal, applying the technical contributions approach, also rejected the application.\textsuperscript{734} Following English practice of construing the invention as a whole, the court concluded as follows: “[c]learly the whole operation revolves around the computer program and the question for decision is whether there is a technical contribution so that it cannot be said that the invention consists of a computer program as such.”\textsuperscript{735}

The result—rejecting an “invention” that used computer imaging to make the process of assembling crystal structures faster—seemed to contradict the EPO’s \textit{Vicom} decision. The Court of Appeal in \textit{Fujitsu} distinguished the computer imaging found patentable in \textit{Vicom} by the way it enhanced the image produced.\textsuperscript{736} The court noted that the only advance made by the invention was to “enable[] the combined structure to be portrayed quicker.”\textsuperscript{737} The operator was still required to produce two displays of the crystal structures and the appropriate way for

\textsuperscript{727} See \textit{id}. at 315.
\textsuperscript{728} Fujitsu Ltd.’s Application, [1997] R.P.C. 608 (A.C.) 610 (Eng.).
\textsuperscript{729} \textit{id}. at 612.
\textsuperscript{730} \textit{id}. at 610.
\textsuperscript{731} \textit{id}. at 612.
\textsuperscript{732} \textit{id}. at 610; see Patents Act, c. 37, § 1(2).
\textsuperscript{734} \textit{id}. at 618.
\textsuperscript{735} \textit{id}. at 618–19.
\textsuperscript{736} \textit{id}. at 619.
them to be superimposed. The Court of Appeal found that Fujitsu’s use of computer imaging simply made the computer function more rapidly, but did not change the computer’s ordinary function. The court concluded that a claim to a method for carrying out a calculation, or a method of performing a mental act, cannot be valid subject matter merely because the process is completed on a computer unless there is a technical contribution present. The fact that the computer performed the operation more quickly is not sufficient.

During the first decade of English jurisprudence regarding computer programs, English courts tracked the jurisprudence of the EPO. By requiring the invention to be construed as a whole, however, English application of the technical contribution test resulted in a stricter review of claims, and thus fewer patent grants. Of the three English Court of Appeal cases applying Patents Act Section 1(2), none found the claimed invention to have sufficient technical character to fall outside the exclusions. The different results from the English Court of Appeal in Fujitsu and the EPO Technical Board of Appeal in Vicom are particularly difficult to reconcile. Both cases used a computer program that caused computer images to be completed more rapidly and efficiently than could be done manually.

C. Recent Jurisprudence (Divergence and Convergence)

Nearly a decade had passed since Fujitsu when the English Court of Appeal issued its next EPC Articles 52(2) and (3) patent subject-matter decision. During that period, EPO jurisprudence had re-

738 Id. at 619–21.
739 See id. at 621.
741 See id.
742 See Ballardini, supra note 7, at 568.
743 See id.
745 See Fujitsu, [1997] R.P.C. at 618 (attempting to distinguish the facts of Vicom from the facts of Fujitsu in order to reach a different outcome).
747 Aerotel Ltd. v. Telco Holdings Ltd., [2006] EWCA (Civ) 1371, [75–76], [2007] 1 All E.R. 225 (A.C.) at 229 (Eng.); Fujitsu, [1997] 114 R.P.C. at 608. In Aerotel, Justice Jacob referred directly to the EPC, rather than to the similar language that had been implemented in the U.K. Patents Act 1977. See Aerotel, [2007] 1 All E.R. at 230. Various reasons contributed to this decision: differences in wording could lead to erroneous construction; the EPO Technical Board of Appeal decisions have strong persuasive authority; and the
jected Vicom’s “technical contribution approach” in favor of the “any hardware approach” of PBS, Hitachi, and Microsoft. In 2006, a Court of Appeal decision rejected the EPO’s “any hardware” approach in Aerotel, affirming the rupture between U.K. and EPO practices which began in 2002 with PBS. Aerotel was actually two cases joined together on appeal: Aerotel Ltd. v. Telco Holdings Ltd. and Re Macrossan’s Application. In the first case, Aerotel sued Telco for infringing its U.K. patent on a telephone system that provided prepayment for telephone calls. Telco counterclaimed for revocation on the basis that the invention was excluded as a method for doing business. The trial court agreed with Telco and revoked the patent. Aerotel appealed. In the second consolidated case, Macrossan, the court considered a computerized method of obtaining the forms needed for incorporating a company. The examiner rejected the application as a method of doing business. The High Court affirmed that merely automating a general purpose computer to produce documents necessary to incorporate an entity, where there is no underlying technical improvement, was not valid subject matter because the essence of the invention was the automation of a mental act. Macrossan also appealed.

Despite acknowledging the weight properly placed on EPO board decisions, the Court of Appeal declined to follow the trilogy of PBS,
Hitachi, and Microsoft, characterizing them as the “‘any hardware approach.’”759 In a lengthy opinion and appendix, Justice Jacob highlighted the differences and inconsistencies of the approaches articulated in those cases.760 Key in his criticism was that the EPO cases treated the various categories of Article 52(2) exclusions as being limited to “something abstract or intangible.”761 The English Court of Appeal disagreed, noting that the categories “are disparate with differing policies behind each.”762 Taking computer programs as an example, the court observed that the trio of cases takes a “narrow view” of the Article 52 computer program exclusion in only excluding abstract sets of instructions.763 The court implicitly found that such a broad definition of valid subject matter encompassed the instructions on a disk or hard drive which “causes a computer to execute the program.”764 The court determined that the framers meant to exclude computer programs “in a practical and operable form . . . not just an abstract series of instructions.”765 Thus, the court declined to adopt the EPO’s narrower view of the exclusion into English law.766

The court went on to emphasize that it was bound by its own precedents and obligated to follow the technical contributions approach from Vicom as interpreted in previous English Court of Appeal cases.767 Justice Jacob synthesized the English approach, which he labeled the “technical effect approach with the rider,” as a structured four-step approach to analyzing claims. The steps include: “(1) properly construe the claim; (2) identify the actual contribution; (3) ask whether it falls solely within the excluded subject matter; (4) check whether the actual or alleged contribution is actually technical in nature.”768

Applying the four-step approach to the Aerotel claim, the court reversed the trial court and found for the patentee.769 In construing the system claim, the court found that the system was “actually a claim to a physical device consisting of various components.”770 Although the in-

760 See id. at 238, 254–63.
761 Id. at 238.
762 Id.
763 Id.
764 Id.
765 Aerotel, [2007] 1 All E.R. at 238.
766 Id.
767 Id. at 236–37, 239.
768 Id. at 239–40.
769 Id. at 241–43.
770 Id. at 242.
vention used conventional telephone exchanges, the patentee added an extra piece of equipment called a “special exchange.” The actual contribution made was a new system requiring a new physical combination of hardware—more than a method of doing business. Aerotel’s use of hardware, even though it was known digital communications exchange hardware, provided the technical contribution. Justice Jacob stated, “it is true that it could be implemented using conventional computers, but the key to it is a new physical combination of hardware. It seems to us clear that there is here more than just a method of doing business as such.”

By contrast, Macrossan’s invention did not fare as well. The court found the invention ineligible for patent protection because it was both a method of doing business as such and a computer program as such. In applying the third step, determining whether the claim’s contribution was to excluded subject matter, the claim failed. Justice Jacobs reasoned that “Mr. Macrossan’s method is for the very business itself, the business of advising upon and creating appropriate company formation documents.” The court thus rejected Macrossan’s method as a quintessential business method. The court distinguished this result from Aerotel, in which a free standing device implemented the business method. The court found that the contribution under step two—providing a computer program, or interactive website, to carry out the method—was a contribution exclusively to excluded matter, and therefore was not a technical contribution. Additionally, the claim failed the fourth step in that there was no technical contribution “beyond the mere fact of the running of a computer program.”

The divergence from EPO approaches in Aerotel may be due to weak logic and inconsistencies in EPO jurisprudence rather than U.K. idiosyncrasies. One scholar suggests that the primary reason the Aerotel court chose to follow Vicom was not strict adherence to English prece-

772 Id.
773 See id.; Leith, supra note 6, at 151.
774 Aerotel, [2007] 1 All E.R. at 242 (emphasis added).
775 Id. at 245 (disagreeing in part with the trial court that found the claim to a method of performing mental acts “as such,” not a method of doing business, and a computer program “as such”).
776 Id. at 245–47.
777 Id. at 247.
778 Cornish & Llewelyn, supra note 18, at 828.
779 See id.
780 Aerotel, [2007] 1 All E.R. at 247.
781 Id.
dent, but because the court found none of the EPO “any hardware” approaches valid in light of EPC text. On this reading, the Court of Appeal found it impossible to reconcile PBS, Hitachi, and Microsoft. Although PBS had addressed and rejected the method claim on fairly conventional grounds, the apparatus claim signaled a “substantial departure from previous case law” in holding “that a computer programmed to carry out the unpatentable method was not within the categories of art[icle] 52(2).” In Hitachi, the Board of Appeal held that the apparatus claim was neither a business method nor a computer program as such because it “comprise[d] clearly technical features, such as a ‘server computer’, ‘client computers’ and ‘a network.’” According to Justice Jacob, the Hitachi logic “most dramatically articulates the departure from earlier [EPO] reasoning—a computer when programmed to conduct a business method is not excluded by Art[icle] 52(2).”

The Aerotel court was also highly critical of treating excluded matter as part of the prior art. The court used an example outside the context of computer programs and business methods to illustrate the point: “Consider for instance . . . a claim to a book . . . containing a new story the key elements of which are set out in the claim.” Justice Jacob characterized deeming the story part of the prior art by applying the PBS or Hitachi case reasoning as “simply not intellectually honest.” He also seemed “puzzled” as to why the EPO rejected applications for non-compliance with Article 56 particularly when doing so led to the same outcome as applying the test from Vicom at the stage of Article 52(2) analysis. The court criticized Microsoft on a number of points, but in particular targeted its very narrow definition of excluding computer programs as “just the abstract set of instructions” rather than a broader view that the term covers instructions on any medium that causes the computer to execute the program. Justice Jacob found this result inconsistent with decisions in both the United Kingdom and the EPO, and stated that this result would “seem to open the way in practice to the patentability in principle of any computer pro-

782 Shemtov, supra note 7, at 510.
784 Id. at 258.
785 Id.
786 Id. at 237.
787 Id.
788 Shemtov, supra note 7, at 510.
789 Aerotel, [2007] 1 All E.R. at 262.
gram.” 790 Hence, the court’s real concern with the any hardware approach was the elevation of form over substance. 791

Following Aerotel, patent practices employed by the EPO and the UKIPO diverged significantly. 792 The Aerotel court did not intend its decision to be a radical departure from prior case law. Nevertheless, in applying the four-part test, the UKIPO rejected most claims directed to computer programs, even if the claim would have been valid patent subject matter prior to Aerotel. 793 In summarizing the differing approaches between the UKIPO and EPO in a guide for patent practitioners, one scholar posits that the difference in current subject-matter treatments reveals different policy approaches to the subject-matter analyses. 794 In the United Kingdom, the policy enshrined in the third step of the Aerotel test requires that if the computer-related invention falls within an excluded category, then the examiner must reject the entire claim even if parts of the claim are novel, achieve an inventive step, and make a technical contribution. 795 By contrast, EPO policy favors granting patents for computer-related inventions if these requirements are met. 796 In practice, applicants for computer program patents in the United Kingdom favor the EPO because the probability of receiving a patent for a computer-implemented invention in the United Kingdom remains low. 797 But this divergence in approaches creates the possibility that a patent examined and granted by the EPO and registered as a patent in the United Kingdom could be invalidated under U.K. law. Since Aerotel, the EPO Technical Board of Appeal has issued several decisions reaffirming the approach in Hitachi. Duns Licensing 798 responded directly to the Aerotel court’s criticism of EPO case law and condemned the four-step approach as not “consistent with a good-faith interpretation of the European Patent Convention.” 799 In light of the increasing tension between U.K. and EPO law and practice, the English

790 Id.
791 Aplin, supra note 7, at 382.
792 See Ballardini, supra note 7, at 569.
794 See Franks, supra note 692, at 72.
795 See id.
796 See id.
797 See id.; see also Wallis, supra note 7, at 4.
799 See id. at 70.
Court of Appeal took a second case only two years after Aerotel. The composition of the appellate panel in Symbian Ltd v. Comptroller-General of Patents indicated the importance of the decision. Two of the three justices, Justice Jacob and Justice Neuberger, sat on the Aerotel panel. Lord Neuberger took the third spot on the panel as a special visitor to the Court of Appeal from the supreme judicial body of the United Kingdom, the House of Lords. The tone of the Symbian case, while markedly more deferential to European Patent Office precedent than Aerotel, ultimately did not change English patent law.

The claim in Symbian covered a computer program for a method of “[m]apping dynamic link libraries in a computing device.” Link libraries are a package of small programs relating to general computer functions. Providing libraries allows function programs to be called up when needed, rather than replicated by each computer program. For instance, when a word processing program needs to use a printing function, the function can be called up from the library so that it does not need to be included in the word processing program. Dynamic link libraries in the prior art could be either “linked by name” or “linked by ordinal” systems. The patent application claimed that Symbian Ltd.’s invention would avoid difficulties and potential unreliability of prior art linked by ordinal systems. The UKIPO denied the claim as a claim to a computer program “as such.” The English High Court ruled that by providing a technical contribution, the invention

---

801 See id. at 1; see also Symbian Legal Appeal Throws UK Software Patents into Confusion, CELLULAR-NEWS (Dec. 15, 2008), http://www.cellular-news.com/story/35181.php [hereinafter Symbian Legal Appeal] (noting Judge Lord Neuberger came down from the highest U.K. court, the House of Lords, expressly to hear this case).
804 Compare Symbian, [2009] R.P.C. at 14 (electing to follow established English precedent while politely declining to follow EPO decisions), and de Mauny, supra note 7, at 150 (noting that Symbian was dismissed, in part, to leave precedent standing), with Aerotel, [2007] 1 All E.R. at 238 (declining to follow EPO precedent without apology).
806 See de Mauny, supra note 7, at 148.
807 See id.
808 See id.
809 See id.
811 See id. at 6.
was not precluded from registration under Section 1(2) of the Patent Act and Article 52(2) of the EPC. The Court of Appeal affirmed.

The Court of Appeal began with a recitation of the statutory provisions and articulated its obligations to follow previous decisions as precedent. The court also noted that it had the freedom to depart from its previous decisions in the field of patent law if the EPO Board had formed a settled view on that point of law that differed from previous decisions, but that it was not bound to do so. The Court of Appeal concluded that “we should try to follow previous authority, we should seek to steer a relatively unadventurous and uncontroversial course, and we should be particularly concerned to minimise complexity and uncertainty.” Nevertheless, despite three EPO cases decided after Aerotel applying the test from Hitachi, the Court of Appeal declined to follow the approach. In part, the court chose to do so because the law still seemed to be in a state of flux; the Enlarged Board had not settled the issue, the post-Aerotel decisions by the EPO were inconsistent, and the German judiciary also expressed doubts about the “any hardware approach.”

Instead of following EPO cases, the Symbian court employed the “technical contribution approach” introduced in Gale’s Application. Under this computer-targeted approach, a computer program must be more than just a “better program to qualify as patent subject matter.” Something more is needed, for instance, a change in speed with which the computer works. The claim must also “solve a ‘technical’ problem

---

812 See de Mauny, supra note 7, at 149–50 (explaining the prior history of the case, the judge’s reluctance to grant permission to appeal to the Comptroller of Patents, and recognizing that the appeal was expedited due to the impact it would have on pending British patent applications).
814 See id. at 11–12 (“In principle the Court of Appeal is bound by one of its previous decisions unless that previous decision is inconsistent with a subsequent decision of the House of Lords . . . is inconsistent with an earlier Court of Appeal decision . . . or can be shown to have been arrived at per curiam (i.e. without reference to the relevant statutory provision or authority).”).
815 See id. at 12.
816 See id. at 16. The court reached its conclusion despite acknowledging, in the previous paragraph, that “the boundary between what is and is not a technical contribution is imprecise” and may not be soluble in a wholly satisfying way. Id. at 15.
817 See id. at 14.
818 See id.
820 See id. at 15.
821 See id. at 16–17.
822 Id. at 16.
lying within the computer itself.” The court concluded that a computer with the claimed program operated better than the prior art and was thus valid patent subject matter. The court considered its conclusion in light of the four-part test articulated in Aerotel and began with the second step because they had already sufficiently characterized the patent claim. The program’s actual contribution identified under the second step was that it made a computer operate faster and more reliably than the prior art by virtue of the claimed feature. Addressing the third step of the analysis, the court found that the claim was not solely to excluded subject matter because it included the “knock-on” effect of a computer working better. The court concluded that the invention was technical “on any view as to the meaning of the word technical.”

Despite acknowledging the inevitability of the EPO granting software and business method patents where the UKIPO would not, the Court of Appeal minimized the differences in their approaches. What differs, according to the court, is where the “technical” determination is made. In the United Kingdom it remains part of the Article 52 analysis, while in the EPO it is completed with reference to Article 56. The court emphasized “the strong desirability of the approaches and principles in the two offices marching together as far as possible,” concluding that “where there may be a difference of approach or of principle, one must try to minimize the consequent differences in terms of the outcome in particular patent cases.”

Achieving a common result is exactly what the Symbian court accomplished. The EPO Examining Division had already indicated it would grant a patent for Symbian’s invention. Had the Symbian court found the invention not valid subject matter under previous precedents and the Aerotel test, the conflict in approaches between the United

---

823 Id. (quoting Gale’s Application, [1991] R.P.C. at 328) (noting that the Symbian invention meets this requirement).
824 See id. at 17.
826 See id. at 16–17.
827 See id. at 17.
828 Id.
829 Id. at 17.
830 See id. at 7.
831 See PBS, [2001] O.J.E.P.O. at 456 (denying a patent because it did not meet the inventive step criterion defined in Article 56).
833 Id. at 18.
834 Id. at 17.
835 See id. at 17–18; Aerotel, [2007] 1 All E.R. at 240; de Mauny, supra note 7, at 151.
Kingdom and the EPO would have widened considerably. By ruling that software that improves the operation of a computer is valid patent subject matter, Symbian moderated the UKIPO practice of rejecting applications that did not have an external effect.\footnote{See Symbian, [2009] R.P.C. at 17; INTELLECTUAL PROP. OFFICE, PRACTICE NOTICES: PATENTS ACT 1977: PATENTABILITY OF COMPUTER PROGRAMS ¶ 3 (Dec. 8, 2008) available at http://www.ipo.gov.uk/pro-types/pro-patent/p-law/p-pn/p-pn-computer.htm [hereinafter IPO].} On December 8, 2008, the UKIPO issued a Practice Notice\footnote{See IPO, supra note 836, ¶ 1.} based on Symbian that confirmed the four-step test.\footnote{See Symbian, [2009] R.P.C. at 1.} The Practice Notice concluded that “a program that results in a computer running faster or more reliably may be considered to provide a technical contribution even if the invention solely addresses a problem in the programming.”\footnote{See IPO, supra note 836, ¶ 5.} Some commentators have concluded that Symbian will result in the UKIPO issuing more patents on computer-implemented inventions,\footnote{Wallis, supra note 7, at 4.} particularly because patent attorneys will make sure to emphasize the “knock-on” technical effects in patent applications.\footnote{Taylor, supra note 793, at 15.}

Although the U.K. and EPO results converged in Symbian, the saga is far from over.\footnote{See de Mauny, supra note 7, at 151; Renao Marchini, Patently Better, 83 EUR. LAW.14, 15 (2008).} The approaches to determining technicality still differ. What might be a technical solution to a technical problem under EPO Article 56 analysis will not necessarily constitute a technical contribution under U.K. Article 52(2) analysis.\footnote{See Shemtov, supra note 7, at 512.} Because a European patent may be challenged for validity in an English court,\footnote{Patents Act, c.37, § 74.} courts could find themselves in a dilemma if one of those instances presaged in the Symbian decision\footnote{Symbian, [2009] R.P.C. at 14–15.} occurred, in which the different approaches lead to divergent results. Further, at least one commentator describes the UKIPO Practice Notice as “a grudging concession” that computer-implemented inventions may be patented if there is no external effect.\footnote{See Symbian Legal Appeal, supra note 801; see also Wallis, supra note 7, at 4.} The Practice Notice explicitly states that “examiners will object to the computerization of what would be a pure mental act if done without the aid of a computer as both a mental act and a computer program as such.”\footnote{IPO, supra note 836, ¶ 8.} Nevertheless, it is possible the UKIPO may reject
applications based on the mental acts exclusion where it would have previously done so under the computer program exclusion.

D. German Patent Subject-Matter Treatment of Software and Business Method Claims

A robust analysis of German law and practice in the area of computer software and business methods is beyond the scope of this Article. Nevertheless, a brief discussion serves to highlight problems within the European patent community, accentuating the need for clear definitions of the EPC exceptions and an appellate body to resolve interpretive differences among the national jurisdictions.

The Bundesgerichtshof (BGH) has attempted to articulate an appropriate test for technicality in the context of computer-implemented inventions over the last decade. Despite the BGH’s repeated assertions that its decisions are in line with EPO case law, its interpretations may differ. Commentators have characterized Germany’s approach to software patents as more cautious than the EPO. As a result, the validity of EPO patents is frequently contested in infringement proceedings in the German Patent Court.

The BGH has developed a two-part test that appears roughly equivalent to the general contours of EPO practice, in which technicality is assessed as part of both Article 52 and Article 56 analyses. In German practice, the invention must possess technical character, which is assessed independently from novelty and inventive step criteria. This second step requires an objective technical contribution over the prior art and a solution to a technical problem.

---

848 See Hans Wegner, Germany, 12 in 1 Software Patents Worldwide (Supp.4 Dec. 2009), supra note 437. Although, as a civil law system, these cases do not have the precedential weight, examiners of the German Patent and Trademark Office typically refer directly to decisions of the Federal Patent Court and the BGH, giving these cases practical importance. See id. at 10.

849 Id. at 11. But see Aerotel, [2007] 1 All E.R. at 269 (noting that the BGH did not consider the computer program exclusion from EPO case law, and so neglected to take EPO case law into account).


851 See id.

852 See Wegner, supra note 848, at 6.

853 Id.

854 Id. at 6, 17 (describing a test for an objective technical contribution and summarizing BGH case law requiring a solution to a technical problem).
specific technical problem. The two criteria have changed in importance over the last several years. Currently the latter step is more important.

The April 2010 decision, *Dynamische Dokumentenverwaltung (Siemens)* is the most recent attempt to articulate the contours of German law. As one scholar points out, the fact that the BGH has decided two cases in 2009 and 2010 indicates that the Federal Patent Court is “trying to understand the boundaries” of patent law in this area. The Federal Patent Court determines which appeals to send to the BGH, and only certifies appeals if there is “a legal question of fundamental importance” or a BGH decision is “needed for consistent interpretation” of the law. Thus, having the BGH decide two cases in this time span is significant. Commentators and even legal practitioners are unable to agree on the impact of the *Siemens* decision. Whereas some commentators hail the *Siemens* case as a “landmark decision” that marks a break from past decisions and goes much further than EPO practice, other commentators view the decision as “a continuation of a long line of thinking by the German courts” that is very similar to the EPO approach.

Regardless of *Siemens*’s impact, looking to German decisions for a definition of “technical” is not fruitful. In 2000, the BGH expressly acknowledged that there is no objective definition of technology in the context of determining whether a particular claimed invention is technical. Therefore, “technical” has a dynamic meaning which can be

---


856 See id. at 186.


859 Id.


862 See Harrison, *supra* note 858.

treated differently in the context of technological developments. In extrajudicial comments, Judge Melullis of the BGH marginalized the word “technical” because “when assessing software as such, the program’s interdependence with the technical device makes the technical content too hard to deny.” Hence, in the context of computer-implemented inventions, the meaning of technical as used by German courts remains unclear.

**Conclusion**

In addressing Justice Ginsburg’s query about whether Europe provides a solution to the U.S. business method and software patent conundrum, our analysis clearly answers in the negative. As the *Bilski v. Kappos* petitioner stated with respect to Europe’s technology requirement, “technology . . . can be a difficult term.” Technology is a difficult term, particularly as it relates to whether and under what conditions business methods, computer programs, algorithms, and mental acts that are embodied in machines should receive patents as a matter of policy and practice. Instead of serving as a panacea for U.S. patent examiners and courts, the European technical requirement fails to provide a meaningful constraint for software patents and many business method patents on either side of the Atlantic.

Uncertainty and evolving standards characterize all patent systems examined in this Article. The United States treats all business methods and software as eligible patent subject matter, whereas the EPC, U.K., and German patent systems exclude all business methods that are not computer implemented. Beyond these basic observations, ambiguity about the nature and extent of patent subject matter permeates all these patent systems. *Bilski* injected significant uncertainty in a relatively settled area of law by rejecting *State Street Bank*’s “useful, concrete, and tangible result” test and inviting the CAFC to develop a new subject-matter test based on the abstract ideas exclusion. In Europe, the certainty of the business method and software exclusions vanishes where patent claims include a computer or software component. For such applications, European patent courts analyze the technicality of the com-

---

864 See Ballardini, supra note 7, at 572.
867 See supra Parts II–IV.
ponent, where there is no commonly recognized definition of technical.869 Thus, in the United States uncertainty of subject-matter exclusion exists for business methods only, whereas in Europe business methods and software are excluded unless they constitute parts of a mixed claim involving “something more.”870

It is this “something more” that is so difficult to characterize and that results in uncertainty. In the United Kingdom, the analysis of technicality is part of the subject-matter test, where English courts require a “technical contribution” or external “knock on” effect of a computer working better.871 The closest U.S. analog to the U.K. external effect requirement would be a weak physical transformation test.872 German courts and the EPO have moved the computer-implemented business method and software technicality determination to their inventive step or obviousness analysis.873 This test asks whether the invention provides a technical solution to a technical problem.874 This appears somewhat analogous to the now-discredited test proposed in Parker v. Flook, which required examiners to evaluate the novelty and non-obviousness contributions of an invention only after completely discounting any contribution due to a mathematical algorithm.875 In order to provide a closer analogy to tests used in Europe, the Flook test would need to treat the business method component of patent claims as contributions that are well known in the art.876

The European test—pejoratively labeled the “any hardware test”—has evolved to bar naked business methods and little else.877 Like the U.S. pattern, where incrementally more types of inventions receive sub-

869 See supra text accompanying notes 359, 861; see also CFPH LLC’S Application, [2005] EWHC (Pat) [14], [2006] R.P.C. 259, 267 (Eng.) (“[T]he word ‘technical’ is not a solution. It is merely a restatement of the problem in different and more imprecise language.”).

870 See, e.g., Case T-208/84, Computer-Related Invention/VICOM, [1987] O.J.E.P.O. 14, 21–22 (Technical Bd. Appeal 3.5.01, July 15, 1986), available at http://archive.epo.org/epo/pubs/cj1987/p001_046.pdf (reasoning that a traditionally patentable technical process involving computer software and a computer apparatus should not be excluded solely by virtue of the fact that software is part of the claim, thereby suggesting that the hardware component of the claim was “something more”).

871 See supra text accompanying notes 820–828.

872 See In re Bilski, 545 F.3d 943, 962–63 (Fed. Cir. 2008), aff’d sub nom. Bilski, 130 S. Ct. 3218.

873 See Wegner, supra note 848, at 6.

874 See id.; Melullis, supra note 855, at 186.


876 See id.

877 See Aerotel Ltd. v. Telco Holdings Ltd., [2006] EWCA (Civ) 1371, [75–76], [2007] 1 All E.R. 225 (A.C.) at 237 (Eng.); Steinbrenner, supra note 437, at 66.
ject-matter treatment, following the 2006 EPO decision in Microsoft even computer programs are patentable provided they achieve a further technical effect. The European analogy to U.S. patent law development differs in two significant respects. First, there is no evidence that European patent law will continue expanding patent protection to business methods. No case in any of the examined jurisdictions has granted patent protection to “naked” business methods or business methods “as such.” Second, English patent courts appear only grudgingly to acquiesce to the EPO practice of granting patents on computer programs that lack an external effect. Hence, whereas European patent reviewing bodies are in consensus about business methods, there is considerable uncertainty surrounding the boundaries of software patentability—and likely significant discord between patent jurisdictions over the proper treatment of software patents. Thus, other than Europe’s unified support for its business method exclusion—an exclusion that the Supreme Court in Bilski refused to embrace—Europe has little to offer the United States that can enhance clarity and certainty in U.S. patent subject-matter practice.

Unfortunately, the U.S. approach to business method and software subject-matter review could use an infusion of outside guidance. Rather than increase clarity, Bilski has increased ambiguity and uncertainty about what inventions constitute eligible subject matter. By rejecting the machine or physical transformation test as the threshold requirement for patentability and likely rejecting the CAFC’s useful, concrete and tangible test, the USPTO and practitioners are left with no test—other than the abstract ideas exclusion—to provide guidance. Bilski provides neither a clear rule nor a prohibition on the patenting of naked business methods unless they constitute abstract ideas. Further, as the Kennedy plurality observes, the suggestion that the “machine-or-transformation test is useful ‘for evaluating processes similar to those in


879 See supra Parts III–IV; see also Bakels & Hugenholtz, supra note 6, at 21–22 (“Too many European observers business method patents represent a horrific prospect—yet another example of unwanted ‘Americanisation’. Even those who are in favour of software patenting usually are vehemently opposed to patenting business methods.”).

880 See supra text accompanying note 846.

881 See 130 S. Ct. at 3229.

882 See id. at 3225.

883 See id. at 3232 (Stevens, J., concurring).

884 See id. at 3229 (majority opinion) (“In searching for a limiting principle, this Court’s precedents on the unpatentability provide useful tools.”).
the Industrial Age,’ but is less useful ‘for determining the patentability of inventions in the Information Age’...” indicates that the courts may employ different subject-matter tests for new technologies.885

This Article contends that such movement is in error. Information-age technology is not qualitatively different from industrial-age technology. It still must qualify as either a process, machine, manufacture, or composition of matter.886 Pre-information-age patent jurisprudence is capable of providing the clarity sorely lacking in current U.S. patent jurisprudence. Rather than rendering cautious decisions based on future unimaginable technological developments, U.S. courts should rely on legislative bodies to address gaping legal deficiencies with respect to new technologies if such technologies develop.

Ironically, both the EPO Enlarged Board of Appeal and the U.S. Supreme Court fail to provide or increase clarity concerning the subject-matter treatment of software and business method patents in their recent opinions.887 What prevents the finding of a consistent approach to patent subject-matter determinations is the absence of any significant policy guidance in a highly politicized arena. This absence has been particularly problematic in Europe due to the plethora of national patent systems governed predominantly by a civil-law tradition that does not adhere to the doctrine of stare decisis. The result has been a multiplicity of approaches to the software patent subject-matter question in the EPO and national courts, with no clear definition of what is “technical” over the last twenty years.888 The absence of sufficient political willpower to provide policy guidance is evident from the failure of the European Union to implement a directive harmonizing the treatment of computer-implemented inventions. Moreover, the inability to remove “programs for computers” from the list of excluded subject matter in EPC 2000, despite years of granting patents for computer software products, further evidences this inertia.889

885 Id. at 3235 (Stevens, J., concurring) (describing the plurality's suggestion).
887 See Bilski, 130 S. Ct. at 3228–30 (declining to adopt general principles of patent protection and issuing a narrow holding for the case at hand); Case T-154/04, Estimating Sales Activity/DUNS LICENSING ASSOCS., [2008] O.J.E.P.O. 46, 66 (Technical Bd. of Appeal 3.5.01, Nov. 15, 2006), available at http://archive.epo.org/epo/pubs/oj008/02_08/02_0468.pdf (“Thus it will remain incumbent on office practice and case law to determine whether subject-matter claimed as an invention has a technical character.”); see also Ballardini, supra note 7, at 563.
888 See Ballardini, supra note 7, at 567.
889 Cf. EPC 2000, supra note 14, art. 52 (maintaining computer program exception in the amended provision).
By contrast, changes in the U.S. approach have spanned a much greater time period and have not suffered the European problem of conflicting contemporaneous treatments of the patent subject-matter question. The U.S. constitutional mandate is to grant patents that promote the useful arts.\textsuperscript{890} The legislative mandate includes patent protection for “new and useful processes.”\textsuperscript{891} No inventions or advances are explicitly excluded under U.S. law. This approach differs sharply from the EPC. EPC Article 52 explicitly excludes business methods and computer programs from patent subject matter.\textsuperscript{892} Nonetheless, the Article 52 exclusion includes the cryptic “as such” modifier.\textsuperscript{893} The meaning of this terse phrase is the root of the controversy in Europe.

In the absence of legislative guidance, U.S. common law developed a seemingly sound approach to dealing with subject-matter issues for patent process claims. This case law excluded claims for abstract ideas, laws of nature, and mathematical expressions and algorithms.\textsuperscript{894} The machine or physical transformation test was a product of this case law.\textsuperscript{895} Due to the absence of clear standards, however, the U.S. Supreme Court has been hesitant to extend these principles to new, and ostensibly different, information-age technology. In the past, this reticence facilitated the expansion of patent subject matter to include all new inventions regardless of application or type.\textsuperscript{896}

In \textit{Bilski}, the Court also failed to provide clarification or guidance in this area, by “not commenting on the patentability of any particular invention, let alone holding that any of the above-mentioned technologies from the Information Age should or should not receive patent protection.”\textsuperscript{897} Whether or not the Court’s reticence is warranted, its reluctance to act emphatically is likely due to the legislative nature of the requested decision. The task of drawing patent subject-matter boundaries is a policy decision that the courts have been uncomfortable making. In \textit{Bilski}, the Supreme Court continued its practice of asking Con-

\textsuperscript{890} See U.S. Const. art. I, § 8, cl. 8.
\textsuperscript{891} See, e.g., Diamond v. Chakrabarty, 447 U.S. 303, 308 (1980).
\textsuperscript{892} See EPC 2000, supra note 14, art. 52.
\textsuperscript{893} See id. art. 52(3).
\textsuperscript{894} See \textit{Bilski}, 130 S. Ct. at 3225 (acknowledging the existence of only three exclusions from patent subject matter: laws of nature, physical phenomena, and abstract ideas) (citing \textit{Chakrabarty}, 447 U.S. at 309); \textit{see also} Gottschalk v. Benson, 409 U.S. 63, 71–73 (1972) (finding mathematical formulas and algorithms to be non-patentable).
\textsuperscript{895} See, e.g., Cochrane v. Deener, 94 U.S. 780 (1876).
\textsuperscript{896} See \textit{Bilski}, 130 S. Ct. at 3225 (acknowledging “broad patent-eligibility principles” with relatively few exceptions).
\textsuperscript{897} Id. at 3228.
gress to address “the great challenge in striking the balance between protecting inventors and not granting monopolies over procedures that others would discover by independent, creative application of general principles.”

Thus far, Congress has not responded to such entreaties. In fact, other than adding a section to the Patent Act that provides a prior-use defense for business method patents, Congress has not addressed business method or software patent claims at all.

Europe has lagged in this seemingly inexorable march to remove most restrictions on patent subject matter, but may not be far behind. The trend in Europe is toward recognizing claims as valid patent subject matter as long as they are computer implemented. The United Kingdom—with its strong common law tradition of limiting patent subject matter—has offered the most resistance to this trend, but may be forced to yield under pressure to make its patent law consistent with EPO practice. It is telling that the Enlarged Board of Appeal recently ruled that there is no conflict in patent subject-matter treatment under the EPC, despite multiple Technical Board of Appeal decisions with seemingly conflicting approaches. First, nothing but the elusive “technical” requirement is left to prevent EPO case law from continuing its evolution in the direction of the more lax U.S. approach to patent subject-matter treatment. Whereas Europe is unlikely to adopt the complete absence of restrictions characteristic of U.S. patent practice, the dissipation of any remaining restrictions on software patents in Europe is a real possibility. Second, it is clear that the United States will not find any answers to its software and business method patent conundrum from the European patent system.

The curious reference to In re Bilski in the Enlarged Board of Appeal decision suggests that the EPO may have been looking to the United States for leadership and that it approved an approach that demonstrated a movement toward harmonization. The type of judicial leadership that Europe would most likely accept is an unequivocal policy statement that provides concrete and practical restrictions to business method and computer-implemented inventions. Bilski did not pro-

898 Id.
899 35 U.S.C. § 273(b)(1) (2006); see Hill & Cangolosi, supra note 366, § 7 (concluding that there have been few legislative amendments to U.S. patent law since 1952).
900 See Grosche, supra note 359, at 273–74; Thomas & DiMatteo, supra note 1, at 21–23.
901 See Aplin, supra note 7, at 379–80; Shemtov, supra note 7, at 510.
903 See id. at 16.
vide such a clear elucidating standard.\textsuperscript{904} It is thus unlikely that European courts will consider anything in the decision worth emulating.

U.S. courts, including the \textit{Bilski} Court, erroneously suggest that case law is inadequate to deal with rapidly changing technology.\textsuperscript{905} The Supreme Court’s final two statutory subject-matter decisions of the twentieth century ruled on cutting edge patent claims. In \textit{Diamond v. Diehr}, the Court validated a computer-controlled rubber curing process.\textsuperscript{906} In \textit{Diamond v. Chakrabarty}, the Court validated a patent for a genetically engineered microorganism capable of oil decomposition processes.\textsuperscript{907} Although the technology in the two cases was pioneering, both claims fit into categories that were clearly envisioned during the time of the first Patent Act. The \textit{Diehr} claim covered an industrial process\textsuperscript{908} and the \textit{Chakrabarty} claim covered a manufacture,\textsuperscript{909} two types of innovation that are listed in Section 101 of the Patent Act.\textsuperscript{910} By contrast, the alarm limit and BCD processes invalidated in \textit{Flook} and \textit{Benson}, respectively, would not qualify as industrial processes.\textsuperscript{911} \textit{Bilski} is in line with these results because the machine-or-transformation test remains a “useful and important clue” to patentability. The petitioners’ claims to the concept of “hedging,” and its reduction to a mathematical formula, were unanimously invalidated.\textsuperscript{912} Nevertheless, the plurality refused to draw a clear line for a case more difficult than \textit{Bilski}.\textsuperscript{913}

The Supreme Court missed the opportunity to bring U.S. patent law closer to that of the EPO by failing to resurrect the moribund business method patent exclusion that the CAFC nullified in \textit{State Street Bank}.\textsuperscript{914} Nevertheless, in holding that the machine-or-transformation test is “a useful and important clue” to the patentability of processes, and not endorsing the \textit{State Street Bank} test,\textsuperscript{915} the Court moved toward

\textsuperscript{904} See \textit{Bilski}, 130 S. Ct. at 3228.
\textsuperscript{905} See id. at 3227; see also \textit{Benson}, 409 U.S. at 72–73.
\textsuperscript{906} \textit{Diamond v. Diehr}, 450 U.S. 175, 177 (1981).
\textsuperscript{907} See \textit{Chakrabarty}, 447 U.S. at 305.
\textsuperscript{908} See \textit{Diehr}, 450 U.S. at 184.
\textsuperscript{909} See \textit{Chakrabarty}, 447 U.S. at 309.
\textsuperscript{911} See \textit{Flook}, 437 U.S. at 594–95; \textit{Benson}, 409 U.S. at 71–72.
\textsuperscript{912} See \textit{Bilski}, 130 S. Ct. at 3231, 3232 (Stevens, J., concurring), 3257–58 (Breyer, J. concurring). The majority held that the claim was not patentable because it represented an abstract idea, but Justices Stevens and Breyer would have held that the method was not a process, and therefore not patentable. \textit{Id.} at 3231–32, 3257–58.
\textsuperscript{913} See id. at 3231 (Stevens, J., concurring).
\textsuperscript{914} See id.; \textit{State St. Bank} & Trust Co. v. \textit{Signature Fin. Grp. Inc.}, 149 F.3d 1368, 1375 (Fed. Cir. 1998).
\textsuperscript{915} See \textit{Bilski}, 130 S. Ct. at 3227; \textit{State St. Bank}, 149 F.3d at 1375.
limiting the patentability of naked business methods. Still, a more clearly delineated test that could provide guidance in both the United States and Europe would be far better than the *Bilski* outcome.

This Article suggests that judicial bodies in the United States and Europe have weakened statutory subject-matter standards in favor of protecting business methods and software patent claims, due to the lack of clear policy mandates. Even the United Kingdom, perhaps the last holdout against the movement in favor of expansive coverage, has started to relent under pressure to be consistent with EPO practice. The U.S. Supreme Court had the opportunity to institute legal standards that provide welfare-enhancing innovation incentives and reduce the anti-competitive effects from broad recognition of business method and software patents. Although the Court recognized the need to balance protecting inventors with protecting the community against government-sanctioned monopolies, the Court declined to indicate “where that balance ought to be struck.” Unfortunately, the United States Supreme Court missed the opportunity to create clear limits to statutory subject matter. Doing so may have encouraged Europe to reform its patent policy, and perhaps brought both sides of the Atlantic closer to a truly harmonized patent policy.

---

916 *Bilski*, 130 S. Ct. at 3228.

THOMAS K. CHENG*

Abstract: This Article undertakes a comparative study of corporate veil piercing doctrines under U.S. corporation and English company law. The Article highlights some fundamental differences between the doctrines in terms of jurisprudential approaches, treatment of specific case types, and other related issues. The Article demonstrates that despite these substantial differences, many English corporate veil cases in fact share a similar analytical approach to the instrumentality doctrine under U.S. law. Therefore, it is possible to construct an English instrumentality doctrine that will bring structure and clarity to the English corporate veil doctrine. The Article concludes with a revival of the much-maligned single economic unit theory, first propounded by Lord Denning. This reformulated theory will provide a more systematic approach to veil piercing cases involving corporate groups.

INTRODUCTION .......................................................... 330
I. A BRIEF HISTORY OF THE ENGLISH CORPORATE VEIL DOCTRINE .... 334
II. SOME GENERAL COMPARISONS BETWEEN THE TWO CORPORATE VEIL DOCTRINES ................................................................. 342
   A. Jurisprudential Approaches .................................................. 343
      1. General Deference to the Separate Corporate Personality Principle .............................................. 343
      2. Reliance on Traditional Common Law Concepts ........... 346
      3. Absence of Overarching Theory ................................. 348
      4. The Role of Policy Considerations ............................... 350
      5. The Role of Justice......................................................... 353
      6. Summing Up .................................................................. 355
   B. Treatment of Specific Types of Cases ............................... 357
      1. Fraud and Misrepresentation ........................................ 357
      2. Tort Cases .................................................................... 361

* © 2011, Thomas K. Cheng. The author is an assistant professor in the Faculty of Law of the University of Hong Kong. He received his B.A. from Yale College, his J.D. from Harvard Law School, and his B.C.L. degree (equivalent of LL.M.) from University of Oxford.
Introduction

The corporate veil doctrine—and the related rules of separate corporate personality and limited liability—has been a much-studied subject in corporation law since the early part of the twentieth century. A perennial challenge facing the corporate veil doctrine has been the attempt to increase its predictability. The doctrine—being an exception to the general rule of limited liability—was created to prevent in-
As such, the application of the doctrine has always been fact-specific and open-ended. Justice requires a flexible legal standard that allows room for the weighing of equity and policy considerations. Some degree of open-endedness is thus probably inherent in the doctrine. It is perhaps because of this open-endedness that the corporate veil doctrine has remained fertile ground for academic research.

On the opposite side of the Atlantic, the corporate veil doctrine under English company law has similarly attracted considerable academic attention. Although the academic discussion in the United States has focused on making sense of the voluminous, and often contradictory, case law—the importance of the doctrine having been firmly accepted in the United States—the debate in the United Kingdom has taken on a different tone. Although some commentators have argued in favor of the doctrine, the general perception is that English courts are loathe to apply the doctrine. Only under exceptional circumstances is veil piercing permitted. Judicial attitude toward the doctrine, however, has not always been unaccommodating. Until the late 1970s, English courts demonstrated considerable willingness to pierce the veil when justice so required. Most notably, Lord Denning pronounced the single economic unit theory, which allows a court to treat a

---


4 See, e.g., Moore, supra note 2, at 182–83; Murray A. Pickering, The Company as a Separate Legal Entity, 31 Mod. L. Rev. 481, 481 (1968); Michael Whincup, Inequitable Incorporation—the Abuse of a Privilege, 2 Company Law 158, 159–60 (1981). The equivalent concept of a corporation under English law is called a company. See, e.g., Whincup, supra. This Article will employ both terminologies, using corporation when referring to the corporate entity in the United States and company when referring to the corporate entity in the United Kingdom. In the United Kingdom, the corporate veil doctrine is generally called lifting of the corporate veil as opposed to piercing of the corporate veil. For simplicity, this Article will use piercing of the corporate veil to encompass both.

6 See Douglas & Shanks, supra note 1, passim.

7 See Robert B. Thompson, Piercing the Corporate Veil: An Empirical Study, 76 Cornell L. Rev. 1036, 1039 (1991) (observing that limited liability and separate corporate personality have been features of U.S. corporate law in many jurisdictions since the mid-19th century).

8 Compare Gallagher & Ziegler, supra note 3, at 293 (noting the importance of the doctrine in preventing injustice), with Moore, supra note 2, at 181 (arguing that current law is “doctrinally unsustainable” and in need of reform).

9 See Andrew Beck, The Two Sides of the Corporate Veil, in Contemporary Issues in Company Law 69, 90–91 (John Farrar ed., 1987) (positing that the doctrine, as employed, promotes justice); Gallagher & Ziegler, supra note 3, at 293 (arguing that the doctrine prevents injustice).

10 See Moore, supra note 2, at 180–81; Whincup, supra note 5, at 158–59.
corporate parent and its wholly owned subsidiaries as a single entity, a
theory that would be considered expansive even under U.S. law.11

What has plagued the English corporate veil doctrine and partly
contributed to its lukewarm reception by the English courts is its lack of
an overarching analytical framework.12 Judicial reluctance to pierce the
veil can be partly attributed to a perceived haphazardness in the case
law. One famous commentator of English company law has described
the doctrine as “palm-tree justice.”13 Unlike their U.S. counterparts,
English courts have not developed a systematic approach to the cases,
and have instead largely relied on traditional common law concepts to
resolve corporate veil issues.14 Academic analysis of the doctrine has
not proceeded much beyond categorization of cases based on a hodge-
podge of criteria, such as the underlying claims and the legal concept
invoked in the case.15 Even the single economic unit theory is poorly
conceived, and its rationale unclear.16 Given the important role played
by the corporate veil doctrine as an exception to separate corporate
personality and limited liability—general principles which occasionally
produce harsh and unjust results—it is important that the doctrine be
reconsidered and rejuvenated. This requires a structured analytical
framework for the doctrine that is nonetheless consistent with the En-
lish cases.

Although the English and the U.S. corporate veil doctrines have
been well studied in their respective jurisdictions, little comparative re-
search has been done. In fact, this author is not aware of any compara-

W.L.R. 852 (A.C.) at 860 (Eng.); Thompson, supra note 7, at 1041 (noting that U.S. courts
will disregard the corporate entity when it is used for “illegitimate purposes”).

12 See Pickering, supra note 5, at 483 (noting lack of consensus on categorization of case
law and scope of doctrine); K.W. Wedderburn, Company Law—Member’s Rights—Oppression of
Minority, 1958 CAMBRIDGE L. J. 152, 155 (“What is urgently needed is a principle from which
litigants can predict when the courts will, and will not, lift the veil of the corporate entity.”);
K.W. Wedderburn, A Corporation’s Ombudsman?, 23 MOD. L. REV. 663, 666 (1960) (“There is
an urgent need for some principles to be injected into this area of law.”).

13 See J. H. Farrar, Fraud, Fairness and Piercing the Corporate Veil, 16 CAN. BUS. L.J. 474,
1979)).

14 See David H. Barber, Piercing the Corporate Veil, 17 WILLAMETTE L. REV. 371, 376 (1980)
(summarizing the two-prong test required to pierce the veil in U.S. courts); Whincup, supra
note 5, at 159. But see Gallagher & Ziegler, supra note 3, at 292–93 (acknowledging the lack of
a settled approach to English corporate veil cases and suggesting a similar confusion in U.S.
cases).

15 See Paul L. Davies et al., GOWER AND DAVIES PRINCIPLES OF MODERN COMPANY
LAW 202–08 (8th ed. 2008); Gallagher & Ziegler, supra note 3, at 293.

16 See Sealy & Worthington, supra note 3, at 66.
This is surprising given that the United States and the United Kingdom are the two most important common law jurisdictions, and that U.S. corporation laws and English company law are arguably the two most important bodies of corporation law in the world. In economic terms, the United Kingdom is one of the most important economies in Europe and London continues to be a major financial center. In legal terms, to this day, English company law continues to influence the development of corporation law in a number of jurisdictions, including Hong Kong, Singapore, Australia, New Zealand, India, and to a lesser extent, Canada. This Article aims to fill this void in the academic literature, and in the process, shed light on the English corporate veil doctrine and provide a more formal structure to it. Two recent cases decided by the English courts, *Beckett Investment Management Group v. Hall* and *Stone & Rolls Ltd. v. Moore Stephens*, indicated renewed flexibility toward separate corporate personality on the part of the English courts. They present a good opportunity to reconsider the doctrine.

Part I of this Article begins with a brief overview of the English corporate veil doctrine—including its evolution over time and its current state—to provide background for the U.S. reader. Part II provides a detailed comparison between the English and U.S. corporate veil doctrines, focusing on differences in their jurisprudential approaches, treatment of specific case types, and related issues. Drawing on this comparison, Part II discusses the English and U.S. methods of judicial reasoning in the corporate veil piercing context. Part III argues that despite the myriad differences between the two doctrines, they in fact share some crucial similarities. Notably, evidence of the U.S. instru-

---

17 There is only one exception. See Karen Vandekerckhove, *Piercing the Corporate Veil* 75–76 (English law), 93–94 (U.S. law) (2007). This book is a vast multi-jurisdictional survey of the corporate veil doctrine. Id. at 1–2. There was little direct comparison between U.S. and English law, however. See id. at 156, 163, 265, 285, 358, 371, 476, 499, 529–30 (discussing independently but rarely comparing the two bodies of law).


22 See infra notes 29–94 and accompanying text.

23 See infra notes 95–342 and accompanying text.

24 See id.

25 See infra notes 343–406 and accompanying text.
mentality doctrine can be found in a number of important English corporate veil cases. Hence, it is possible to construct an English version of the instrumentality doctrine that brings clarity and structure to the English doctrine. Part IV proposes a revival of the single economic unit theory first propounded by Lord Denning. Drawing on economic analysis and recent cases in English company law and European Union (EU) competition law, Part IV advances a structured approach to the theory. This approach will allow English courts to analyze veil piercing cases involving corporate groups in a systematic manner.

I. A Brief History of the English Corporate Veil Doctrine

Unlike its U.S. counterpart, which has enjoyed steady judicial acceptance throughout the years, the English corporate veil doctrine has had a topsy-turvy career. The attitude of English courts toward the doctrine has oscillated from enthusiasm to outright hostility. The history of the English doctrine can be roughly divided into three periods. The first period lasted from 1897, when Salomon v. Salomon was decided, to around the Second World War. This period can be called the early experimentation period, during which English courts experimented with different approaches to the doctrine. The second period began after the War and continued until 1978, the year when Woolfson v. Strathclyde Regional Council was decided. This period can be regarded as the heyday of the doctrine. Much of the vitality of the doctrine during this period can be attributed to Lord Denning, who was an enthusiastic advocate and practitioner of veil piercing and one of the most influential English jurists of the second half of the twentieth cen-

28 See infra notes 407–580 and accompanying text.
29 See Thompson, supra note 7, at 1049 (charting the success rate for plaintiffs under the doctrine to demonstrate that it has largely remained in the pre-1960s level).
30 [1897] A.C. 22 (H.L.) at 22 (Eng.).
31 See, e.g., Gilford Motor Co. v. Horne, [1933] Ch. 955 (A.C.) at 956 (Eng.) (piercing the veil for attempting to evade a legal obligation); In re Darby, Brougham, [1911] 1 K.B. 95 at 103 (Eng.) (piercing the veil because of misrepresentation).
32 (1978) S.C.(H.L.) 90, 90 (appeal taken from Scot.).
Woolfson marked the beginning of the third period, which has lasted to this day and has seen the doctrine fall into disfavor.34

Given Salomon’s revered status in English law—the 1897 House of Lords decision firmly established the primacy of separate corporate personality and limited liability—one may be excused for forgetting that limited liability was highly controversial at its inception.35 Limited liability was not established in Salomon itself. Rather, the British Parliament granted limited liability to English companies in the Limited Liability Act of 1855.36 Salomon merely affirmed its availability to so-called one-man companies.37 Limited liability met vociferous opposition when it was first proposed in the mid-nineteenth century.38 In an 1824 editorial, The Times of London pronounced that:

Nothing can be so unjust as for a few persons abounding in wealth, to offer a portion of their excess for the formation of a company, to play with that excess—to lend the importance of their whole name and credit to the society, and then, should the funds of the incorporated body prove insufficient to answer all demands, to retire into the security of their unhazardous fortune, and leave the bait to be devoured by the poor deceived fish.39

Even after the passage of the Act, limited liability remained controversial, at least as applied to “one-man companies.”40 This is evident from the appellate decision in Salomon v. Salomon, where the English Court of Appeal unanimously imposed personal liability on Mr. Salomon for his company’s debts.41 Although the House of Lords affirmed the separate legal personality of one-man companies in Salomon, it did not settle the issue once and for all. English courts began to pierce the

---

33 Cf. Sealy & Worthington, supra note 3, at 53 (noting Lord Denning’s “positive enthusiasm” for veil piercing).
36 Halpern et al., supra note 35, at 118–19.
38 See Davies et al., supra note 15, at 194 (referring to “the battle for legislative acceptance of the principle of limited liability in the middle of the nineteenth century”).
39 Editorial, Times (London), May 25, 1824.
40 See Halpern et al., supra note 35, at 119.
41 Broderip v. Salomon, [1895] 2 Ch. 323 (A.C.) at 340–41, 347 (Eng.).
The corporate veil soon after *Salomon*. These early cases included *Apthorpe v. Peter Schoenhofen Brewing*[^42] and *St. Louis Breweries v. Apthorpe*, both of which were taxation cases involving very similar facts.[^43] Veil piercing was not confined to single-shareholder companies. The courts regularly pierced the veil of companies with multiple shareholders.[^44]

What followed was a period of considerable enthusiasm for the corporate veil doctrine. Successful veil piercing cases in the first half of the twentieth century included *Gilford Motor v. Horne*,[^45] *In re Darby, Brougham*,[^46] *Trebanog Working Men’s Club and Institute, Ltd v. MacDonald*,[^47] and *Rainham Chemical Works, Ltd. v. Belvedere Fish Guano Co.*[^48] The lack of a well-defined approach to the doctrine meant that English courts had to experiment with existing common law concepts such as agency, trusteeship, and tort liability principles to resolve corporate personality issues.[^49] These experiments failed to yield a generally applicable framework. But this lack of a general framework did not prevent the courts from piercing the veil when the circumstances so warranted.[^50] This is not to say that the plaintiffs were always successful. There were instances in which the courts refused to pierce the veil.[^51] Nevertheless, the corporate veil doctrine was robust in this period. The sentiment of the time was perhaps best captured in Professor Kahn-Freund’s famous 1944 article, in which he characterized *Salomon* as a “calamitous” decision and

[^42]: [1899] 15 T.L.R. 245 (A.C.) at 245 (Eng.).
[^43]: [1898] 15 T.L.R. 112 (Q.B.) at 112 (Eng.).
[^45]: [1933] Ch. at 943.
[^46]: [1911] 1 K.B. at 100.
[^47]: [1940] 1 K.B. 576 at 582 (Eng.). In *Trebanog* the court did not pierce the veil to find shareholder liability, but separate corporate personality was set aside nonetheless. Specifically, the court found that the incorporated company was acting as an unincorporated trustee for the alcohol of its members, and therefore did not represent an illegal sale by a distinct (incorporated) legal entity. See id.; Sealy & Worthington, *supra* note 3, at 59–60.
[^48]: [1921] 2 A.C. 465 (H.L.) at 466–67 (appeal taken from Eng.).
[^49]: Cf. Pickering, *supra* note 5, at 482–83 (recognizing that English courts created a tangled history regarding corporate personality in the pre- and post-WWII eras).
[^51]: See, e.g., Macaura v. N. Assurance Co., [1925] A.C. 619 (H.L.) at 630 (Lord Sumner) (appeal taken from Ir.) (upholding the separate legal personality of Macaura’s company despite his complete control and ownership, resulting in his inability to collect on his insurance policy); Gramophone & Typewriter, Ltd. v. Stanley, [1908] 2 K.B. 89 (A.C.) at 96 (Eng.) (upholding separate personality of a wholly owned subsidiary, and holding that complete share ownership does not automatically turn a subsidiary into a parent company’s agent).
advocated reforms that would dramatically narrow the scope of limited liability.\textsuperscript{52} He even contemplated abrogating \textit{Salomon} by legislation.\textsuperscript{53}

One exception to the lack of systematic approach in this period is \textit{Smith, Stone and Knight v. Birmingham}, which represented the first attempt by an English court to lay down comprehensive criteria for veil piercing.\textsuperscript{54} Judge Atkinson observed that whether a subsidiary can be said to be carrying on a business on behalf of its parent, which would justify veil piercing, depends on the facts of each case. Judge Atkinson proceeded to identify six guiding questions for each case: (1) Who was really carrying on the business?; (2) Were the profits treated as the profits of the parent company?; (3) Was the parent company the head and the brain of the trading venture?; (4) Did the parent company decide what should be done and how much investment to make in the business?; (5) Did the parent company make a profit based on its skill and direction?; and (6) Was the parent company in effectual and constant control?\textsuperscript{55} These guiding questions are indistinguishable from those applied by U.S courts.\textsuperscript{56} In fact, there is substantial overlap between these criteria and the eleven circumstances identified by Frederick Powell as indicating that a corporation is a mere instrumentality.\textsuperscript{57}

Despite their comprehensive scope, these criteria have been largely overlooked in subsequent U.K cases.\textsuperscript{58} One explanation for this fact is that in applying the corporate veil doctrine, English courts have generally preferred to resort to traditional common law concepts.\textsuperscript{59} This attempt at a comprehensive list of veil piercing criteria probably constituted judicial overreaching in the eyes of English judges. Another explanation is perhaps the unique factual circumstances of that case: it involved the parent company pleading for its own veil to be pierced in

\begin{itemize}
\item \textsuperscript{53} See id. at 57.
\item \textsuperscript{54} See [1939] 4 All E.R. 116 (K.B.). at 121 (Eng.) (concluding that the parent company did control the business and was entitled to compensation for the compulsory purchase of its business).
\item \textsuperscript{55} Id.
\item \textsuperscript{56} See Barber, supra note 14, at 374–75 (1980) (compiling criteria applied by U.S. courts).
\item \textsuperscript{57} See Smith, Stone & Knight Ltd. v. Birmingham, [1939] 4 All E.R. 116 (K.B.) at 121 (Eng.); Cathy S. Krendl & James R. Krendl, \textit{Piercing the Corporate Veil: Focusing the Inquiry}, 55 DENV. L.J. 1, 16–17 (1978) (listing Frederick J. Powell’s published indicators that a subsidiary is a mere instrumentality).
\item \textsuperscript{58} Compare Smith, Stone & Knight, [1939]4 All E.R. at 121 (specifying six factors a court should use in its analysis about whether to pierce a corporate veil), with Jones v. Lipman, [1962] 1 W.L.R. at 836–37 (piercing the corporate veil without specifically using \textit{Smith, Stone & Knight’s factors}), and In re F.G. (Films), [1953] 1 W.L.R. at 486 (same).
\item \textsuperscript{59} See, e.g., In re FG (Films), [1953] 1 W.L.R. at 485.
\end{itemize}
order to obtain compensation from the government. These facts allowed subsequent courts to distinguish their cases from Smith, Stone and Knight and decline to follow its approach.

After the Second World War, the English corporate veil doctrine entered its golden era. Notable veil piercing cases in this period included In re FG (Films), Jones v. Lipman, Firestone Tyre and Rubber v. Lewellin, and Merchandise Transport v. British Transport Commission. As mentioned earlier, the vigor of the doctrine was in no small part due to Lord Denning. He took part in a string of corporate veil cases between the 1950s and 1970s, including Scottish Cooperative Wholesale Society v. Meyer, Littlewoods Mail Order Stores v. Inland Revenue Commissioners, Wallersteiner v. Moir, and lastly—and most well known of them all—D.H.N. Food Distributors Ltd. v. Tower Hamlets London Borough Council. His enthusiasm for the doctrine was best encapsulated in his judgment in Littlewoods, where he warned against blind adherence to Salomon:

The doctrine laid down in Salomon v. Salomon & Co. [1897] A.C. 22, has to be watched very carefully. It has often been supposed to cast a veil over the personality of a limited company through which the courts cannot see. But that is not true. The courts can and often do draw aside the veil. They can, and often do, pull off the mask. They look to see what really lies behind. The legislature has shown the way with group accounts and the rest. And the courts should follow suit.

Lord Denning’s call for judicial flexibility toward corporate personality is commendable. There is a significant difference, however, between requiring group companies to report their accounts on a consolidated basis, and collapsing group companies for liability purposes (thereby ignoring their separate legal personality). Financial reporting chiefly serves informational purposes and the legislature might have required

---

60 See Smith, Stone & Knight, [1939] 4 All E.R. at 118.
61 Compare id. with Merch. Transp. v. British Transp. Comm’n, [1962] 2 Q.B. 173 (A.C.) at 180 (Eng.) (relying on Salomon while overlooking Smith, Stone & Knight), and In re FG (Films), at 485 (ignoring Smith, Stone & Knight).
63 [1962] 1 W.L.R. at 832.
64 [1957] 1 W.L.R. 464 (H.L.) at 464 (appeal taken from Eng.).
67 [1969] 1 W.L.R. 1241 (A.C.) at 1254 (Eng.).
68 [1974] 1 W.L.R. 991 (A.C.) at 993, 1013 (Eng.).
69 [1976] 1 W.L.R. 852 (A.C.) at 857 (Eng.).
account consolidation for reasons that have nothing to do with limited liability. Nonetheless, Lord Denning’s call for judicial flexibility toward corporate entities would eventually culminate in his much-criticized decision in *D.H.N.*

As early as 1956, in *Lee v. Sheard*—not itself a corporate veil piercing decision—Lord Denning foreshadowed his logic in *D.H.N.* by analogizing the relationship between a shareholder and his company to a partnership.71 In *D.H.N.*, a unanimous English Court of Appeal allowed a parent company to claim compensation for disturbance of business under the Land Compensation Act, even though the business and the land on which it sat were owned by different corporate entities.72 Lord Denning’s judgment opened with these famous words: “This case might be called the ‘Three in One.’ Three companies in one. Alternatively, the ‘One in three.’ One group of three companies.”73 Lord Denning continued on to describe the single economic unit theory,74 declaring that “[t]his group is virtually the same as a partnership in which all the three companies are partners. They should not be treated separately so as to be defeated on a technical point.”75 In the same year, a prominent commentator declared that “modern English company law has abandoned the exaggerated view of *Salomon*’s case . . . . English law is now prepared to admit qualifications of, and exceptions to, this principle, by lifting the veil of corporateness.”76 That year marked the height of the corporate veil doctrine. Subsequent development proved that this optimism was misplaced and saw the English judiciary turning increasingly frosty toward the doctrine.

Two years later, the House of Lords openly questioned the reasoning in *D.H.N.* In his judgment in *Woolfson v. Strathclyde Regional Council*, Lord Keith of Kinkel expressed doubt as to “whether in this respect the Court of Appeal properly applied the principle that it is appropriate to

---

71 See *Lee*, [1956] 1 Q.B. at 196; see also *D.H.N.*, [1976] 1 W.L.R. at 860. (“[T]his group is virtually the same as a partnership in which all three companies are partners.”).
73 Id. at 857.
74 See id. at 860. In supporting his description of the single economic unit theory, Lord Denning quotes Professor Gower—a leading authority on English company law at the time—who noted the emerging tendency in English law to treat group companies as a single economic entity. See id.
75 Id. It is important to note that Lord Denning never explicitly referred to the three companies in *D.H.N.* as a single economic entity in his judgment. See id. The term “single economic unit” was subsequently coined by commentators and the courts. See, e.g., *Pierrelli Cable Holding NV v. Inland Revenue Comm’rs*, [2006] UKHL 4, [2006] 1 W.L.R. 400 (H.L.) [73] (Eng.).
pierce the corporate veil only where special circumstances exist indicating that it is a mere façade concealing the true facts.”

Woolfson signaled the beginning of the decline of the doctrine. Although Lord Keith stopped short of overruling D.H.N., subsequent cases made it clear that the single economic unit theory and the corporate veil doctrine were falling out of favor. In Bank of Tokyo v. Karoon, the English Court of Appeal explicitly rejected the single economic unit argument, noting that “we are concerned not with economics but with law. The distinction between the two is, in law, fundamental and cannot here be abridged.”

The corporate veil doctrine arguably reached its nadir in Adams v. Cape Industries plc, in which the English Court of Appeal declared that the use of the corporate structure to limit future liabilities is an inherent feature of English company law and practically ruled out veil piercing in tort cases.

It has not been all gloom and doom for veil piercing plaintiffs after Adams. Two years after that case, the court in Creasey v. Breachwood Motors Ltd. pierced the veil between two companies after their common owners had transferred the assets of the first company to the second in order to avoid an impending judgment. Five years later, in a case involving very similar facts, the Admiralty Court, after explicitly endorsing the reasoning in Creasey, invalidated the sale of a vessel by one member of a corporate group to another. Unfortunately, Creasey was subsequently overruled in Ord & Anor v. Belhaven Pubs.

The current state of the English corporate veil doctrine is best summed up by a leading commentator, who opined that:

The doctrine of lifting the veil plays a small role in British company law, once one moves outside the area of particular contracts or statutes. Even where the case for applying the doc-

---

78 [1987] A.C. 45, 64 (Lord Goff L.J.); see also F. G. Rixon, Lifting the Veil Between Holding and Subsidiary Companies, 102 L. Q. Rev. 415, 415 (1986) (suggesting the demise of the single economic entity in English case law and the piercing of the corporate veil doctrine with respect to group companies).
79 See [1990] Ch. 433 (A.C.) at 544 (Eng.). The English Court of Appeal reviewed various grounds for veil piercing, including fraud, agency, and the single economic entity, and rejected them all. See id. at 494, 545–47, 532–39.
81 See The “Tjaskemolen” (Now Named “Visvliet”), [1997] 2 Lloyd’s Rep. 465 (Q.B.) at 470–71 (discussing the facts of Creasey), 471–73 (comparing the Creasey facts to the similar facts of Tjaskemolen), 474 (holding that the sale of a vessel by one corporation to another was invalid because the transaction was a sham) (Eng.).
82 See [1998] B.C.C. 607 (A.C.) at 616 (Eng.). Although the Belhaven court did not overrule Tjaskemolen, the continued validity of that case is now in doubt. See id.
trine may seem strong, as in the undercapitalised one-person company, which may or may not be part of a larger corporate group, the courts are unlikely to do so.\textsuperscript{83}

After the high-profile cases in the 1990s that came out decidedly against the doctrine, the consensus is that veil piercing has become a rarity under English law.\textsuperscript{84}

Nevertheless, two recent cases suggest a possible change of attitude toward veil piercing. In \textit{Beckett Investment Management Group v. Hall}, the English Court of Appeal pierced the veil between the parent company and its subsidiaries to give effect to a covenant not to compete in an employment contract.\textsuperscript{85} The significance of this case for the corporate veil doctrine is underscored by a number of statements in Lord Justice Kay’s judgment, which will be discussed below.\textsuperscript{86} Likewise, in \textit{Stone & Rolls v. Moore Stephens}, the sole shareholder and director of a company had set up a fraudulent scheme, which its external auditors failed to detect, and swindled huge sums of money from some banks.\textsuperscript{87} The company subsequently went into liquidation and the liquidator brought claims against the auditors for professional negligence.\textsuperscript{88} At issue was whether the culpable shareholder’s intentions should be attributed to the company, which would prevent it from pursuing its claims against the auditors.\textsuperscript{89} The majority of the House of Lords set aside its separate legal personality and imputed the shareholder’s fraudulent intentions to the company.\textsuperscript{90} The company’s claims against the auditors were hence barred.\textsuperscript{91}

The future of the English corporate veil doctrine remains to be seen. \textit{Beckett} and \textit{Moore Stephens} may augur a reversal of fortune for the doctrine. Although \textit{Adams} left the applicability of the doctrine to tort claims in doubt,\textsuperscript{92} \textit{Beckett} reaffirmed the availability of veil piercing in

\textsuperscript{83} Davies et al., \textit{supra} note 15, at 208–09.
\textsuperscript{84} See id.
\textsuperscript{86} See infra text accompanying notes 444–447.
\textsuperscript{87} See [2009] UKHL 39, [2009] 1 A.C. 1391 at 1447–48 (Eng.). Even though the case was primarily concerned with the attribution of a director’s knowledge to the company, it had clear implications for the separate corporate personality principle and the corporate veil doctrine. See id. at 1460.
\textsuperscript{88} See id.
\textsuperscript{89} See id.
\textsuperscript{90} See id. at 1455.
\textsuperscript{91} See id.
\textsuperscript{92} See [1990] Ch. at 547 (refusing to pierce the veil of a British parent company for tort liability of its U.S. subsidiary).
contractual cases. Moreover, *Moore Stephens* may signal a willingness on the part of the English courts to pierce the veil against fraudulent single-shareholder corporations. What is needed for a rejuvenation of the doctrine is a more formal analytical framework than currently exists in the case law. Before proposing such a framework, this Article will first turn its attention to a comparative analysis of the corporate veil doctrine on both sides of the Atlantic.

II. SOME GENERAL COMPARISONS BETWEEN THE TWO CORPORATE VEIL DOCTRINES

At first glance, there seems to be an unbridgeable gulf between the corporate veil doctrines on the two sides of the Atlantic. Although courts in both countries have asserted the role of the judiciary in preventing abuses of the corporate entity—and the discussion about limited liability and the corporate veil doctrine in both jurisdictions often starts with *Salomon v. Salomon*—the attitude of the English judiciary is distinctly more conservative. In contrast, U.S. courts have demon-

---

94 See [2009] W.L.R. at 515 (Lord Walker of Gestingthorpe) (indicating willingness to pierce the veil of a company primarily held and controlled by a single fraudulent actor).
95 See, e.g., Anderson v. Abbott, 321 U.S. 349, 366–67 (1944) (“If the judicial power is helpless to protect a legislative program from schemes for easy avoidance, then indeed it has become a handy implement of high finance. Judicial interference to cripple or defeat a legislative policy is one thing; judicial interference with plans of those whose corporate or other devices would circumvent that policy is quite another.”); *In re a Company*, [1985] 1 B.C.C. 99421 (A.C.) at 99425 (Cumming-Bruce L.J.) (Eng.) (“In our view the cases before and after Wallersteiner v. Moir [1974] 1 W.L.R. 991 show that the court will use its powers to pierce the corporate veil if it is necessary to achieve justice irrespective of the legal efficacy of the corporate structure under consideration.”); *In re Polly Peck Int’l plc*, [1996] 2 All E.R. 433 (Ch.) at 447 (Walker J.) (Eng.) (explaining that lifting the corporate veil is justified when the corporate structure is used as a façade “in an unconscionable attempt to evade existing obligations or to practice some other deception”).
97 Compare Moore, *supra* note 2, at 180–81 (stating that, with few exceptions, English courts have “steadfastly refused” to disregard the principle of limited liability), with Thompson, *supra* note 7, at 1048 (showing that over several decades of the twentieth century, U.S. courts have pierced the corporate veil approximately forty percent of the time).
strated greater willingness to intervene when the situation so warrants.\textsuperscript{98} The differences between these two jurisdictions, however, go deeper than mere disparities in plaintiff success rate. The divergent attitudes of the English and the U.S. courts reveals fundamental differences in their conceptions of judicial decision making, judicial prerogatives to fashion new legal doctrines, the weight given to policy considerations, and the role of justice and its potential conflict with doctrinal rules.

\textbf{A. Jurisprudential Approaches}

1. General Deference to the Separate Corporate Personality Principle

Corporate veil cases come in a great variety. The paradigmatic corporate veil case is one in which the separate corporate personality is disregarded and the shareholders are held liable for the corporation’s debts.\textsuperscript{99} In this type of case, both the rule of separate corporate personality and limited liability—which provide that a shareholder’s responsibility for a corporation’s liabilities is limited to the value of his equity investment—are overridden.\textsuperscript{100} Such cases may be called shareholder liability cases because veil piercing results in shareholder liability for corporate debts.\textsuperscript{101} There are other types of corporate veil cases in which separate corporate personality is disregarded without the imposition of shareholder liability. For example, a subsidiary may be ignored as a separate legal entity to allow a court to exercise jurisdiction over the corporate parent or to compel the production of documents from a subsidiary.\textsuperscript{102} Separate corporate personality may be set aside so that proceeds from fraud that have been deposited into a company can be

\textsuperscript{98} Compare Moore, supra note 2, at 180–81, with Thompson, supra note 7, at 1048. Professor Thompson’s survey of U.S. corporate veil cases suggests that U.S. courts have pierced the veil in approximately forty percent of the cases. Thompson, supra note 7, at 1048. Meanwhile, the general perception within the English legal academia and profession is clearly that the corporate veil doctrine will not often be invoked by the courts. See Moore, supra note 2, at 180–81; Whincup, supra note 5, at 158–59.

\textsuperscript{99} Vandekerckhove, supra note 17, at 11 (defining piercing the corporate veil as “the situation where a shareholder is held liable for its corporation’s debts despite the rules of limited liability and/or separate personality”).

\textsuperscript{100} Sealy & Worthington, supra note 3, at 51.

\textsuperscript{101} See Vandekerckhove, supra note 17, at 13 (explaining the distinction made by many commentators between corporate veil cases where shareholders are held liable and corporate veil cases where shareholders are not held liable).

recovered. The corporate entities of the parent and the subsidiary may also be collapsed for the purpose of determining the scope of a managing director’s responsibilities. Some commentators have called these “identification” cases, while others have used the more colorful metaphor of “peeping behind the veil.”

While the bulk of the corporate veil cases in the United States have been shareholder liability cases, shareholder liability is rarely imposed in the English cases. For example, of the seven cases excerpted under the corporate veil doctrine section of a leading text on English company law, only three attempted to impose shareholder liability by lifting the corporate veil. In the remaining four, separate corporate personality was set aside for the following reasons: to prevent a U.S.-produced film from being registered as a British film under a relevant statute, to allow a corporate parent to claim compensation for the loss of its business in a compulsory purchase (government takings) proceeding, to determine the enemy character of a company, and to avoid a prohibition on the sale of alcohol without a license. Even among the three cases that attempted imposition of shareholder liability, the two in which the veil was pierced involved companies that had been incorporated with the express purpose of evading an existing law.

103 See Wallersteiner v. Moir, [1974] 1 W.L.R. 991 (A.C.) at 1016–17 (Eng.); Gencor ACP Ltd. v. Dalby, [2000] 2 B.C.L.C. 734 (Ch.) at 734 (Eng.).
105 See Vandekerckhove, supra note 17, at 13. The reason for calling these cases identification cases is that the shareholders are identified with the corporation. Id.
109 Id. at 61–63, 66–68; see Adams, [1990] Ch. 433 (refusing to lift corporate veil to impose shareholder liability); Gilford Motor, [1933] Ch. 935 (lifting corporate veil to prevent evasion of contractual obligation); In re Darby, [1911] 1 K.B. 95 (lifting corporate veil where company used to perpetrate fraud).
110 See In re FG (Films), [1953] 1 W.L.R. at 486.
112 See Daimler, [1916] 2 A.C. at 308.
gal obligation or perpetrating fraud. Neither involved a *bona fide* company: a company with an actual business operation, as opposed to a mere shell company. Indeed, *Adams* is the only case listed in this leading text that fits in the paradigmatic corporate veil case, and the veil was left intact in that case. In fact, among the thirty-five English corporate veil cases surveyed in this Article, only six involved *bona fide* companies with legitimate business operations.

There are two main implications from this discussion. First, unlike its U.S. counterpart, the principal focus of the English corporate veil doctrine was never the imposition of shareholder liability. Rather, its focus has been considerably broader, with identification cases forming

---

114 *Gilford Motor*, [1933] Ch. at 961–62. In *Gilford Motor*, an employee had signed a non-compete agreement with its employer car dealer before leaving the company to start his own business. *Id.* at 953–55. To escape the reach of the non-compete clause, the employee set up a company to operate his car dealership. *Id.* at 954–56. The court pierced the veil to apply the non-compete agreement to his company. *Id.* at 961–62.

115 *In re Darby*, [1911] 1 KB at 101–03.


the bulk of the English corporate veil cases.118 What brings an English case under the rubric of the corporate veil doctrine is not that shareholder liability was imposed, but that there was an attempt to set aside separate corporate personality.119 Second, it is highly unlikely that attempts to pierce the veil of a *bona fide* company will be successful under English law. One of the tacit premises of the *Adams* court’s decision to refuse to pierce the veil was that the U.S. subsidiary was a *bona fide* corporate entity with a full-fledged asbestos operation.120 Although the House of Lords did pierce the veil against a *bona fide* company in *Rainham Chemical Works*, the Law Lords’ reasoning relied heavily on common law concepts such as occupation rights.121 Additionally, *Rainham* has since received little attention from both courts and commentators. Meanwhile, U.S. courts have been willing to pierce the veil of a *bona fide* company if elements of the doctrine are met.122 At a jurisprudential level, the English approach to the corporate veil doctrine reflects a substantially more deferential attitude toward limited liability compared with the approach taken by U.S. courts.

2. Reliance on Traditional Common Law Concepts

Beyond this general difference in attitude toward the corporate veil doctrine, the English and U.S. courts’ approaches to the doctrine diverge in a few specific ways. One striking difference between the judicial approaches of these two jurisdictions is that while U.S. courts have been ready to fashion new doctrines—such as the instrumentality doctrine and the alter ego doctrine123—to analyze corporate veil cases, their English counterparts have often confined themselves to applying traditional common law concepts and principles, such as agency and trust.124 In *In re (FG) Films*, the court deemed the British company in-

---

118 See supra text accompanying notes 100–117.
120 See *Adams*, [1990] Ch. at 544.
121 See *Rainham*, [1921] 2 A.C. at 475.
122 See *Barber*, supra note 14, at 373–75.
123 See *Thompson*, supra note 7, at 1063 (analyzing plaintiff’s likelihood of success after courts invoked instrumentality and alter ego doctrines); see also *Barber*, supra note 14, at 376–77 (outlining two-prong corporate veil test developed by courts); Krendl & Krendl, *supra* note 57, at 11–15 (discussing attempts to adopt systematic approach through instrumentality doctrine and alter ego test).
124 See, e.g., *Broderip v. Salomon*, [1895] 2 Ch. 323 (A.C.) at 332, 338 (applying agency and trust concepts to hold Mr. Salomon liable for the company’s debts); *Trebanog* [1940] 1
corporated by its U.S. parent for film registration purposes as the latter’s agent. In *Broderip v. Salomon*, the English Court of Appeal characterized Mr. Salomon himself as the beneficiary, and his company as “a trustee improperly brought into existence by him to enable him to do what the statute prohibits.” In both instances, the court set aside the separate legal personality of the companies because of the improper purpose of incorporation. Although some U.S. courts have also applied agency concepts to corporate veil cases, the number remains small. Other common law concepts are rarely employed to decide corporate veil cases. The English courts made a notable attempt at fashioning a new doctrine in *D.H.N.* This attempt, however, only enjoyed a brief success. In fact, *D.H.N.* itself was the perfect illustration of the important role played by traditional common law concepts in the English corporate veil cases. In addition to the single eco-

K.B. at 582 (analogizing the company to a trustee holding trust property for its beneficiary members).

125 *In re F.G. (Films)*, [1953] 1 W.L.R. at 484, 486.

126 *Broderip*, [1895] 2 Ch. at 338. This case was overruled by the House of Lords in *Salomon v. Salomon*, [1897] A.C. 22 (H.L.) at 58 (appeal taken from Eng.). Nevertheless, the House of Lords did not overrule the English Court of Appeal on the grounds that a company could never be its shareholders’ trustee, but that in the instance before the court the trustee relationship was not established. See id. at 56–57.

127 See *Broderip*, [1895] 2 Ch. at 338; *In re F.G. (Films)*, [1953] 1 W.L.R. at 486. In *Broderip*, the improper purpose was to shield Mr. Salomon from liability. [1895] 2 Ch. at 338. In *In re F.G. (Films)*, the improper purpose was to circumvent the restriction on the showing of foreign films under the Cinematograph Films Acts of 1938 and 1948. [1953] 1 W.L.R. at 486.


129 See *Thompson*, supra note 7, at 1063 (noting that of the more than 1600 cases surveyed, only fifty-two mentioned agency as a consideration—representing roughly three percent of the cases—while 181 applied the alter ego doctrine and 550 mentioned domination and control by the shareholders).

130 See id. (mentioning neither trust nor any other common law concepts as the bases for veil piercing by U.S. courts in any case in his comprehensive survey).


132 See *Woolfson*, (1978) S.C.(H.L.) at 96. In his judgment, Lord Keith of Kinkel expressed doubts on “whether in this respect the Court of Appeal [in *D.H.N.*] properly applied the principle that it is appropriate to pierce the corporate veil only where special circumstances exist indicating that it is a mere façade concealing the true facts.” *Id.* One should not be too quick to announce the demise of the single economic unit theory, however. Litigants have continued to invoke the theory after *Woolfson*, and courts have continued to apply it, with mixed results for plaintiffs. See, e.g., *Adams*, [1990] Ch at 544; *In re Polly Peck*, [1996] 2 All E.R. at 447.
nomic unit theory, the *D.H.N.* court invoked trust law concepts such as constructive and resulting trusts.\(^{133}\) Even Lord Denning felt compelled to buttress the single economic unit theory with more mundane common law concepts such as irrevocable license.\(^{134}\) This reliance on existing legal concepts is in keeping with the general sentiment of the English courts, which tend to hold faith in the ability of the existing common law concepts to solve new problems.\(^{135}\) To the extent that existing concepts are applied to decide corporate veil cases, one may argue that there is no distinct content in or structure for the English corporate veil doctrine. The doctrine is merely a label for the cases in which separate corporate personality is considered by the courts and for the myriad non-company law concepts that are applied to decide them.

3. Absence of Overarching Theory

Unlike U.S. corporate veil cases, the English doctrine lacks an overarching theory or analytical framework. In the United States, the instrumentality doctrine was first formulated in 1931 by Frederick J. Powell.\(^{136}\) Many state courts have since adopted some variation of this doctrine.\(^{137}\) In contrast, no general analytical framework can be identified in English cases, nor have the English courts enumerated relevant

\(^{133}\) See *D.H.N.*, [1976] 1 W.L.R. at 859. Property law concepts were called upon because under the relevant statute for the determination of compensation, the corporate parent would have been entitled to more compensation had its interest in the land been more than that of a bare licensee. Two of the judges in the case, Lord Denning and Lord Justice Goff, believed that the parent company had an irrevocable license to the land. *Id.* at 859, 860. Trust concepts were utilized to determine that the parent company had an equitable interest in the land beyond that of a bare licensee. *Id.* at 859, 865.

\(^{134}\) See *id.* at 859.

\(^{135}\) See, e.g., *id.* at 859, 865. Another example is the equitable subordination doctrine, under which a shareholder’s capital contribution to the corporation will be subordinated to its creditors’ claims if the shareholder is found to have committed abuses in his or her management of the corporation. See *Taylor v. Standard Gas & Elec. Co.*, 306 U.S. 307, 323–24 (1938) (applying the equitable subordination doctrine in an early bankruptcy case).


\(^{137}\) Krendl & Krendl, *supra* note 57, at 13 n.40 (listing various state cases adopting the instrumentality doctrine). In addition, the courts have identified a host of relevant factors and considerations for adjudicating veil piercing claims, including lack of substantive separation, failure to comply with corporate formalities, undercapitalization, and overlap of corporate personnel. See Barber, *supra* note 14, at 374–75; Thompson, *supra* note 7, at 1063.
factors or considerations. The prevailing analytical approach by English courts is to classify cases into a number of categories based on a mix of the legal concepts invoked and the factual circumstances in which the case arises. These categories include agency, trusts, fraud, group enterprises, enemy, revenue (taxation), and cases based on provisions of the Companies Act. Agency and trusts are categories premised on the legal concepts applied. Fraud, taxation, and Companies Act cases, on the other hand, are based on the nature or the statutory basis of the underlying claims. The group enterprise category consists of cases sharing a common factual circumstance, namely that the shareholder at issue is a corporation and the plaintiff is attempting to impose enterprise liability on a corporate group. Enemy is a unique category illustrated by *Daimler v. Continental Tyre and Rubber*. At issue in that case was whether a company incorporated in the United Kingdom, all but one of whose shareholders and directors were German, should be deemed an enemy company during the First World War.

These categories would have been helpful if they were predictive of the outcome of cases. With the exception of fraud, however, this has not been the case. Nor is there any common mode of analysis across cases within the same category. Take the category of group enterprises as an example. A review of the leading cases reveals that there is no common analytical framework. *Lonrho v. Shell Petroleum* was not concerned with

---

138 Beck, *supra* note 9, at 71, 73. Part III.B of this Article will argue that even though the English courts do not explicitly identify factors such as failure to comply with corporate formalities or undercapitalization, these factors do underpin many of their decisions. *See infra* text accompanying notes 359–406.

139 *See* Beck, *supra* note 9, at 72 n.7.


141 *See* id. at 60. In fact, many of the cases in this category, such as *D.H.N.* and *Woolfson*, are not concerned with the imposition of shareholder liability at all. They belong to the group enterprise category merely because the cases involve corporate groups.

142 *Id.* at 61.

143 *Daimler*, [1916] 2 A.C. at 310.

144 *See, e.g.*, *In re a Company*, [1985] 1 B.C.C. at 99425; *Wallersteiner*, [1974] 1 W.L.R. at 1014 (“[T]o my mind Dr. Wallersteiner was guilty of misfeasance.”); *In re Darby*, [1910] 1 K.B. at 99; *Trustor*, [2001] 1 W.L.R. at [23]; *Gencor*, [2000] 2 B.C.L.C. at [26] (applying equitable principles to lift the veil because the subsidiary was nothing more than a “creature company” used to receive profits).

145 *See, e.g.*, *Lonrho*, [1980] 1 W.L.R. at 636 (leaving the corporate veil intact in a discovery action); *Meyer*, [1959] A.C. at 343 (reasoning that when a subsidiary is formed with an independent minority of shareholders, the parent must deal fairly with the subsidiary); *Beckett*, [2007] I.C.R. at 1439 (using rules of contract interpretation broadly to pierce the corporate veil); *Ord*, [1998] B.C.C. at 615 (declaring to pierce the veil on limited liability grounds and holding that there were no improprieties alleged); *Adams*, [1990] Ch. at 543 (stating that legal precedent provides “rather sparse guidance”); *Multinational Gas*, [1983]
imposition of liability.\textsuperscript{146} At issue was whether the corporate parent was required to produce documents possessed by its subsidiaries.\textsuperscript{147} The case largely hinged on the interpretation of the word “power” in the relevant rule for discovery in the English High Court.\textsuperscript{148} \textit{Gramophone and Typewriter v. Stanley} was a taxation case in which the pivotal issue was the control exercised by the parent over the subsidiary.\textsuperscript{149} In \textit{In re Polly Peck}, a bankruptcy case,\textsuperscript{150} the court applied a range of concepts, ranging from agency and nomineeship to the single economic unit theory.\textsuperscript{151} The lack of a common analytical framework is unsurprising given the diversity of facts and issues presented in these group enterprise cases. Even if there was a common approach to these cases, it would most likely require modifications in order to be applied to cases ranging from discovery to bankruptcy and taxation.

The foregoing discussion may give the impression that the U.S. and English approaches to veil piercing are mutually exclusive. This is in fact not the case. Unexpected as it may be, a number of the categories under the prevailing English approach share common concerns and attributes with the U.S. instrumentality doctrine formulated by Powell.\textsuperscript{152}

4. The Role of Policy Considerations

A third difference between the U.S. and English judicial approaches to the corporate veil doctrine is the readiness of the former to acknowledge and discuss the policy considerations behind the statute

\textsuperscript{146} \textit{Lonero}, [1980] 1 W.L.R. at 635.

\textsuperscript{147} \textit{Id.}

\textsuperscript{148} \textit{Id.} at 633, 635.

\textsuperscript{149} \textit{Gramophone}, [1908] 2 K.B. at 95–96.

\textsuperscript{150} [1996] 2 All E.R. at 433.

\textsuperscript{151} \textit{Id.} at 445–48.

\textsuperscript{152} See \textit{infra} text accompanying notes 359–406.
or the legal rule at issue. In cases premised on the interpretation of a statute, U.S. courts have often given explicit recognition to the policy rationale of the statute. In *National Labor Relations Board v. Fullerton Transfer & Storage*, in deciding whether to pierce the veil of the two corporations at issue, the Sixth Circuit Court of Appeals repeatedly referred to federal labor policies and focused on whether there was “a specific attempt to thwart labor law obligations.” In *Walkovszky v. Carlton*, there was an in-depth discussion of the policy considerations behind the minimum liability provisions of the New York Vehicle and Traffic Law and the need for the court to refrain from second-guessing or overriding it. The courts have also acknowledged the special policy concerns behind federal income taxation and bankruptcy laws. Even in cases not premised on specific statutes, U.S. courts paid heed to general public policy. For example, in *Swearngin v. Sears Roebuck & Co.*, the court affirmed the trial court where the trial court had pierced the veil of a two-tier corporate structure that had been set up to obstruct product liability suits against the manufacturer. In doing so, the court specifically noted that the policy reasons for shielding distributors or retailers from product liability claims did not apply to the sales division of an enterprise that also manufactured the product.

In contrast, the English courts do not accord the same weight to policy considerations. In *Commissioner of Inland Revenue v. Sanson*, a taxation case, the issue was whether the business at issue belonged to the company or the dominant shareholder, which in turn would affect the tax liability of the respondent. Without relying on relevant taxation policies, the English Court of Appeal decided the case by invoking agency concepts and highlighting the fact that the shareholder had made a loan to the company. The court reasoned that if it were to disregard the separate personality of the company, it would have meant

---

154 NLRB v. Fullerton Transfer & Storage Ltd., 910 F.2d 331, 338 (6th Cir. 1990).
155 See *Walkovszky*, 223 N.E.2d at 9.
157 See 376 F.2d 637, 642–43 (10th Cir. 1967).
158 Id. at 643.
159 See Whincup, *supra* note 5, at 159. *But see* Merch. Transp., [1962] 2 Q.B. at 185 (noting that “Parliament [had] decided that it would be in the public interest that public hauliers should be protected against the inroads of private hauliers with spare capacity” unless the latter complies with some special licensing requirements); *Ebbw Vale*, [1951] 2 K.B. at 373 (recognizing the duty of the relevant authority to secure “an efficient, adequate, economical and properly integrated system of public inland transport”).
160 [1921] 2 K.B. 492 (A.C.) at 496–97 (Eng.).
161 See id. at 505, 509, 515.
that the shareholder had made a loan to himself, which would be logically absurd.  

In \textit{D.H.N.}, Lord Denning did not attempt to justify his decision by the policy rationale of the Land Compensation Act, even though he believed that the parent company was meant to be covered by the statute for the loss of its business.  

An even more striking illustration of this de-emphasis of policy considerations in English cases is \textit{Rainham Chemical Works v. Belvedere Fish Guano}.  

This case featured a direct conflict between the policies of limited liability and tort rules. In his judgment, Lord Buckmaster acknowledged that the objective of limited liability is that “enterprise and adventure . . . be encouraged,” and tacitly recognized the conflict between this and the policy objective of tort law, which is to compensate tort victims. Nevertheless, he proceeded to decide the case by applying the property law concepts of tenancy and occupation rights.

This is not to say that the English judges never consider policy arguments. In an area of law in which the cases are as varied as corporate veil cases, generalizations are fraught with difficulty and subject to reservations and qualifications. It remains true, however, that U.S. courts have demonstrated a greater propensity to take into account policy arguments. This is to be expected given the English judiciary’s general aversion to policy arguments. One implication of U.S. courts’ receptiveness to policy arguments is that it allows them to show heightened sensitivity to the factual circumstances and legal considerations of each

\begin{itemize}
  \item \textit{See id.} at 502, 504–05, 509.
  \item \[1976\] 1 W.L.R. at 857, 860.
  \item \textit{See} \[1921\] 2 A.C. at 467. This case concerned two business people who were operating a plant manufacturing explosives for the British Ministry of Munitions during the First World War. Mindful of the potential for liability of their business, they incorporated a company and transferred the business to it. \textit{Id.} at 474–75. An explosion ensued and caused damages to property adjacent to the plant. \textit{Id.} at 466. The property owner sought to impose liability on both the company and the directors. \textit{Id.}
  \item \textit{Id.} at 475.
  \item \textit{Id.} at 476 (“If the company was really trading independently on its own account, the fact that it was directed by Messrs. Feldman and Partridge would not render them responsible for its tortious acts unless, indeed, they were acts expressly directed by them. If a company is formed for the express purpose of doing a wrongful act or if, when formed, those in control expressly direct that a wrongful act be done, the individuals as well as the company are responsible for the consequences.”).
  \item \textit{Rainham}, \[1921\] 2 A.C. at 478. In fairness, Lord Buckmaster seemed to have subtly resolved the conflict in favor of limited liability when he stated that “[h]aving contracted for the manufacture of munitions, which in itself does not appear to me that liability could not be displaced by the introduction of an independent contractor.” \textit{Id.} at 477.
  \item \textit{See Krendl & Krendl, supra} note 57, at 29–31 (illustrating where U.S. courts have pierced the corporate veil for companies that are blatantly avoiding a public policy).
  \item \textit{See Whincup, supra} note 5, at 159.
\end{itemize}
Comparing the English & U.S. Corporate Veil Doctrines

5. The Role of Justice

There is consensus that a main purpose of the corporate veil doctrine in the United States is to achieve justice. In *Anderson v. Abbott*, the Supreme Court declared that “the courts will not permit themselves to be blinded or deceived by mere forms of law but will deal with the substance of the transaction involved as if the corporate agency did not exist and as the justice of the case may require.”

In *Arnold v. Philips*, the Fifth Circuit Court of Appeals concluded that application of the corporate veil doctrine is premised on the corporate entity being “a sham, a mere obstacle to justice, or instrument of fraud.”

In perhaps one of the most famous corporate veil cases in the United States, *Berkey v. Third Avenue Railway*, Justice Cardozo, then sitting on the New York Court of Appeals, similarly noted the importance of justice as a guide in corporate veil cases. The role of justice has been confirmed time and again in recent cases.

---

170 See *id.* at 160, 161.
171 See *id.* at 161.
176 See *Berkey*, 155 N.E. at 61. When discussing the circumstances under which a subsidiary’s liability will be imputed to its corporate parent, Justice Cardozo observed that “[d]omination may be so complete, interference so obtrusive, that by the general rules of agency the parent will be a principal and the subsidiary an agent. Where control is less than this, we are remitted to the tests of honesty and justice.” *Id.*
177 See *Fullerton Transfer & Storage*, 910 F.2d at 340 (noting that the promotion of justice and the prevention of fraud presented two bases for piercing the corporate veil in the instant case); *Doughty v. CSX Transp., Inc.*, 905 P.2d 106, 110 (Kan. 1995) (“[A]ppellate
In contrast, English judges have demonstrated an ambivalent attitude toward considerations of justice. Some English judges share similar views with their U.S. brethren. In In re a Company, Lord Justice Cumming-Bruce asserted that “[i]n our view the cases before and after Wallersteiner v. Moir show that the court will use its powers to pierce the corporate veil if it is necessary to achieve justice irrespective of the legal efficacy of the corporate structure under consideration.” In Atlas-Maritime Co. v. Avalon Maritime Ltd., Lord Justice Neill asserted that “in the exercise of a discretion in relation to injunctive relief ‘the eye of equity’ can, I think, look behind the corporate veil in order to do justice.” A number of commentators have similarly endorsed the primacy of justice in corporate veil cases. Despite these sympathetic statements, the current consensus within the English judiciary is that it is not “open to this court to disregard the principle of Salomon v. A Salomon & Co Ltd merely because it considers it just to do so.” In Adams, the court’s refusal to pierce the veil meant that the plaintiff asbestos victims could not enforce their judgments against the English corporate parent of a U.S. subsidiary, which had been shut down by the parent as part of an elaborate scheme to evade liability. Unfortunately for the asbestos victims in that case, Adams firmly established that the sanctity of separate corporate personality trumps the judicial prerogative to prevent injustice.

This English perception of the role of justice in corporate veil cases reflects a more formalistic approach to judicial decision-making. It is also consistent with the English courts’ reliance on traditional common law concepts and their aversion to explicit references to public

---

178 [1985] 1 B.C.C. at 99425 (internal citation omitted). This statement was endorsed in Creasey, [1992] B.C.C. at 646. In D.H.N., Lord Justice Shaw similarly emphasized the importance of justice as a basis for veil piercing when he noted that a refusal to grant compensation to D.H.N. for disturbance of its business would have been “a denial of justice.” See [1976] 1 W.L.R. at 854, 867.

179 [1991] 1 Lloyd’s Rep. 563 (A.C.) at 569 (Eng.).

180 See Beck, supra note 9, at 71; Gallagher & Ziegler, supra note 3, at 307; Whincup, supra note 5, at 161.

181 Adams, [1990] Ch. at 537.

182 See id. at 534, 541.

183 See id. at 541–44.

184 See Beck, supra note 9, at 91–92.
policies. \textsuperscript{185} After all, policy considerations are often invoked to modify existing legal concepts or to substantiate a reinterpretation of precedent so that a more just or fair result can be reached. The English judges’ reservations about the overt pursuit of justice may be due to the open-ended nature of the concept. Although there are cases, such as \textit{Adams}, in which there is little doubt about the just outcome, \textsuperscript{186} in some close cases, justice can be said to be in the eye of the beholder. For instance, while Lord Justice Shaw believed that it would be an affront to justice to deny D.H.N. compensation for the value of its business, \textsuperscript{187} a commentator has argued that allowing a company to “play hot and cold” with incorporation would be grossly unfair. \textsuperscript{188} Moreover, acknowledging justice as the ultimate goal of the doctrine is one thing, but applying it as an operational principle is quite another. U.S. courts have fashioned new doctrines—such as the instrumentality doctrine and the alter ego doctrine—to encapsulate the occasional conflict between separate corporate personality and considerations of justice, and to render justice an operational principle for deciding cases. \textsuperscript{189} The English judiciary’s reluctance to create new doctrines means that they must rely on existing legal concepts, which are often not well-suited for that purpose. \textsuperscript{190} Deprived of the better-tailored tools available to their U.S. counterparts, it is not surprising that English judges have shied away from the pursuit of justice in veil piercing cases.

6. Summing Up

The various jurisprudential characteristics of the English corporate veil doctrine are interrelated. The formalistic approach to judicial decision making and the strict adherence to separate corporate personality enshrined in \textit{Salomon} \textsuperscript{191} means that there is little need to take into account notions of justice and policy considerations. This de-emphasis of justice and policy in turn obviates the need to fashion a new analytical framework, thus allowing English courts to rely on traditional common law concepts which were not created with the corporate veil piercing situation in mind, and are ill-suited for deciding corporate veil cases. \textsuperscript{192}

\textsuperscript{185} See id.
\textsuperscript{186} See \textit{Adams}, [1990] Ch. at 434–35.
\textsuperscript{188} See \textit{Rixon}, \textit{supra} note 78, at 420.
\textsuperscript{189} See \textit{Barber}, \textit{supra} note 14, at 377.
\textsuperscript{190} See \textit{Whincup}, \textit{supra} note 5, at 159.
\textsuperscript{191} See id.
\textsuperscript{192} See id.
The jurisprudential development of the U.S. corporate veil doctrine contrasts with its English counterpart and reflects the different judicial philosophy prevailing in the United States. U.S. courts first analyzed corporate veil issues with traditional common law concepts, but quickly found them deficient.\footnote{See id. at 160–61.} They proceeded to develop special doctrines such as instrumentality and alter ego, which were necessary because they allowed the courts to focus on the factors that are truly important in corporate veil cases.\footnote{See Barber, supra note 14, at 377.}

The use of agency concepts by U.S. courts vividly illustrates this difference. Like their English counterparts, U.S. courts have applied these concepts in corporate veil cases.\footnote{See Barber, supra note 14, at 377.} Nevertheless, U.S. courts appreciated the limitations of agency concepts early on. In \textit{Kingston Dry Dock v. Lake Champlain Transportation}, Judge Learned Hand noted that while agency concepts may be useful in consensual agency relationships expressly contracted for by the parties, they do not focus on considerations such as “the parent’s direct intervention in the transaction, ignoring the subsidiary’s paraphernalia of incorporation, directors and officers” that ultimately drive corporate veil decisions.\footnote{31 F.2d at 267.} In \textit{Lowendahl v. Baltimore & Ohio Railroad}, the court similarly observed the following:

\begin{quote}
[E]xcept in cases of express agency, the use of the term “agent” seems unfortunate and may tend to confuse . . . [and] the “instrumentality” rule seems to furnish the most practical and effectively applicable theory for breaking down corporate immunity where equity requires that this be done to circumvent fraud or other legal wrong.\footnote{247 A.D. at 156.}
\end{quote}

Early U.S. commentators similarly noted the inadequacy of agency concepts for resolving corporate veil claims.\footnote{See Douglas & Shanks, supra note 1, at 195.} Some U.S. courts use the term “agent” interchangeably with instrumentality rather than in the traditional common law sense.\footnote{See Barber, supra note 14, at 400.} Unlike its English counterpart, traditional agency concepts have largely faded into the background in U.S. corporate veil jurisprudence.

\footnotesize
\begin{itemize}
\item \footnote{See id. at 160–61.}
\item \footnote{See Barber, supra note 14, at 377.}
\item \footnote{See \textit{House of Koscot Dev.}, 468 F.2d at 67 n.2 (noting that the concepts of agency, instrumentality, and identity are interchangeable); \textit{Kingston Dry Dock}, 31 F.2d at 267; \textit{Walkauszky}, 223 N.E.2d at 10; \textit{Berkey}, 155 N.E. at 61; \textit{Anderson}, 398 S.W.2d at 637.}
\item \footnote{31 F.2d at 267.}
\item \footnote{247 A.D. at 156.}
\item \footnote{See \textit{Douglas & Shanks}, supra note 1, at 195.}
\item \footnote{See Barber, supra note 14, at 400.}
\end{itemize}
Although the English judges’ aversion to the creation of new legal doctrines is consistent with their general judicial philosophy, it is worth remembering that common law concepts were judicial creations in the first place. They originate from precedents that are, in some cases, hundreds of years old. For instance, negligence, which is the dominant basis for tort liability, was a judicially crafted doctrine and only became a fully recognized basis for tort liability in the nineteenth century.\footnote{See Dan B. Dobbs et al., Prosser and Keeton on Law of Torts § 28 (5th ed. 1984).} When the need arose, English judges did not hesitate to create this new doctrine to tackle the pressing problems of the time.\footnote{See Morton J. Horwitz, The Transformation of American Law: 1780 to 1860 89 (1977).} Therefore, their conservatism should not prevent them from coming up with an analytical framework that presents a more systematic approach to corporate veil cases.

B. Treatment of Specific Types of Cases

1. Fraud and Misrepresentation

Promotion of justice and prevention of fraud are the main objectives of the U.S. corporate veil doctrine.\footnote{See Thompson, supra note 7, at 1041.} It is important to note that fraud is given a considerably broader meaning in the corporate veil cases than its general usage. When the courts refer to fraud in these cases, the conduct at issue usually falls short of common law fraud, and is more appropriately deemed a misrepresentation.\footnote{Krendl & Krendl, supra note 57, at 31; Thompson, supra note 7, at 1044–45 n.53.} The courts seem to use fraud and misrepresentation to refer to the same class of conduct.\footnote{Krendl & Krendl, supra note 57, at 31.} Plaintiff success rates in the United States are very high when misrepresentation is present.\footnote{See Thompson, supra note 7, at 1063. Thompson includes in this category cases in which the defendant made misrepresentations as to the assets and financial condition of the corporation and the party responsible for payment. Id. at 1044–45 n.53.} Professor Thompson found that U.S. courts had pierced the veil in ninety-four percent of the cases in which misrepresentation was established.\footnote{Id. at 1063. This is hardly surprisingly because once a court has branded certain conduct as misrepresentation, refusal to pierce the veil would be tantamount to tacit endorsement of the conduct.} Perhaps even more revealingly, they had refused to pierce the veil in over ninety-two percent of the cases in which the absence of misrepresentation was noted.\footnote{Id. at 1064–65 n.141.}
As stated earlier, fraud is the only predictive category within the English corporate veil case classification. English judges have often pierced the veil when the defendant is found to have perpetrated a fraud. Unlike their U.S. counterparts, most of the English cases that fall within the fraud category have generally involved misappropriation of corporate assets or other outright fraudulent conduct. Nevertheless, the English courts have also occasionally expanded the scope of fraud to encompass misrepresentation. This is aptly illustrated by \textit{In re Darby, Brougham}, one of the leading veil piercing cases involving fraud. That case featured two notorious fraudsters in the United Kingdom at the time, Darby and Gyde, who had set up an elaborate scheme to defraud public investors. To attract public investment in the debentures issued by their sham company, they concealed their connection to the scheme from the public. They accomplished this through the incorporation of a Guernsey corporation as the parent of the sham company. Although the ostensible ground for veil piercing in that case was fraud, Judge Phillimore characterized the fraud involved as “what they did through the corporation they did themselves and represented it to have been done by a corporation of some standing and position, at any rate a corporation which was more than and different from themselves.” Judge Phillimore also stated that “they represented that some business was being done by or through the corporation and concealed the fact that it was done by and through Darby and Gyde.”

\begin{footnotes}
\item[209] See, e.g., \textit{Wallersteiner}, [1974] 1 W.L.R. at 999, 1001 (involving a financier’s acquisition of a company through circular check transactions, defrauding the company’s other shareholders and creditors in the process); \textit{Trustor}, [2001] 1 W.L.R. at 1179 (involving the misappropriation of corporate funds by one of the officers); \textit{Gencor}, [2000] 2 B.C.L.C. at 734 (involving the misappropriation of corporate funds and business opportunities by officers to themselves or to companies controlled by them).
\item[210] See [1911] 1 K.B. at 101. In that case, two notorious fraudsters, Darby and Gyde, had incorporated a company in Guernsey, an island jurisdiction in the English Channel off the coast of France. They then used this company to incorporate yet another company, Welsh Slate Quarries Ltd., to offer debentures to the general public. Welsh Slate never had a viable business and the investors lost all their investment. Meanwhile, Darby and Gyde had made a secret profit from their promotion activities. The liquidator for Welsh Slate sought to pierce the veil of the Guernsey corporation to recover the secret profit. \textit{Id.}
\item[211] See \textit{id.}
\item[212] See \textit{id.}
\item[213] See \textit{id.}
\item[214] \textit{Id.}
\item[215] \textit{Id.} at 101.
\end{footnotes}
Therefore, Judge Phillimore’s real objection was to how they concealed their involvement in the scheme and misrepresented that the Guernsey corporation was the entity behind the sham company.\textsuperscript{216} It is noteworthy that the misrepresentation in that case did not concern the amount of assets in support of the company—the usual subject of misrepresentation in U.S. cases\textsuperscript{217}—but the general credibility and creditworthiness of its shareholders.\textsuperscript{218}

Some commentators argue that from a transaction cost perspective, misrepresentation justifies veil piercing because it induces corporate creditors to expend monitoring costs to a socially wasteful level.\textsuperscript{219} If misrepresentation results in shareholder liability, shareholders will be deterred from engaging in it, and creditors will incur lower costs to detect it.\textsuperscript{220} This justification, however, is questionable. A creditor would only investigate if it had grounds for suspicion. If the representation made by the corporation seemed valid and genuine, the creditor would have little reason to investigate it.

Take the example of \textit{In re Darby}. If the investors had no reason to doubt the legitimacy of the Guernsey corporation, they would have little motivation to undertake any investigation, regardless of the availability of shareholder assets for recovery.\textsuperscript{221} The investigation would have been straightforward: the identity of the shareholders of the corporation should have been available at the Companies Registry in Guernsey.\textsuperscript{222} Moreover, the extent to which the investors would have been reassured by shareholder liability is unclear. If they had been aware of Darby and Gyde’s role in the sham company, the investors probably would not have been assuaged by the availability of veil piercing to remain in the company. They would have either not invested at all or tried to sell their investments. The incentive to investigate does not arise from the fact that limited liability confines investor recovery to corporate assets. Rather, the incentive derives from the potential inves-

\textsuperscript{216} See \textit{In re Darby}, [1911] 1 K.B. at 101, 103.


\textsuperscript{218} See \textit{In re Darby}, [1911] 1 K.B. at 101.


\textsuperscript{220} See id. The kind of misrepresentation Posner has in mind is when a parent has represented to a subsidiary that the parent’s assets will be available to the subsidiary’s creditors. See id. at 521.

\textsuperscript{221} See \textit{In re Darby}, [1911] 1 K.B. at 101.

\textsuperscript{222} See id.
tortors’ desire to avoid exposure to fraud in the first place. Contrary to the transaction cost perspective, altering the size of the pool of assets available for recovery will not have significant impact on investor incentives to investigate. Justifying veil piercing in misrepresentation cases on transaction cost grounds is shaky at best.

The obvious response from the transaction cost line of reasoning is that if misrepresentation results in veil piercing, shareholders will refrain from attempting it in the future. There will be no need for potential creditors to investigate. Unfortunately, shareholder liability will not deter potential fraudsters from attempting to defraud creditors. These fraudsters will react to the threat of veil piercing by shifting the proceeds of their fraudulent schemes beyond the reach of the creditors by way of offshore accounts, trusts, or other liability-evading devices. Shareholder liability will raise the costs of fraud, but it will not eliminate it. Corporate creditors will continue to investigate and monitor the corporation regardless of the existence of a misrepresentation exception to limited liability.

It seems that one can best justify a misrepresentation exception with a combination of estoppel and what one commentator refers to as the ideals of Truth and Respect that govern a debtor’s dealing with his creditor. Estoppel refers to the idea that once a shareholder has made the representation that he is responsible for the corporation’s debts, and the creditor has reasonably relied on it, the shareholder should be estopped from repudiating the representation. The ideal of Truth requires the corporate debtor to be truthful and honest in its dealing with its creditors, and the ideal of Respect requires it to accord primacy to the creditors’ claims when engaging in transfers of corporate assets.

The 1961 Texas Court of Appeals case Moore & Moore Drilling Co. v. White may illustrate the court’s tacit utilization of the principles of Truth

226 See Clark, supra note 224, at 509–11. Professor Clark defines the ideal of Truth as a requirement that in connection with transfer of property to others, a debtor cannot tell lies that will lead to non-satisfaction of the creditor’s claim. Id. at 509. The ideal of Respect imposes on the debtor “a moral duty in transferring his property to give primacy to so-called legal obligations, which are usually the legitimate, conventional claims of standard contract and tort creditors, as opposed to the interests of self, family, friends, shareholders, and shrewder or more powerful bargaining parties.” Id. at 510–11.
and Respect.\textsuperscript{227} In that case, a Texas state court refused to pierce the veil based only upon the defendant corporation’s silence about its financial situation.\textsuperscript{228} The court would not impose shareholder liability absent fraud or other active misrepresentation.\textsuperscript{229} This result, while sound, is inconsistent with the transaction cost analysis presented above. From the transaction cost perspective, the \textit{Moore} court should have pierced the veil because doing so would have compelled the corporation to come forward with its own financial information, saving the plaintiff creditor substantial investigation costs. Taking this reasoning to its logical conclusion, from a transaction cost perspective there should be no material difference between active misrepresentation and omission to provide information. In either case, transaction costs will be saved if the threat of veil piercing compels the corporation to come forward with relevant information. It thus goes beyond mere sanctioning of active misrepresentation and would impose an affirmative duty on a corporate debtor to provide full disclosure of relevant financial information to its creditors. Nevertheless, courts would likely find this approach too intrusive.\textsuperscript{230}

A better explanation for the outcome in \textit{Moore} is that the defendant corporation was truthful and respectful in its dealing with the plaintiff creditor and had made no representations of any kind. The defendant did not pursue deceptive conduct that resulted in non-satisfaction of the creditor’s claims.\textsuperscript{231} Nor did it undermine the primacy of the creditor’s claims in any asset transfer by favoring its own associates.\textsuperscript{232} Moreover, the lack of active representation to the creditor means that the estoppel rationale did not apply.\textsuperscript{233} The \textit{Moore} court was thus correct in refusing to pierce the veil.

2. Tort Cases

Economic analysis of limited liability and the corporate veil doctrine has drawn distinctions between different types of claimants. According to this mode of analysis, judicial readiness to override limited

\textsuperscript{228} See id.
\textsuperscript{229} See id.
\textsuperscript{230} See id. (noting that in the absence of a fiduciary relationship between the parties, there is no affirmative duty to disclose financial information of the corporation).
\textsuperscript{231} See Clark, supra note 224, at 508–09.
\textsuperscript{232} See id. at 511–12.
\textsuperscript{233} \textit{Moore}, 345 S.W.2d at 555–56; see Clark, supra note 224, at 541 n.97 (claiming that in the absence of express agency, estoppel, or tort, three elements must be proved to make a parent corporation responsible for the subsidiary: control, fraud, and proximate causation of injury).
liability should vary according to the claimant at issue. In particular, many commentators have argued that veil piercing should be more readily available to tort claimants than to contractual claimants because of the former’s inability to negotiate ex ante for compensation or to seek additional credit protection from the corporation.\textsuperscript{234} Despite this almost unanimous view within the academic community, U.S. courts have been more sympathetic toward contractual claims than tort claims. Professor Thompson found that the plaintiff success rate in contractual corporate veil cases was forty-two percent, compared to only thirty-one percent for tort cases.\textsuperscript{235} Although an eleven percent difference is clearly substantial, a thirty-one percent success rate suggests that tort claimants prevail with sufficient regularity in the United States as well.

The same cannot be said about English tort claimants. The general view in the English judiciary after the \textit{Adams} decision is that the door to tort claimants bringing veil piercing claims is largely closed.\textsuperscript{236} In response to the argument that the corporate parent in that case had intentionally located hazardous asbestos operations in an underfunded subsidiary to avoid potential liabilities, the court proclaimed that:

\begin{itemize}
  \item \textsuperscript{234} See Henry Hansmann & Reinier Kraakman, \textit{Toward Unlimited Shareholder Liability for Corporate Torts}, 100 \textit{Yale L.J.} 1879, 1919 (1991); David W. Leebron, \textit{Limited Liability, Tort Victims, and Creditors}, 91 \textit{Colum. L. Rev.} 1565, 1566–68 (1991); Comment, \textit{Should Shareholders Be Personally Liable for the Torts of Their Corporations?}, 76 \textit{Yale L.J.} 1190, 1195 (1967); see also Barber, supra note 14, at 381–82 (suggesting that because a tort claimant has usually not engaged in prior dealings with the corporation, it seems illogical to require a disregard of corporate formalities before piercing the veil); Phillip I. Blumberg, \textit{Limited Liability and Corporate Groups}, 11 \textit{J. Corp. L.} 573, 616–17 (1986) (claiming that the efficiency advantages of limited liability where there are tort claimants disappears as compared to when there are contract claimants); Frank H. Easterbrook & Daniel R. Fischel, \textit{Limited Liability and the Corporation}, 52 \textit{U. Chi. L. Rev.} 89, 112 (1985) (claiming that courts are more willing to disregard the corporate veil in tort than in contract cases); William P. Hackney & Tracey G. Benson, \textit{Shareholder Liability for Inadequate Capital}, 43 \textit{Pitt. L. Rev.} 837, 867 (1982) (stating that a tort claimant differs from a contract claimant because the tort claimant cannot “select the person he does business with”); Halpern et al., supra note 35, at 145–47 (arguing that some unlimited liability be available for tort claimants because of the risk that corporations will shift the risk of business failures to these claimants). \textit{But see} Posner, supra note 219, at 519–20 (arguing that piercing the corporate veil for tort claimants would create a negative externality in that investment would be discouraged); Robert B. Thompson, \textit{Unpacking Limited Liability: Direct and Vicarious Liability of Corporate Participants for Torts of the Enterprise}, 47 \textit{Vand. L. Rev.} 1, 31–32 (1994) (arguing that liability should not be extended to shareholders because it will have a significant negative effect on the ability of shareholders to diversify and therefore will remove the advantage of standardized share-pricing necessary for liquid financial markets for shares).
  \item \textsuperscript{235} See Thompson, supra note 7, at 1058.
  \item \textsuperscript{236} See \textit{Trustor}, [2001] 1 W.L.R. at 1184–85.
\end{itemize}
[W]e do not accept as a matter of law that the court is entitled to lift the corporate veil as against a defendant company which is the member of a corporate group merely because the corporate structure has been used so as to ensure that the legal liability (if any) in respect of particular future activities of the group (and correspondingly the risk of enforcement of that liability) will fall on another member of the group rather than the defendant company. Whether or not this is desirable, the right to use a corporate structure in this manner is inherent in our corporate law.  

The English courts have not always held such a staunchly pro-corporation view on this issue. In *Rainham Chemical Works*, Lord Sumner described the two shareholders’ attempt to shield themselves from potential tort liabilities through the interposition of a company as “expedient” and “ignoble,” and the House of Lords proceeded to impose shareholder liability for the damages caused by an explosion at their explosives plant. Remarkably, this was done in the absence of any improper conduct or wrongdoing on the part of the shareholders. In this sense, the Law Lords were actually more progressive as compared with their U.S. contemporary, Justice Cardozo, in *Berkey v. Third Avenue Railway*, where Justice Cardozo refused to pierce the veil between the parent and the subsidiary in light of the absence of fraudulent conduct. For whatever reason, *Rainham Chemical Works* has not received much attention in the academic literature or subsequent cases. *Adams* has largely shaped the current view on the applicability of the corporate veil doctrine to tort claims in the United Kingdom.

Some U.S. courts have expressed a similar view to that articulated in *Adams*. The underlying premise of these U.S. courts seems to be

---

237 *Adams*, [1990] Ch. at 544.
238 [1921] 2 A.C. at 483. Lord Sumner added that the use of the corporate entity “was excellent if it could be made good, but unfortunately it required that Messrs. Feldman and Partridge should be at the same time occupiers and not occupiers of the premises, which was absurd.” *Id.* at 484.
239 *Id.* at 466.
240 *Id.* at 486. The decision held that the two shareholders had retained their occupation rights with respect to the plant and therefore remained liable to the tort victim under English precedent *Rylands v. Fletcher*. *Id.* at 478, 479.
241 See *Berkey*, 155 N.E. at 58, 61; *Rainham*, [1921] 2 A.C. at 486. The degree of control exerted by the corporate parent over the subsidiary in *Berkey* was quite similar to that wielded by the two shareholders over the company in *Rainham Chemical Works*.
that the use of the corporate entity to limit future liabilities is legitimate.\footnote{See Craig, 843 F.2d at 150; Harris, 8 Cal. App. 3d at 843.} For these courts, however, the crucial condition is the absence of improper conduct or wrongdoing on the part of the corporation and its shareholders.\footnote{See Craig, 843 F.2d at 150. After accepting that Cape’s liability evasion scheme was fraudulent, the Third Circuit Court of Appeals proceeded to declare that “evasion of tort liability has never, in itself, been sufficient basis to disregard corporate separateness.” \textit{Id}.} The stance taken by English and U.S. courts on the proper use of the corporate entity to limit future tort liabilities is most aptly illustrated by \textit{Adams}\footnote{See Craig, 843 F.2d at 149–50 (“We accept for purposes of this appeal the district court’s findings and conclusion that Cape’s scheme to avoid asbestos-injury liability in the United States constituted the type of ‘fraud or injustice’ that would satisfy that element of the standard for piercing the corporate veil.”). The Third Circuit refused to impose liability on Cape Industries’ shareholder, who had no direct involvement in the liability-evasion scheme or NAAC’s asbestos operations in the United States. \textit{See id}. at 150–52.} and the 1988 Third Circuit Court of Appeals case \textit{Craig v. Lake Asbestos of Quebec}.\footnote{See Whincup, \textit{supra} note 5, at 162.} These two cases in fact arose from the same events and involved the same plaintiffs, employees of NAAC, which was the U.S. asbestos-producing subsidiary of Cape Industries in \textit{Adams}.\footnote{See Rainham, [1921] 2 A.C. at 470 (seeking to pierce the corporate veil in compensation for damage to property following an explosion at a munitions factory).} There was ample evidence before the English Court of Appeal that the decision to shut down NAAC was calculated to escape a default judgment that had been entered against it by a U.S. court.\footnote{See Moore, \textit{supra} note 2, at 182.} Yet by refusing to pierce the veil, the court tacitly condoned it. In contrast, the Third Circuit acknowledged the fraudulent nature of this liability evasion scheme.\footnote{See Craig, 843 F.2d at 146–48; \textit{Adams}, [1990] Ch. at 434–35.} These two cases show that English courts hold a considerably more permissive view of the exploitation of limited liability protection against tort claims. In comparison, most U.S. courts would have treated Cape Industries’ activities as asset stripping.\footnote{Thompson, \textit{supra} note 7, at 1058.}

Nevertheless, it is noteworthy that even before \textit{Adams}, tort claims appeared less frequently among English corporate veil cases. Of the thirty-five prominent English corporate veil cases surveyed in this Article, only one involved tort claims.\footnote{See \textit{id}. at 146–48; \textit{Adams}, [1990] Ch. at 434–35.} This starkly contrasts with the composition of the corporate veil cases in the United States. Of the 1,600 cases Professor Thompson included in his survey, 226 were tort cases.\footnote{See \textit{id}. at 150–52.} It is not entirely clear what explains this dearth of tort cases in
the United Kingdom. It is possible that even before *Adams*, potential tort claimants appreciated that avoidance of future liabilities is an inherent feature of the English corporate entity.

3. Asset Stripping vs. Evasion of Existing Legal Obligations

One class of U.S. corporate veil cases involves asset stripping. Asset stripping is shareholder dissipation of corporate assets by way of excessive dividend disbursements, inflated director salaries, or other improper means.\(^{254}\) The corporation is usually rendered insolvent as a result, and hence becomes unable to meet its creditors’ claim or to satisfy a judgment.\(^{255}\) Asset stripping may take place before or after a judgment is entered against the corporation, but is usually performed after a legal obligation has been incurred.\(^{256}\) It should not come as a surprise that the U.S. courts have long adopted a hostile attitude toward such conduct. As early as 1939, in *Pepper v. Litton*, the Supreme Court subordinated the debts of the controlling shareholder of a corporation after the shareholder was found to have orchestrated a complex scheme to empty the corporate coffers.\(^{257}\) The company’s scheme, which involved voluntary filing of bankruptcy, was executed with the specific intent to evade a judgment debt against the corporation.\(^{258}\) Regardless of whether it happens before or after a judgment has been issued, asset stripping has been roundly condemned by the U.S. courts ever since *Pepper v. Litton*.\(^{259}\)

The closest concept to asset stripping in the English corporate veil doctrine is evasion of existing legal obligations, which encompasses some of the best-known English corporate veil cases.\(^{260}\) These cases bring to the forefront one of the most vexing questions in the corporate veil doctrine: how to distinguish between legitimate uses and abuses of

\(^{254}\) See Vandekerckhove, *supra* note 17, at 178.

\(^{255}\) See id.

\(^{256}\) See id. at 180.

\(^{257}\) 308 U.S. 295, 310–12 (1939).

\(^{258}\) See id. at 299.


the corporate entity.\textsuperscript{261} This is related to the earlier discussion about the reluctance of English courts to pierce the veil of a \textit{bona fide} company.\textsuperscript{262} The English courts have affirmed the legitimacy of using the corporate structure to avoid \textit{future} liabilities.\textsuperscript{263} Therefore, the crucial inquiry seems to be the relative timing of the incurrence of legal obligations and the use of corporate structure to avoid them. The best formulation of this inquiry appeared in a Hong Kong case prior to 1997, when English precedents were still binding on the Hong Kong courts. In \textit{China Ocean Shipping v. Mitrans Shipping}, the Hong Kong Court of Appeal declared that:

Using a corporate structure to \textit{evade} legal obligations is objectionable. The courts’ power to lift the corporate veil may be exercised to overcome such evasion so as to preserve legal obligations. But using a corporate structure to \textit{avoid} the incurring of any legal obligation in the first place is not objectionable. And the courts’ power to lift the corporate veil does not exist for the purpose of reversing such avoidance so as to create legal obligations.\textsuperscript{264}

\textsuperscript{261} See, e.g., \textit{Laborers’ Pension Fund}, 580 F.3d at 611 (stating that the rule for analyzing whether the corporate form would sanction fraud is not clear cut).

\textsuperscript{262} See supra notes 117–121 and accompanying text. An English commentator has labeled these evasion cases as fraud cases. See Jennifer Payne, \textit{Lifting the Corporate Veil: A Reassessment of the Fraud Exception}, 56 CAMBRIDGE L.J. 284, 284 (1997). Nevertheless, there are considerable differences between the underlying conduct at issue in these evasion cases compared with the genuine fraud cases discussed above. All the genuine fraud cases featured outright fraudulent conduct. \textit{Gencor ACP} and \textit{Smallbone} both involved misappropriation of company funds and \textit{Wallersteiner} involved so-called “circular cheque transactions” that were used to inflate the corporation’s book assets. See \textit{Wallersteiner}, [1976] 1 W.L.R. at 991–92; \textit{Trustor}, [2001] 1 W.L.R. at 1177; \textit{Gencor}, [2000] 2 B.C.L.C. at 734. Meanwhile, the evasion cases involved the defendant’s attempt to avoid legal obligations that had been created under fully legitimate and lawful circumstances. \textit{Gilford Motor} and \textit{Creasey} involved an unemployment contract, \textit{jones} involved a conveyance of land, and \textit{Yukong} and \textit{Tjaskemolen} involved charter-parties. See \textit{Yukong}, [1998] 1 W.L.R. at 294; \textit{Gilford Motor}, [1933] Ch. at 935; \textit{Tjaskemolen}, [1997] 2 Lloyd’s Rep. at 465; \textit{Creasey}, [1992] B.C.C. at 638; \textit{jones}, [1962] 1 W.L.R. at 832. It was only when one of the parties to these legal transactions attempted to evade their legal obligations that the issue of veil piercing arose. See Payne supra, at 287–88. Therefore, this Article draws a distinction between these two categories of cases and refers to cases involving evasion of existing legal obligations as “evasion cases.”

\textsuperscript{263} \textit{Adams}, [1990] Ch. at 544 (remarking that it is unobjectionable that “the corporate structure has been used so as to ensure that the legal liability (if any) in respect of particular future activities of the group . . . will fall on another member of the group rather than the defendant company”).

If the legal obligation existed prior to incorporation, the use of the corporate structure would be considered evasion and the veil would be pierced. If the legal obligation was only created after incorporation, the use would be considered as avoidance and separate corporate personality would be upheld. This formulation of the evasion of existing legal obligations rule in fact explains a majority of the English cases in this category.\textsuperscript{265}

Again, it should come as no surprise that English courts have been more tolerant of asset stripping than their U.S. counterparts. Two cases in particular—\textit{Adams}\textsuperscript{266} and \textit{Yukong v. Rendsburg Investments Corp.}\textsuperscript{267}—featured the kind of conduct that would most likely have been condemned by U.S. courts.\textsuperscript{268} Yet English courts refused to pierce the veil. In \textit{Adams}, the English Court of Appeal treated the case as concerning the legitimate incorporation of a subsidiary to limit the asbestos liability exposure of a corporate group.\textsuperscript{269}

In reality, the corporate parent evaded the default judgment that had been entered against the subsidiary by shutting it down and transferring its operations to a newly incorporated subsidiary. To ensure that it would be safely insulated from liability, the parent interposed a Liechtenstein dummy corporation between itself and the new subsidiary.\textsuperscript{270} The asbestos liabilities had materialized when the second subsidiary was incorporated, which means that its veil should have been pierced under the evasion rule. The court chose to focus on the first subsidiary, however, which means that incorporation preceded the incurrence of liabilities and separate corporate personality was upheld.\textsuperscript{271} The court likely reasoned that because the liabilities were incurred by the first subsidiary, it should be the time of its incorporation—not the second subsidiary’s incorporation—that ought to serve as the reference point

\textsuperscript{265} See, e.g., \textit{Gilford Motor}, [1933] Ch. at 938 (upholding a covenant not to compete because the legal obligation had been created prior to the incorporation of defendant’s company); \textit{Tijshemolen}, [1997] 2 Lloyd’s Rep. at 469 (extending liability to the acquiring company because the alleged breach of contract had occurred before the vessel was transferred); \textit{Creasey}, [1992] B.C.C. at 638 (allowing the plaintiff to seek the assets of the acquiring company because the alleged wrongful termination occurred before the takeover); \textit{Jones}, [1962] 1 W.L.R. at 835 (requiring a sham company to convey property pursuant to the purchaser’s pre-merger contract with original company).

\textsuperscript{266} [1990] Ch. at 433.

\textsuperscript{267} [1998] 1 W.L.R. at 294.

\textsuperscript{268} See Craig, 843 F.2d at 149 (accepting the district court’s finding that Cape’s liability evasion scheme was fraudulent.)

\textsuperscript{269} See [1990] Ch. at 544.

\textsuperscript{270} See id. at 450.

\textsuperscript{271} See id. at 539–44.
when applying the evasion rule. Although this reasoning may be logically sound, it has essentially left the door wide open for evasion by judgment debtors.\textsuperscript{272} This dilemma can be sidestepped if the court shows greater flexibility in the choice of reference point. In a case such as \textit{Adams}, given the substantial connection between the first and the second corporations—the latter had received all of the former’s assets and hired all of its employees—it would be appropriate to choose the creation of the second corporation as the reference point.

\textit{Yukong} exposed a further, and perhaps more serious, limitation of the evasion rule as a tool against asset stripping.\textsuperscript{273} Just as in \textit{Adams}, the defendant shareholder shifted assets from the first to the second company after a legal obligation had been created.\textsuperscript{274} Again, the court refused to pierce the veil.\textsuperscript{275} Unlike the company in \textit{Adams}, the second company in \textit{Yukong} was incorporated before the legal obligation incurred.\textsuperscript{276}

The second company existed when the breach of contract at issue took place. Therefore, no flexibility in the choice of reference point would have resulted in veil piercing. \textit{Yukong} demonstrates how a shareholder can evade liabilities without being subject to veil piercing. The key is to ensure that incorporation precedes the incurrence of the legal obligation. But such ease of evasion is highly unsatisfactory. Therefore, the evasion rule should be modified to reach not only the incorporation of a new company, but also the use of an existing company to evade legal obligations. The reference point should be modified from

\begin{footnotesize}
\textsuperscript{272} See Ord, [1998] B.C.C. at 615 (demonstrating the impact on evasion post-\textit{Adams}).

\textsuperscript{273} See \textit{Yukong}, [1998] 1 W.L.R. at 309–10. \textit{Yukong} concerned a charter-party between Rendsburg, a Liberian company controlled by the defendant shareholder, and Yukong Line, a Korean shipping company. \textit{Id.} at 297–98. Rendsburg repudiated the contract after market conditions had deteriorated, after which Yukong sued for breach of contract. \textit{Id.} at 298. Upon filing of the suit, the shareholder promptly transferred Rendsburg’s assets to another company also controlled by him. \textit{Id.} at 298–99. Yukong sought to hold the defendant shareholder personally liable by piercing the veil between him and Rendsburg. \textit{Id.} at 299.

\textsuperscript{274} \textit{Id.} at 298–99; see \textit{Adams}, [1990] Ch. at 443–44. In fact, the removal of assets took place on the same day that the contract was repudiated. \textit{Yukong}, [1998] 1 W.L.R. at 302. The court refused to accept that this sequence of events was a coincidence, and concluded that “the purpose of the transfer was to put Rendsburg’s assets beyond the reach of Yukong in the probable event of litigation.” \textit{Id.}

\textsuperscript{275} \textit{Id.} at 310. This decision was all the more remarkable given the judge’s characterization of Rendsburg as the shareholder’s “alter ego.” \textit{Id.} at 302. The judge further found that the shareholder had completely controlled both companies, that his conduct before and during the trial was disreputable, and that he contradicted himself on several key issues of the case. \textit{Id.}

\textsuperscript{276} Compare \textit{id.} at 298, with \textit{Adams}, [1990] Ch. at 444.
\end{footnotesize}
the time of incorporation to the time a corporate entity is utilized. Such use may be by way of incorporation, as in *Adams*,277 or by way of transfer of funds to the company, as in *Yukong*.278

The different approaches to asset stripping on both sides of the Atlantic vividly illustrate the limitations of the English formalistic approach to veil piercing. The emphasis of the evasion rule under English law is on contractual obligations and tort liabilities. What matters is when a contract is breached or when a tort is committed vis-à-vis when a company is incorporated.279 In contrast, the U.S. corporate veil doctrine focuses on the underlying economic reality that justifies veil piercing, notably, the fact that a shareholder has taken deliberate action to impair creditor interests.280 This focus is more in keeping with the policy-driven and justice-conscious approach to veil piercing adopted by U.S. courts.281 In light of all the complexities of the evasion rule, it would seem that if the goal was to deter evasion of legal obligations, the U.S. approach to asset stripping should be preferred. By focusing on whether the shareholder has acted deliberately to undermine creditor interests instead of comparing the relative timing of the incurrence of legal obligation and the use of the corporate entity, the U.S. approach focuses the inquiry on the crux of an asset stripping case.

Given the entrenched status of the evasion rule under English law, however, the best one can do is to improve its administrability. In fact, *Adams* raised one further issue concerning its application. Although the asbestos injuries had occurred prior to the incorporation of the secondary subsidiary, liability was only established after incorporation.282 Therefore, whether the evasion rule should have resulted in veil piercing depends on one’s view of when a legal obligation arises. There are a number of possibilities. One of them is that a legal obligation exists once liability is established.283 Applying this formulation, one commentator endorsed the outcome in *Adams* by arguing that liabilities were merely speculative when the second subsidiary was incorporated.284 One prob-

---

277 See [1990] Ch. at 443–44.
280 See Krendl & Krendl, supra note 57, at 28.
281 See id. at 2.
283 Hansmann & Kraakman, supra note 234, at 1896. Hansmann and Kraakman call this the judgment rule. *Id.*
284 Payne, supra note 262, at 289. It is not clear that the liabilities were speculative. First, the event that gave rise to liability had already taken place when the corporate structure was used as a liability limiting device. *Adams*, [1990] Ch. at 449–50. Second, Cape had
lem with this approach is that it would create a gaping loophole for shareholders to exploit. If legal obligations only exist when liability is judicially established, the time elapsed can be used to transfer assets out of the company. The second possibility is that a legal obligation exists when a claim is filed in court. Nevertheless, this possibility still leaves much room for strategic behavior on the part of the shareholders. The timeframe for liability evasion strategies would be shorter, but the possibility still exists. The best rule seems to be that a legal obligation is incurred when the event that gives rise to potential liability takes place, and the company knows, or has reason to know, that the event may expose it to liability. This formulation would have changed the results in Adams. In Adams, the corporate parent was clearly aware of the potential for further liability when it set up the second corporation. After all, it had settled similar asbestos claims.

On the other hand, whether a contractual obligation arises requires a different analysis. A contractual obligation is created when the parties enter into the contract. The time of breach is more precisely described as the time when a potential liability arises, and judgment date as the time when liability is established. When the evasion rule is applied to contractual cases, the more suitable reference point is the time of breach. For contractual claims, the main concern of the evasion rule is that the shareholders will attempt to avoid judgment after they know or have reason to know that liability will be incurred: i.e., the breach. This concern can be adequately addressed by choosing the time of breach as the reference point. Adopting the time of contract formation as the reference point would be the most complete solution as it would cover practically all attempts at evasion.

settled some similar asbestos claims prior to incorporating the second subsidiary. Id. at 449.

285 See Hansmann & Kraakman, supra note 234, at 1896.
286 See id. at 1896–97.
287 Id. at 1896.
288 See id. at 1897. This is similar to Hansmann and Kraakman’s “information-based rule.” Id. Their rule is more complicated, however, because it is an ex ante unlimited liability rule for corporate torts. It thus takes into account issues that are not particularly pertinent to ex post veil piercing. See id.
289 [1990] Ch. at 449.
290 See Restatement (Second) of Contracts § 9 (1981).
291 See id. §§ 243, 346, 357.
Nevertheless, adopting the time of contract formation as the reference point would prove over-inclusive by catching many legitimate uses of the corporate entity. It is important to note that the choice of the reference point should reflect the likelihood that at that point in time, the shareholders had a specific intent to evade liability. Most contractual parties do not contemplate contractual breach or evasion of liability at the time of contract formation. Moreover, choosing contract formation as the reference point would mean that shareholder liability could result from any asset transfer after the contract was entered into regardless of whether the shareholders had a specific intent to evade liability. In fact, using contract formation as the reference point would be tantamount to imposing an obligation on the corporation not to transfer its assets to any affiliated company after the contract was entered into. This would clearly be an excessively burdensome obligation.

A further modification of the evasion rule is needed for contractual obligations. The necessity of this modification is due to the fact that the breaching party can control the timing of the breach. A corporation bent on evading its contractual obligation may incorporate a second company or transfer funds to it just before it repudiates a contract, thereby avoiding veil piercing. To forestall this possibility, if the use of the corporate entity takes place close in time to the breach, there should be a rebuttable presumption of evasion. To avoid veil piercing, the company must rebut this presumption by adducing evidence that the use of the corporate structure was for a legitimate purpose.

Applying this modified evasion rule would have changed the result in Yukong. In Yukong, the shareholder knew that the contractual repudiation would expose him to liability and he promptly transferred assets from the first to the second company after he had decided to repudiate the contract. As the court found, that was why he siphoned funds out of the first company on the day of repudiation. Given that the time when the shareholder knew, or should have learned, about the impending breach preceded the transfer of assets to the second company, the veil would have been pierced under the modified evasion rule.

---

293 See id.
294 See id.
4. Voluntary Piercing Cases

One notable feature of the English corporate veil cases is the preponderance of voluntary piercing cases. Voluntary piercing is also referred to as reverse piercing. This is a misnomer, however, because strictly speaking, reverse piercing refers to the imposition of liability on the corporation for the shareholder’s debts. This contrasts with the usual piercing situation—forward piercing—in which the shareholders are held liable for the corporation’s debts. Voluntary piercing, however, refers to the invocation of the doctrine by the shareholders themselves to set aside the corporate entity. It is voluntary because it is instigated by the shareholder itself. The typical veil piercing claim is involuntary in that it is filed by a creditor or corporate outsider, and the shareholders are subject to liability against their will. Therefore, even though reverse piercing is often voluntary, it need not be so. An example of an involuntary reverse piercing claim is when a corporate parent’s creditors seek recovery from its subsidiaries because, perhaps, the latter have more assets in their possession than the parent.

Interestingly, in a majority of English voluntary piercing cases, the veil piercing claims are advanced so that the shareholder can claim

\[\text{References}\]


298 See Kingston Dry Dock, 31 F.2d at 267. In this case, the plaintiff attempted to reach for the subsidiary’s assets because the parent had failed to satisfy the debts. Id.; see also Floyd v. Internal Revenue Serv., 151 F.3d 1295, 1299–1300 (10th Cir. 1998) (remanding case in which IRS sought to reverse-pierce corporation to satisfy individual’s debts); Towe Antique Ford Found. v. Internal Revenue Serv., 999 F.2d 1387, 1390–91 (9th Cir. 1993) (allowing United States’ levy of corporation’s automobiles to satisfy individual taxpayer’s liabilities because corporation was individual’s alter ego); LFC Mktg. Group v. Loomis, 8 P.3d 841, 843 (Nev. 2000) (finding that the alter ego doctrine may be applied to recover an individual debt from the assets of a corporation determined to be the alter ego of the individual debtor); Zisblatt v. Zisblatt, 693 S.W.2d 944, 955 (Tex. App. 1985) (finding that most of the assets the corporation acquired during the time of individual’s marriage were in fact part of the individual community estate and subject to division upon divorce).
compensation or benefit of some kind.\footnote{See, e.g., Macaura, [1925] A.C. at 625–28; Smith, Stone & Knight, [1939] 4 All E.R. at 120–22.} These cases may be referred to as compensation cases, three of which, in particular, arose in a compulsory purchase (government takings) context.\footnote{See, e.g., Woolfson, (1978) S.C.(H.L) at 90; D.H.N., [1976] 1 W.L.R. at 854; Smith, Stone & Knight,, [1939] 4 All E.R. at 117. Compulsory purchase is the equivalent of government takings under eminent domain in the United States. See Ben Beckman, Note, The Wholesale Decommissioning of Vacant Urban Neighborhoods: Smart Decline, Public-Purpose Takings, and the Legality of Shrinking Cities, 58 CLEV. ST. L. REV. 387, 392 (2010) (“‘[C]ompulsory purchase’ [is] primarily used in British, while ‘eminent domain’ [is] exclusively American.”).} These cases raise interesting theoretical issues for two reasons. First, the corporations themselves are pleading for veil piercing.\footnote{See Zisblatt, 693 S.W.2d at 947.} Second, unlike the typical veil piercing cases, where shareholder liability is at stake, the doctrine is invoked in these cases to obtain financial benefits for the corporation.\footnote{See LFC Marketing Group, 8 P.3d at 844.} The English courts have expressed almost uniform hostility toward voluntary piercing. In \textit{Tunstall v. Steigman}, Lord Justice Danckwerts stated that “[the owner] cannot say that in a case of this kind she is entitled to take the benefit of any advantages that the formation of a company gave to her, without at the same time accepting the liabilities arising therefrom.”\footnote{[1991] Ch. 512 at 531.} Vice-Chancellor Browne-Wilkinson expressed this sentiment pointedly in \textit{Tate Access Floors Inc. v. Boswell}:

> If people choose to conduct their affairs through the medium of corporations, they are taking advantage of the fact that in law those corporations are separate entities, whose property and actions are in law not the property or actions of their incorporators or controlling shareholders. In my judgment controlling shareholders cannot, for all purposes beneficial to them, insist on the separate identity of such corporations but then be heard to say the contrary when discovery is sought against such corporations.\footnote{Gaertner, supra note 296, at 681.}

There is greater diversity of views among U.S. courts toward voluntary piercing. Some courts believe that voluntary piercing should be rejected on principle, while some apply the same analysis to both voluntary and involuntary piercing claims.\footnote{[1962] 2 Q.B. at 601.} Yet other courts apply a more

---

\footnote{See, e.g., Macaura, [1925] A.C. at 625–28; Smith, Stone & Knight, [1939] 4 All E.R. at 120–22.}
open-ended policy analysis to voluntary piercing claims.\(^{309}\) Despite this variety of analytical approaches, voluntary piercing has a considerably lower likelihood of success. In his comprehensive survey, Professor Thompson discovered that only 13.4% of voluntary piercing claims prevailed, while the success rate was 42.3% when the plaintiff was a creditor.\(^{310}\) These figures suggest that English and U.S. courts are similarly averse to voluntary piercing.

Although the argument that the shareholders should not be allowed to have it both ways carries considerable intuitive appeal, it is worth pondering whether it admits of exceptions or qualifications. This is illustrated by \(D.H.N.\)^\(^{311}\) The benefit allegedly obtained by \(D.H.N.\) from its corporate structure was savings on a stamp duty.\(^{312}\) The cost was the loss of compensation for the disturbance of business.\(^{313}\) Although the opinion contained no information on the amount of compensation in dispute, it is likely to have exceeded the savings on stamp duty.\(^{314}\) Given this discrepancy between the realized costs and benefits of incorporation, one may defend voluntary piercing in \(D.H.N.\) on the grounds that there was a significant mismatch between the realized costs and benefits from incorporation, and that compulsory purchase is not something that \(D.H.N.\) could have reasonably foreseen. The argument is that when a potential incorporator considers whether to set up a corporation, it forms expectations of the benefits and costs of incorporation. It weighs these expected benefits and costs, and decides to incorporate if the former exceed the latter. Given that the incorporator cannot anticipate

\(^{309}\) Id.

\(^{310}\) Thompson, supra note 7, at 1057.

\(^{311}\) See \(D.H.N.\), [1976] 1 W.L.R. at 854–55, 857–60. The facts of this case are as follows: \(D.H.N.\), the corporate parent, operated a grocery business on a premise in Tower Hamlets, a borough in London. Id. at 857. \(D.H.N.\) had acquired the premise through a series of rather complex transactions that resulted in the land being owned by one of its wholly owned subsidiaries, Bronze. Id. at 854–55. The Tower Hamlets Borough Council issued a compulsory purchase order for the redevelopment of the land and offered compensation for the value of the land to the subsidiary. Id. at 854. The Council, however, refused to compensate the parent for the value of the business itself. Id. at 858. Under the Land Compensation Act, \(D.H.N.\) would only have been entitled to compensation if it had been more than a bare licensee of the land. Id. at 857–58. The English Court of Appeal unanimously pierced the veil between \(D.H.N.\) and Bronze and allowed compensation for the disturbance of business on a variety of grounds, including that the companies constituted a single economic entity, that \(D.H.N.\) had an irrevocable license to the land, and that \(D.H.N.\) had an equitable interest in the land. Id. at 859–60.

\(^{312}\) See id. at 858.

\(^{313}\) See id. at 857–58.

\(^{314}\) See id. at 858, 860.
every contingency, it follows that it should not be held accountable for consequences that it should not be expected to have foreseen.

Although this may initially seem like a convincing argument, it is flawed in a number of ways. First, it is difficult to determine what the incorporator should have and should not have foreseen. Likewise, it is very difficult for a court to make this determination ex post. Second, and more importantly, while there are unexpected costs, there are also unexpected benefits. Voluntary piercing may permit a court to save an incorporator from a substantial unexpected cost. Nevertheless, no legal doctrines exist to allow a court to remove the unexpected benefits. Therefore, permitting voluntary piercing under these circumstances creates a windfall for the incorporator. In the end, the best rule is to require an incorporator to bear all the costs, expected and unexpected, of incorporation, which means that a court should not sanction voluntary piercing in a compensation case.315

C. Issues External to the Doctrine

There are two further differences between the English and the U.S. corporate veil doctrines that do not directly bear on the judicial approaches to the doctrine, but are nonetheless interesting to note.

1. Dearth of Complementary Doctrines

The first difference is the lack of related substitutes for the corporate veil doctrine under English law. One commentator has argued that fraudulent conveyance laws, dividend restraint statutes, the equitable subordination doctrine, and the corporate veil doctrine serve similar purposes, namely, the protection of creditors from shareholder dissipation of corporate assets. This commentator characterizes the latter three as “functional substitutes or technically necessary complements to

315 See id. at 857–58 (explaining the argument advanced by the party opposing voluntary piercing). Even if one holds the view that a corporation should not bear the unexpected costs of incorporation, it is doubtful whether D.H.N. deserved compensation in this case. Lord Justice Shaw referred to the fact that D.H.N. could have maintained all the benefits of incorporation and conveyed the land to itself the day before the compulsory purchase order was issued, and would have received full compensation. Id. at 867. D.H.N. had become aware of the possibility of a compulsory purchase as early as February 1970, and had eight months until October 1970—when the order was finally issued—to convey the land back to itself. See id. at 858. No explanation was given for why D.H.N. had failed to do so over this long period of time. See id. at 856. Having been given ample opportunity to protect itself and failing to do so, D.H.N.’s voluntary piercing claim should have been rejected. See id. at 857–58.
the law of fraudulent conveyances.” Given the narrow scope of the English corporate veil doctrine, one would expect creditor protection to be taken on by these other doctrines. Veil piercing and equitable subordination, in particular, have substantial overlaps with each other in the United States. Both doctrines are triggered by a host of similar considerations, such as fraudulent conduct by the controlling shareholder and inadequate capitalization of the corporation. Courts applying both doctrines have invoked the same language of “instrumentality” or “alter ego.”

Yet there is no equitable subordination under English insolvency law that allows the courts to subordinate a shareholder’s advances to the corporation below the creditors’ claims. In light of the functional equivalence of these two doctrines and the English courts’ reservations about veil piercing, the lack of an equitable subordination doctrine in English law is not altogether surprising. There are, however, a host of statutory provisions under English law that perform similar functions as the fraudulent conveyance laws in the United States. Therefore, the significance of a lack of equitable subordination in English law should not be overstated.

---

316 Clark, supra note 224, at 505.
317 See Landers, supra note 96, at 590.
318 See Clark, supra note 224, at 518; Landers, supra note 96, at 597–98.
319 See Landers, supra note 96, at 598.
320 See Clark, supra note 224, at 505.
321 See Landers, supra note 96, at 590.
322 O’Donovan, supra note 320, at 504.
323 Id. at 458. A number of examples include Sections 238 through 241, and 423 through 425 of the Insolvency Act. Sections 238 through 241 allow the courts to invalidate transactions that constitute an unfair preference for a particular creditor or that are undervalued. Id. at 458–59. Sections 423 through 425 set aside transactions that are at an undervalue and entered into for the purpose of putting corporate assets beyond the reach of its creditors. Roy Goode, Principles of Corporate Insolvency Law 503–05 (2005). These sections replace the fraudulent conveyance provisions in the Law of Property Act, the application of which to the mere payment of money was not clearly established. See Sir Kenneth Cork, Insolvency Law and Practice—Report of the Review Committee 276 (1982). These provisions seem largely to serve the same purpose as the fraudulent conveyance laws in the United States. See O’Donovan, supra note 320 at 458.
324 O’Donovan, supra note 320 at 504 (“Indeed, with the arsenal of remedies and defences available against lenders for misrepresentation, economic duress, breach of fiduciary duty, wrongful and fraudulent trading, fraudulent and voidable dispositions, [and] transactions at an undervalue, . . . there is scarcely any need for a doctrine of equitable subordination in English law.”).
2. Absence of Federalism Concerns

The second further difference between English and U.S. corporate veil cases is the absence of federalism concerns in the former. One issue that has been raised in some U.S. corporate veil cases, especially those involving the enforcement of a federal statute, is the extent to which the policy considerations behind the federal statute trump state corporation law.\(^{325}\) It has been argued that in some statutory veil piercing cases, the rule of separate corporate personality frustrates federal regulatory policy.\(^{326}\) *United States v. Bestfoods* showcases the conflict between state corporation law and federal regulatory policy.\(^{327}\) *Bestfoods* addressed the extent to which a parent corporation could be held liable for a subsidiary’s operation of a polluting facility under the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA).\(^{328}\) The Sixth Circuit applied Michigan corporate veil law and absolved the corporate parent in that case from liability in light of the court’s finding that the parent had not used the subsidiary to perpetrate fraud or to “subvert justice.”\(^{329}\) The Sixth Circuit further held that direct liability under CERCLA would only apply if the parent had sole or joint venture operation of a polluting facility.\(^{330}\) This holding is consistent with the traditional state law notion of separate corporate personality.

The Supreme Court, however, overturned the Sixth Circuit’s decision, and held that a parent corporation could be liable under CERCLA even when its participation in the polluting facility is short of sole and joint venture operation.\(^{331}\) This decision can be understood as reflecting a recognition that the demarcation of a corporation’s boundaries under state law would have obfuscated the federal policy behind CERCLA, which had been characterized as a “remedial statute de-


\(^{326}\) Ronald G. Aronovsky & Lynn D. Fuller, *Liability of Parent Corporations for Hazardous Substance Releases Under CERCLA*, 24 U.S.F.L. Rev. 421, 463–67 (1990); Note, *Piercing the Corporate Law Veil: The Alter Ego Doctrine Under Federal Common Law*, 95 Harv. L. Rev. 853, 871 (1982). But cf. Thompson, supra note 96, at 392. Thompson describes federalism as “an example of Latty’s triumph of a minor premise,” and argues that the alleged conflict between state corporation law and federal regulatory policy is “a conflict which does not exist.” *Id.* Thompson’s article was written before *Bestfoods* was handed down, however, and he expressly stated that it does not incorporate the *Bestfoods* decision. *Id.* at 379. It is possible that *Bestfoods* would have changed his view on the conflict.

\(^{327}\) See *Bestfoods*, 524 U.S. at 60.

\(^{328}\) *Id.*

\(^{329}\) *United States v. Cordova Chem. Co. of Mich.*, 113 F.3d 572, 582 (6th Cir. 1997).

\(^{330}\) *Id.* at 81.

\(^{331}\) *Bestfoods*, 524 U.S. at 66–67. The Supreme Court, however, did not rule on the parent corporation’s liability and instead remanded the issue to the trial court. *Id.* at 72–73.
signed to protect and preserve public health and the environment.” The Supreme Court circumvented this problem by adopting a more expansive interpretation of the term “operator” in the statute to override separate corporate personality under state law.

The absence of federalism concerns is hardly surprising given that the United Kingdom is not a federalist polity and does not comprise a federal legal system alongside subordinate state legal systems. Although the United Kingdom actually consists of two legal systems—England, Wales, and Northern Ireland, on the one hand, and Scotland, on the other—the two systems function in parallel. Moreover, there is only one body of company law for the entire United Kingdom. The Companies Act (2006) applies to the entire United Kingdom and there is no separate Scottish company law with which statutory policies pursued by Parliament may clash.

This is not to say that the notion of separate corporate personality under English company law does not clash with statutory policies. It only means that the interface between company law and statutory policies will often be different from the kind of interaction underpinned by federalism concerns that is seen in U.S. cases. In the United States, the interaction is often that of deference by state law to federal regulatory law. In the United Kingdom, company law need not yield to statutory policies as both are promulgated at the same level in the polity. Even then, given the hesitation of the English courts explicitly to discuss policy considerations in corporate veil cases, whenever a clash between statutory policies and separate corporate personality arises, it is seldom resolved with a direct recognition of the conflict. Instead, it is mediated through traditional common law concepts. For instance, In re (FG) Films could have been treated as a case featuring a conflict between separate corporate personality and the statutory policy of the

332 Cordova, 113 F.3d at 577.
334 See S.H. Smith et al., The Modern English Legal System 4 (5th ed. 2007).
335 Id. at 4.
337 See, e.g., Floyd, 151 F.3d at 1300 (explaining the problems associated with reverse piercing, but noting that ultimately the decision whether to adopt the doctrine rests with state corporate law).
340 See id.
Comparing the English & U.S. Corporate Veil Doctrines

Cinematograph Films Act to protect the British film market.\(^{341}\) There was little recognition of this conflict, however, and the court decided the case applying agency concepts.\(^{342}\)

III. A Reconciliation of the Two Corporate Veil Doctrines

Part II highlighted a number of important differences between the English and U.S. corporate veil doctrines. In particular, Part II explained that the prevailing approach to the corporate veil doctrine under English law involves classification into common law categories such as agency, trust, and fraud.\(^{343}\) This approach is distinctly unfamiliar to the U.S. corporation lawyer.\(^{344}\) Despite these differences, the two doctrines are in fact more similar than they seem at first glance. Many English corporate veil cases share the same mode of analysis used by U.S. courts under the instrumentality doctrine and the alter ego doctrine. In these cases, English courts have focused on the lack of economic substance of the company at issue and the presence of improper conduct, which are the two main prongs of the instrumentality doctrine.

A. The Instrumentality and Alter Ego Doctrines in U.S. Corporate Veil Jurisprudence

The instrumentality doctrine and the alter ego doctrine are the two most systematic analytical frameworks applied by U.S. courts in corporate veil cases. In Professor Thompson’s survey, the instrumentality doctrine was applied in seventy-five cases. Alternatively, the alter ego doctrine was applied in 181 cases.\(^{345}\) One of the most authoritative formulations of the instrumentality doctrine was laid down in Lowendahl v. Baltimore & Ohio Railroad.\(^{346}\) The Lowendahl court stated that three elements must be satisfied to establish a corporate veil claim:

1. Control, not mere majority or complete stock control, but complete domination, not only of finances, but of policy and business practice in respect to the transaction attacked so that

\(^{341}\) See id.

\(^{342}\) See id.

\(^{343}\) See supra text accompanying notes 124–135 (discussing English courts’ reliance on principles of agency and trust), 208–223 (discussing English courts’ application of fraud to the corporate veil doctrine).

\(^{344}\) See supra text accompanying notes 123–130 (discussing U.S. courts’ readiness to adopt new doctrines in corporate veil cases).

\(^{345}\) Thompson, supra note 7, at 1063.

\(^{346}\) See 247 A.D. at 156 (N.Y. App. Div. 1936), aff’d, 6 N.E.2d 56 (N.Y. 1936).
the corporate entity as to this transaction had at the time no separate mind, will or existence of its own; and (2) Such control must have been used by the defendant to commit fraud or wrong, to perpetrate the violation of a statutory or other positive legal duty, or a dishonest and unjust act in contravention of plaintiff’s legal rights; and (3) The aforesaid control and breach of duty must proximately cause the injury or unjust loss complained of.\footnote{347}

Accordingly, the instrumentality doctrine involves three elements. First, a corporation must be devoid of independence and economic substance.\footnote{348} Second, there must be an improper purpose or conduct.\footnote{349} And third, there must be a showing that the “instrumentalization” of the corporation proximately caused injury to the plaintiff.\footnote{350} Frederick Powell, who first articulated this doctrine, identified eleven circumstances that tend to show that a corporation is an instrumentality,\footnote{351} and seven situations that qualify as improper purposes.\footnote{352} The doctrine was originally conceived to apply only to parent-subsidiary relationships, but has since been extended to individual shareholder cases.\footnote{353} One commentator has proposed a two-prong test for veil piercing that

\footnote{347}Id.\footnote{348} See id.\footnote{349} See id.\footnote{350} See id.\footnote{351}Frederick J. Powell, Parent and Subsidiary Corporations 9 (1931). These circumstances are: (1) the parent corporation owns all or most of the capital stock of the subsidiary; (2) the parent and subsidiary corporations have common directors or officers; (3) the parent corporation finances the subsidiary; (4) the parent corporation subscribes to all of the capital stock of the subsidiary or otherwise causes its incorporation; (5) the subsidiary has grossly inadequate capital; (6) the parent corporation pays the salaries and other expenses or losses of the subsidiary; (7) the subsidiary has substantially no business except with the parent corporation or no assets except the ones conveyed to it by the parent corporation; (8) in the papers of the parent corporation or in the statements of the officers, the subsidiary is described as a department or division of the parent corporation, or its business or financial responsibility is referred to as the parent corporation’s own; (9) the parent corporation uses the property of the subsidiary as its own; (10) the directors or executives of the subsidiary do not act independently in the interest of the subsidiary but take their orders from the parent corporation in the latter’s interest; and (11) the formal legal requirements of the subsidiary are not observed. \textit{Id.}\footnote{352} Id. at 54. These situations are: (1) actual fraud, (2) violation of a statute, (3) stripping the subsidiary of its assets, (4) misrepresentation, (5) estoppels, (6) torts, and (7) other cases of wrong or injustice. \textit{Id.}\footnote{353}Krendl & Krendl, \textit{supra} note 57, at 11.
focuses on “disregard of formalities and fairness.” The proposed test is almost identical to the eleven factors identified by Powell.

The alter ego doctrine focuses on the lack of independence of a corporation from its shareholders. To show that a corporation is an alter ego, a plaintiff must show that the corporation and its shareholders are fundamentally indistinguishable. In Hamilton v. Water Whole International Corp., the U.S. Court of Appeals for the Tenth Circuit identified nine relevant factors for determining when a corporation becomes an alter ego. These factors largely overlap with the eleven identified by Powell. In many ways, the alter ego doctrine replicates the first prong in the instrumentality doctrine.

B. Construction of an Instrumentality Doctrine from the English Corporate Veil Cases

There are traces of the instrumentality and alter ego doctrines in some English corporate veil cases, which similarly focused on the economic substance of the corporation and the presence of impropriety. These cases, however, paid little attention to proximate causation. In cases in which the English courts pierced the veil, there was little doubt that the improper conduct caused the plaintiff’s injury, which largely obviated the need for a separate showing of proximate causation.

354 Barber, supra note 14, at 397.
355 See id. at 397–98.
356 Transition Healthcare Assocs., Inc. v. Tri-State Health Investors, 306 F. App’x 273, 280 (6th Cir. 2009).
357 302 F. App’x 789, 793 (10th Cir. 2008). These factors are:

(1) whether the dominant corporation owns or subscribes to all the subservient corporation’s stock, (2) whether the dominant and subservient corporations have common directors and officers, (3) whether the dominant corporation provides financing to the subservient corporation, (4) whether the subservient corporation is grossly undercapitalized, (5) whether the dominant corporation pays the salaries, expenses or losses of the subservient corporation, (6) whether most of the subservient corporation’s business is with the dominant corporation or the subservient corporation’s assets were conveyed from the dominant corporation, (7) whether the dominant corporation refers to the subservient corporation as a division or department, (8) whether the subservient corporation’s officers or directors follow the dominant corporation’s directions, and (9) whether the corporations observe the legal formalities for keeping the entities separate.

Id.

358 Compare id. with Powell, supra note 351, § 6 (sharing most factors in common).
359 See infra text accompanying notes 400–403.
In each of the three main categories of English corporate veil cases—agency, fraud, and evasion of existing legal obligations—there are cases in which the analysis parallels the first two prongs of the instrumentality doctrine. Examples among the fraud cases include *Gencor v. Dalby*,360 *Wallersteiner v. Moir*,361 and *Trustor v. Smallbone*.362 In these cases, the defendant shareholder committed fraud and therefore impropriety was clearly present.363 The courts, however, did not base their corporate veil analysis solely on the existence of fraud, even though the conduct at issue in those cases was outright fraudulent.364 In addition to the improper conduct, the courts emphasized the lack of economic substance of the companies.

In *Gencor*, one of the main issues was whether the sham company controlled by one of the officers, Mr. Dalby, was accountable to Gencor ACP for the commission paid to Dalby but deposited into the company.365 The commission could only have been recovered if the veil of the sham company was pierced.366 The court did so with little hesitation, focusing on Dalby’s complete control over the company and its lack of genuine business substance such as “sales force, technical team or other employees capable of carrying on any business.”367

In *Wallersteiner*, holding the fraudster in the case, Mr. Wallersteiner, liable entailed piercing the veil of a number of sham companies he had formed in Liechtenstein.368 Lord Denning characterized the companies as Mr. Wallersteiner’s agents on the grounds that Wallersteiner had eschewed corporate formalities, consummating “contracts of enormous magnitude on their behalf on a sheet of notepaper,” and that he had commingled his funds with theirs, using “their moneys as if they were his own.”369 Although Lord Denning characterized the company as an agent, he was not using the term in the traditional common law sense, but in the U.S. alter ego sense.370 In *Smallbone*, the court pierced the veil of the company after describing it as “simply a vehicle used for re-

360 Gencor ACP Ltd. v. Dalby, [2000] 2 B.C.L.C. 734 (Ch.) at 742–44 (Eng.).
362 Trustor AB v. Smallbone, [2001] 1 W.L.R. 1177 (Ch.) at 1177 (Eng.).
366 Id. at 740–42.
367 Id. at 744.
369 Id.
370 See id.
ceiving money from Trustor,” again emphasizing its lack of economic substance.

Within the agency category, the English case that best fits into the instrumentality doctrine is In re (FG) Films. In that case, a British company was created in an attempt to bypass a prohibition against the showing of non-British films in the United Kingdom. This constituted circumvention of a statute and would meet the impropriety prong of the instrumentality doctrine. Although the court characterized the English company as the U.S. company’s agent, it was clear that no genuine agency relationship existed between them. Meanwhile, the court emphasized the fact that the company had “no place of business apart from their registered office, and . . . did not employ any staff,” that it only had a capital of £100 when the film it had purportedly produced had cost £80,000 to make, and that it did not play any role in the production of the film. In the eyes of the court, the company was inadequately capitalized and lacked a genuine business operation and physical presence. These would likely be sufficient to satisfy the economic substance prong of the instrumentality doctrine.

Among the evasion of existing legal obligations cases, a number of them also applied analysis that largely parallels the first two prongs of the instrumentality doctrine. In these cases, the improper purpose prong is satisfied because upholding separate corporate personality would allow the defendant to evade its legal obligation, which would result in a wrong or injustice. In Gilford Motor Co. v. Horne, Judge Farwell set aside the company incorporated by Mr. Horne to escape his non-compete obligations with his former employer by focusing on the following facts: (1) the company only had two directors and shareholders, Horne’s wife and his former colleague; (2) the company had no physical premise of its own, and was only located in a garage in his home; and (3) the letterheads and notepapers of the company were practically

372 [1953] 1 W.L.R. 483 (Ch.) (Eng.).
373 See id. at 484.
374 See id.
375 Pickering, supra note 5, at 495 (explaining that the term agency “was used descriptively and no attempt was made to define the relationship in question in any precise legal sense”).
376 In re F.G. (Films), [1953] 1 W.L.R. at 486.
377 Id.
378 See id.
379 See id.
the same as those previously used by Horne himself.\textsuperscript{380} There was further evidence that his wife took no active role in the company’s management and that he ran the company as if it were his own.\textsuperscript{381} These facts would support a finding of a lack of economic substance under the U.S. corporate veil doctrine.

In \textit{Jones v. Lipman}, the defendant had contracted to sell property to the plaintiff, but subsequently changed his mind.\textsuperscript{382} To put the property beyond the plaintiff’s reach, the defendant transferred it to a company he had bought for this very purpose.\textsuperscript{383} Calling the company “the creature of the first defendant, a device and a sham, a mask which he holds before his face in an attempt to avoid recognition by the eye of equity,”\textsuperscript{384} the court did not hesitate to pierce the veil and ordered specific performance of the contract.\textsuperscript{385} The non-compliance with corporate formalities was serious. The company did not issue share capital or possess any real physical premise.\textsuperscript{386} Nor was any director appointed.\textsuperscript{387} It was obvious that this company had no genuine economic substance and was used solely to evade the defendant’s contractual obligation.\textsuperscript{388}

Aside from a wealth of cases in which the English courts employed analysis resembling the instrumentality doctrine, the doctrine also receives tacit support from a leading judgment by the House of Lords. There is a close parallel between the first two prongs of the instrumentality doctrine and Lord Keith’s statement in \textit{Woolfson} that veil piercing is only justified if the corporate entity is used as “a mere façade concealing the true facts.”\textsuperscript{389} This is one of the most authoritative formulations of a test under the English corporate veil doctrine, and has been repeatedly invoked by English courts in corporate veil cases.\textsuperscript{390} By referring to companies whose veil is pierced as a façade, Lord Keith emphasized

\begin{itemize}
\item \textsuperscript{380} [1933] Ch. 935 (A.C.) at 943, 954 (Eng.).
\item \textsuperscript{381} \textit{Id.} at 943.
\item \textsuperscript{382} [1962] 1 W.L.R. 832 (Ch.) at 833 (Eng.).
\item \textsuperscript{383} \textit{Id.} at 835.
\item \textsuperscript{384} \textit{Id.} at 836.
\item \textsuperscript{385} \textit{Id.} at 837.
\item \textsuperscript{386} \textit{Id.} at 835.
\item \textsuperscript{387} \textit{Id.}
\item \textsuperscript{388} \textit{Jones}, [1962] 1 W.L.R. at 835.
\item \textsuperscript{389} \textit{Woolfson} v. Strathclyde Reg’l Council, (1978) S.C.(H.L) 90, 96 (appeal taken from Scot.).
\end{itemize}
their lack of economic substance. A company with a genuine business operation and a high degree of economic independence would not likely be called a façade. The reference to concealment of the true facts at least covers fraud and misrepresentation, two of the seven situations qualifying as an improper purpose under Powell’s formulation.\textsuperscript{391} The reference can also be understood more broadly to encompass generally dishonest and deceptive conduct, such as asset stripping and circumvention of a statutory provision.\textsuperscript{392}

It should be clear by now that there are greater similarities in the approaches to veil piercing on both sides of the Atlantic than was previously assumed. In a way, this parallel is to be expected. If the question is whether the separate legal personality of a corporation should be set aside, it is difficult to avoid assessments of its economic substance and the propriety of the purpose for which the corporation is used. Labeling the corporation an agent, an instrumentality, or an alter ego does not alter the fundamental nature of the analysis.

There is ample support from case law for English courts to fashion their own version of the instrumentality doctrine. Doing so would help English courts provide clearer guidance to future litigants and would sharpen the focus of their analysis. Some modifications will be needed, however. First, given the English courts’ demonstrated reluctance to pierce the veil of a \textit{bona fide} company,\textsuperscript{393} the economic substance prong of an English instrumentality doctrine should be crafted in such a way that the existence of a \textit{bona fide} business operation precludes finding a lack of economic substance. Second, the preponderance of English cases involving evasion of an existing legal obligation\textsuperscript{394} means that evasion should be expressly incorporated as one of the improper purposes under the impropriety prong. A determination of whether an evasion has taken place can be made with reference to the modified evasion rule proposed above.\textsuperscript{395} Third, a decision needs to be made as to whether voluntary piercing should be permissible under this English instrumentality doctrine, and perhaps under the English corporate veil doctrine generally. The instrumentality doctrine was not formulated with volun-

\textsuperscript{391} See Powell, supra note 351, § 13.

\textsuperscript{392} Concealment of the true facts may be thought to cover circumvention of a statute because the corporate entity is used to conceal factual circumstances that would have called for an application of the statutory provision in question.


\textsuperscript{395} See supra text accompanying notes 283–294.
tary piercing in mind, and its focus on improper purposes arguably renders it unsuitable for these cases. Voluntary piercing usually entails the company pleading to pierce its own veil to obtain some benefits. There is usually no improper purpose or conduct of the kind that Powell enumerated. If English courts decide to continue to allow voluntary piercing claims, a separate analytical framework is probably needed. Nonetheless, in light of the discussion of the demerits of voluntary piercing above, there is a strong argument in favor of precluding voluntary piercing claims.

A remaining difference between the English cases and the U.S. instrumentality doctrine is the absence of discussions of proximate causation in the English cases. This omission is in some ways unremarkable given that many U.S. courts and commentators do not insist on a proximate causation requirement. Such a requirement was arguably

396 See Krendl & Krendl, supra note 57, at 11–14 (discussing the early development of the instrumentality doctrine through involuntary piercing cases); cf. Gaertner, supra note 296, at 685–87 (examining cases in which courts apply traditional veil piercing theory in the context of reverse piercing, although factors courts consider generally do not include “improper purpose”).

397 See Gaertner, supra note 296, at 681.

398 See supra text accompanying notes 310–315.


400 See, e.g., Laborers’ Pension Fund v. Lay-Com, Inc., 580 F.3d 602, 610 (7th Cir. 2009); Hamilton, 302 F. App’x at 793; Dusharm, 302 F. App’x at 572; Sonora Diamond Corp. v. Superior Court, 99 Cal. Rptr. 2d 824, 836 (Cl. App. 2000).

401 For example, Professor Barber’s two-prong test does not contain a proximate causation requirement. Barber, supra note 14, at 376–77. Many states likewise do not require proof of causation in a corporate veil case. See, e.g., Laborers’ Pension Fund, 580 F.3d at 610 (“Under Illinois law, [t]he standard test for piercing the corporate veil is two-pronged. The plaintiff must demonstrate both that there is such unity of interest and ownership between the individual or entity and the corporation that the separate personalities of the corporation and the individual [or entity] no longer exist, and that adherence to the fiction of separate corporate existence would sanction a fraud or promote injustice.”); Hamilton, 302 F. App’x at 793 (“The Oklahoma Supreme Court has held that a ‘corporation is so organized and controlled and its affairs so conducted that it is merely an instrumentality or adjunct of another corporation’ the two corporations will no longer be considered distinct legal entities.”); Dusharm, 302 F. App’x at 572 (To establish a corporate veil claim, “[p]laintiffs must prove . . . that observance of the corporate form would sanction a fraud or promote injustice.”) (quoting Gatecliff v. Great Republic Life Ins. Co., 821 P.2d 725, 729 (Ariz. 1991)); Sonora Diamond, 99 Cal. Rptr. 2d at 836 (“In California, two conditions must be met before the alter ego doctrine will be invoked. First, there must be such a unity of interest
unnecessary in the English cases discussed above. In most of them, there was little doubt that the defendant’s control of the company directly contributed to the plaintiff’s loss. In *Gencor*, *Wallersteiner*, and *Smallbone*, the company was an integral part of the fraudulent scheme to embezzle funds. In *Gilford Motor* and *Lipman*, the company played a pivotal role in the defendant’s attempt to evade its existing legal obligation. The question remains, however, whether such a requirement should be incorporated into an English instrumentality doctrine. Under most circumstances, requiring proof of proximate causation will create additional obstacles to plaintiff recovery and limit the availability of veil piercing. Given the English courts’ generally conservative attitude toward veil piercing, the inclusion of a proximate causation requirement of some kind seems appropriate.

If such a requirement is included, what should the causation link consist of? On one side is obviously the plaintiff’s loss. On the other side are two possibilities: the company’s lack of economic substance or impropriety. Some commentators argue that the plaintiff should be required to show that the lack of economic substance proximately caused its loss. This view, however, should be rejected. It would be extremely difficult for a plaintiff to prove that the overlap of corporate personnel or the sharing of physical premises proximately caused its loss. In most cases, the link between the lack of economic substance and the plaintiff’s loss is too tenuous to be demonstrated in court. Instead, the plaintiff should only be required to show that the improper conduct caused its loss. This represents a much more reasonable causation requirement.

What remains to be determined is the appropriate stringency of this requirement. Some commentators have criticized Powell’s formula-and ownership between the corporation and its equitable owner that the separate personalities of the corporation and the shareholder do not in reality exist. Second, there must be an inequitable result if the acts in question are treated as those of the corporation alone.”).  


404 See *Hackney & Benson*, supra note 234, at 854, 857–58.  

405 See, e.g., id. at 859–60 (“It is only logical that if a business cannot have assets taken out by its owner to the point where what is left is an unreasonably small capital, then the business should not commence activity in corporate form with such an unreasonably small capital.”); Hamilton supra note 156, at 986 (“In contract cases the separate existence of a minimally capitalized corporation should usually be recognized [sic].”); Posner, supra note 219, at 521 (“In general, a corporation’s creditors should be allowed to reach a shareholder’s assets when the shareholder, whether an individual or another corporation, has represented to the creditor that those assets are in fact available to satisfy any claim that the creditor may assert against the debtor corporation.”).
tion of the third prong—“wrongful acts resulting from the parent’s domination being the proximate cause of the plaintiff’s loss”—as too restrictive. Instead, such commentators have proposed a more liberal standard of “some reasonable connection between plaintiff’s injury and the action of the defendant.” The stringency of the third prong can be calibrated to reflect the courts’ view of the desirable scope of the corporate veil doctrine. It will need to be determined in future decisions.

IV. Revival of the Single Economic Unit Theory

Thus far, the discussion has emphasized the relative conservatism of the English corporate veil doctrine. There is in fact one area within the corporate veil doctrine in which the English courts have demonstrated greater progressiveness than even U.S. courts: the single economic unit theory. Under this theory, the court may impose general enterprise liability on a corporate group, something which even U.S. courts have been hesitant to do. Despite repeated suggestions of enterprise liability within the academic circle, U.S. courts have not developed distinct approaches for veil piercing within corporate groups. Professor Berle first proposed enterprise liability—the reassignment and imposition of liabilities within a corporate group based on a reconstruction of the economic boundary of the enterprise—in the 1940s. Subsequent commentators have further developed Berle’s proposal and called for some version of enterprise liability. Other commentators

406 Krendl & Krendl, supra note 57, at 27.
408 See Adolf A. Berle, Jr., The Theory of Enterprise Entity, 47 COLUM. L. REV. 343, 352–54 (1947); Thompson, supra note 96, at 388–92.
409 Berle, supra note 408, at 354.
410 See Blumberg, supra note 234, at 630 (“If a subsidiary corporation constitutes only one of a number of components of a corporate group collectively conducting a fragmented unitary business, the very basis for the establishment of limited liability as a matter of general legal policy disappears.”); Easterbrook & Fischel, supra note 234, at 111 (“Allowing creditors to reach the assets of parent corporations does not create unlimited liability for any people. Thus the benefits of diversification, liquidity, and monitoring by the capital market are unaffected.”); Landers, supra note 96, at 623 (“Moving beyond the threshold problem of limited liability for the protection of ultimate investors in the enterprise, separate consideration must be given to the application of limited liability to corporate groups. When limited liability is applied to multi-tiered corporate groups, layer upon layer of insulation from liability can result.”). But cf Posner, supra note 219 (arguing against Landers’ proposal); Thompson, supra note 234, at 38–39 (defending limited liability for corporate groups). Some scholars of English law have also advocated some version of enterprise liability. See Karl Hofstetter, Parent Responsibility for Subsidiary Corporations:
have advocated the imposition of shareholder liability on the corporate parent for its subsidiary's torts. Notwithstanding these proposals, U.S. courts have largely applied the same analytical framework to corporate and individual shareholders. In fact, Professor Thompson found that U.S. courts appear more hesitant to pierce the veil against a corporate shareholder compared to an individual shareholder.

In contrast, the single economic unit theory was specifically developed to address corporate groups. It was first propounded by Lord Denning in *D.H.N.* Unfortunately, he did not provide further elaboration of the theory in the opinion. His reasoning for treating the three group companies as one entity was thin. Nor did he suggest how to distinguish circumstances where the theory applies from circumstances where it does not. Furthermore, *D.H.N.* was a compensation case, and it was never entirely clear how the single economic unit theory would apply to attempts to impose shareholder liability on the corporate

---


413 See Thompson, *supra* note 7, at 1055–56. The plaintiff’s success rate against corporate parents was found to be thirty-seven percent and that against individual shareholders forty-three percent. See id. at 1055.


415 See *D.H.N.*, [1976] 1 W.L.R. at 850. *D.H.N.* notwithstanding, English judges have asserted on various occasions that there is no presumption that group companies be treated as a single economic unit or an enterprise under English company law. See Harold Holdsworth & Co. v. Caddies, [1955] 1 W.L.R. 352 (H.L.) at 363 (appeal taken from Scot.) (“[E]ach company in the group is in law a separate entity.”); *In re Southard & Co.*, [1979] 1 W.L.R. 1198 (A.C.) at 1208 (Eng.) (“English company law possesses some curious features, which may generate curious results. A parent company may spawn a number of subsidiary companies, all controlled directly or indirectly by the shareholders of the parent company. If one of the subsidiary companies, to change the metaphor, turns out to be the runt of the litter and declines into insolvency to the dismay of its creditors, the parent company and the other subsidiary companies may prosper to the joy of the shareholders without any liability for the debts of the insolvent subsidiary.”).


417 See id.

418 See id.
parent. Two years after *D.H.N.* was decided, the House of Lords openly questioned the validity of the theory, although it stopped short of overruling it. And the theory has been repeatedly criticized in subsequent cases. Despite its theoretical infirmities and questionable validity, the single economic unit theory has continued to feature in English corporate veil cases. If the theory can be reformulated with a more clearly defined scope and provided a theoretical foundation, it can be rendered a useful tool for English courts in corporate veil cases.

What follows is an attempt to justify the revival of the single economic unit theory in light of recent case law under both English company law and EU competition law, along with general insights from economic theory. Based on lessons drawn from these sources, Part IV.D proposes a more formal structure for the theory so that it is capable of more predictable application.

A. A Critique of the Theory

One needs to look no further than Lord Denning’s judgment itself to discern a central flaw in the single economic unit theory. He offered scant guidance on when group companies may constitute a single economic unit. To substantiate his conclusion, he relied on two main arguments. First, *D.H.N.*, the corporate parent, “can control every movement of the subsidiaries. These subsidiaries are bound hand and foot to the parent company and must do what the parent company says.”

Second, the group “is virtually the same as a partnership in which all the three companies are partners.” Control alone cannot be the deciding factor, for that would mean that the veil will be pierced between every wholly owned subsidiary and its parent. Moreover, no justification was

---

419 See id. at 857, 860.
421 See Vandekerckhove, supra note 17, at 72; see also Davies et al., supra note 15, at 202–03.
424 Id.
425 Lord Denning’s basis for concluding that *D.H.N.* controlled the subsidiary was that it had complete ownership of it. See id. at 856, 860. This argument was firmly rejected in *Gramophone & Typewriter, Ltd. v. Stanley*, See [1908] 2 K.B. 89 (A.C.) at 95–96 (Eng.). Many U.S. courts have similarly held that control alone does not justify veil piercing. See, e.g., Towe Antique Ford Found. v. Internal Revenue Serv., 999 F.2d. 1387, 1393 (2d. Cir. 1993); Kingston Dry Dock Co. v. Lake Champlain Transp. Co., 31 F.2d. 265, 267 (2d. Cir. 1929).
provided for applying partnership concepts to veil piercing cases. There is in fact an inherent contradiction between Lord Denning’s two arguments. Generally, partners are co-equals in a partnership and wield the same management powers. There is no presumption that one partner controls the others. By analogizing the three companies as partners, Lord Denning implicitly suggested that they are autonomous entities, which contradicts his first argument that the parent controls the subsidiaries.

B. Jurisprudential Support for a Revival of the Theory

For the purpose of reformulating the single economic unit theory, it is important to distinguish between voluntary piercing and involuntary piercing. The doctrinal framework applicable in these two situations will be different. So are the relevant theoretical considerations. In particular, economic analysis of limited liability is only relevant for involuntary piercing cases. Although voluntary piercing implicates separate corporate personality, it almost never bears on limited liability. As argued above, there is little theoretical justification for voluntary piercing. It was urged that English courts reconsider whether voluntary piercing should be permitted at all. Nevertheless, English courts continue to allow it, as was evident in the most recent corporate veil case Beckett Investment Management v. Hall.

1. Voluntary Piercing

There has always been firm jurisprudential support for applying the single economic unit theory to voluntary piercing cases. After all, the theory originated in such a case. As mentioned above, even though the validity of the theory was openly questioned in Woolfson, the House of Lords stopped short of overruling it. Beckett, a case decided

427 Henn & Alexander, supra note 217, at 70.
429 See supra text accompanying notes 310–315.
430 See supra text accompanying note 315.
by the English Court of Appeal in 2007, gave the theory a boost. In that case, the English Court of Appeal treated members of a corporate group as one entity in a voluntary piercing context. Specifically, the veil between group members was pierced at their behest to give effect to a covenant not to compete. Mr. Hall, the employee at issue in the case, signed an employment contract with the parent holding company, promising not to solicit its customers within a certain time period after departure from his employment. Unfortunately for the group, the parent did not have any customers. All of the group’s services were provided by the subsidiaries. Nevertheless, the covenant expressly only referred to the holding company’s customers. A literal construction of the covenant would have rendered it meaningless. This predicament seemed to have been the result of a drafting oversight by the group’s lawyers. The court came to the employer’s rescue, and treated members of the group as one entity for purposes of enforcing the covenant.

Even though the Beckett court did not invoke the single economic unit theory by name, the outcome of the case fits squarely within the theory. The court treated the parent and the two wholly owned subsidiaries as a single entity. The facts in Beckett were similar to those in D.H.N. in that both cases required the court to treat members of a corporate group as a single entity in order to allow the group to reap certain benefits. In D.H.N., the benefit was compensation for the disturbance of the parents’ business. In Beckett, it was the right to enforce a contractual covenant against a former employee. The justification for veil piercing is arguably even weaker in Beckett because the group’s predicament was entirely of its own making. In D.H.N. at least the compulsory purchase was beyond the group’s control. The fact that the English Court of Appeal pierced the veil in Beckett provides considerable support for the single economic unit theory.

435 Id. at 1541–43.
436 Id. at 1544.
437 See id.
438 See id. at 1545–46. The trial judge had refused to enforce the covenant because it only referred to the company, whereas other parts of the employment contract referred to the company and its subsidiaries. If the employer had intended the covenant to extend to the entire group, the judge reasoned, it would have so stated. Id. at 1545.
439 See id. at 1545–46.
This conclusion is further bolstered by a number of statements made by Lord Justice Kay, who decided *Beckett*. These statements indicate that his decision was based on more than mere contractual interpretation. Rather, he treated the companies as a single economic unit.\(^444\) To substantiate his decision, he cited Lord Denning’s judgment in *Littlewoods Organisation Ltd. v. Harris*,\(^445\) in which Lord Denning reasoned that “the law to-day has regard to the realities of big business. It takes the group as being one concern under one supreme control.”\(^446\) Lord Justice Kay proceeded to declare that “unlike the [High Court] judge, I do not feel inhibited by a purist approach to corporate personality.”\(^447\) One may argue that *Beckett* was not a genuine corporate veil case because there was no imposition of shareholder liability. Nevertheless, the English corporate veil doctrine has always encompassed more than mere shareholder liability cases.\(^448\) It has been regularly applied to situations similar to that in *Beckett*.\(^449\) The fact that the English Court of Appeal permitted a voluntary piercing claim—generally disfavored after *Woolfson*—further underscores the significance of this case and the support it lends to the single economic unit theory.

2. Involuntary Piercing

The support from recent English cases for applying the single economic theory to involuntary piercing is less direct. One case that supports such application is *Stone & Rolls Ltd. v. Moore Stephens*, in which the House of Lords tacitly applied the single economic unit theory to allow the attribution of the controlling shareholder’s fraudulent intentions to the company at issue. Outside of the United Kingdom, application of the theory to involuntary piercing finds support in a string of EU competition law cases. In fact, under EU law, there is a presumption of single economic unit if the parent owns all the shares of a subsidiary.

a. *The Stone & Rolls Ltd. v. Moore Stephens Decision*

In the summer of 2009, in its last case before its judicial functions were transferred to the recently constituted Supreme Court of the United Kingdom, the House of Lords handed down a controversial and

\(^{444}\) See id. at 1545–46.
\(^{445}\) [1977] 1 W.L.R. 1472 (A.C.) at 1482.
\(^{446}\) See *Beckett*, [2007] I.C.R. at 1546.
\(^{447}\) Id.
\(^{448}\) See supra text accompanying notes 100–119.
\(^{449}\) See supra sources listed in note 119.
sharply divided decision in Stone & Rolls Ltd. v. Moore Stephens, in which it pierced the veil between a company and its sole beneficial shareholder to impute his fraudulent intentions to the company. This imputation meant that the company’s professional negligence claims against its external auditors were barred under the tort law principle of ex turpi causa. The company at issue was Stone & Rolls (S&R), which was set up by Mr. Stojevic as part of a fraudulent scheme. The scheme resulted in a loss in excess of £94 million for many bank lenders. Stojevic had hired Moore Stephens to audit the company specifically to lend legitimacy to his fraudulent scheme. After the fraud was uncovered, the company promptly became insolvent and entered liquidation proceedings. The liquidators brought professional negligence claims against Moore Stephens, arguing that the auditors had performed their external audits negligently. For the ex turpi causa principle to apply to bar these claims, Stojevic’s fraudulent intentions must be treated as the company’s.

The ownership structure of S&R was complicated. The company had a corporate parent, an Isle of Man company, which in turn was owned by Stojevic’s family trust. The relationship between him and his family trust was shrouded in mystery. What was known was that Stojevic did not directly own any shares in S&R. Applying the ex turpi causa principle would require the House of Lords to treat Stojevic as S&R’s shareholder, which the majority did. In fact, the majority deemed him the company’s sole controlling shareholder. In doing so, the House implicitly treated S&R and its parent, Isle of Man company, as one en-

450 Stone & Rolls Ltd. v. Moore Stephens, [2009] UKHL 39, [54]–[56] [2009] 1 A.C. 1391 (H.L.) [54]–[56] (appeal taken from Eng.). It was a 3–2 decision in which Lord Scott of Foscote and Lord Mance dissented. Id. at [88], [206].

451 See id. at [86]. This principle states that a claimant cannot recover under a tort claim if pleading that claim requires the claimant to invoke its own illegal conduct. Id. at [20].

452 See id. at [1].

453 See id. at [3].

454 Id. at [4].

455 Id. at [3].

456 See Stone & Rolls, 1 A.C. 1391 (H.L.) at [1], [3] (noting liquidators’ argument that had Moore Stephens exercised due care, they would have uncovered the fraud earlier and the company’s losses would have been reduced).

457 Id. at [90].

458 See id. at [90], [113].

459 See id. at [20], [51], [86].

460 See id. at [2]. The minority disagreed with the majority on this point. Lord Mance, for one, thought that the ownership interest of S&R was not entirely clear. See id. at [34]. Nevertheless, there was no dispute that Stojevic was the sole director of the company and controlled its affairs. See id. at [209].
As Stojevic had no direct ownership interest in S&R, the only way to deem him a shareholder was to collapse the entities at issue into one economic unit. The House of Lords proceeded to impute Stojevic’s fraudulent state of mind to S&R and barred its claims against external auditors. The single economic unit theory was thus tacitly invoked to permit the majority to achieve its desired result. What was most surprising, given the English courts’ generally strict adherence to separate corporate personality, was how facilely the three Law Lords in the majority treated S&R and its parent company as one entity. After a brief mention of the Isle of Man parent in the beginning of the case, its existence was almost completely ignored for the remainder.

Even though *Stone & Rolls* was primarily concerned with attribution of shareholder intentions to the company, there was no mistake that separate corporate personality was implicated. The majority set aside the separate legal personalities of the various companies to treat them as one entity. It then pierced the veil of this entity to impute Stojevic’s intentions to it.

Interestingly, this is not a case of forward veil piercing—the imposition of shareholder responsibility for corporate liabilities—

---

461 See id. at [2], [113].
462 See *Stone & Rolls*, 1 A.C. 1391 (H.L.) at [54], [86].
463 Id. at [86]. The House of Lords clearly appreciated that veil piercing was at stake in the case. What divided the majority and the minority was whether S & R’s separate personality was to be respected. Lord Scott characterized the majority’s decision as “lifting the corporate veil and treating S & R as if it were Mr Stojevic himself who was seeking to repel the ex turpi causa defence” on the basis of his “absolute beneficial owner[ship] . . . coupled with his undoubted absolute managerial control.” Id. at [118]. Lord Scott proceeded to argue that the veil should not be pierced because it was unclear whether Stojevic was in fact an absolute beneficial shareholder. Id. Lord Mance similarly objected to the majority’s veil piercing, arguing that:

> I am also aware of no “policies or principles”, generally understood or not, which might limit a company’s recovery for a wrong done to it by reference to whatever loss its innocent shareholders might, if the corporate veil were lifted, be said themselves to have suffered. The suggestion that this could be the measure of a company’s recovery again ignores the company’s separate legal identity and interests.

Id. at [255]. Lord Scott echoed this sentiment, asserting that “S & R is a legal persona in its own right.” Id. at [118]. The disagreement over S & R’s status as a separate legal entity was highlighted by different characterizations of its role in the fraudulent scheme by the majority and the minority. Lord Phillips and Lord Walker were unequivocal in calling S & R a perpetrator of the fraud, whereas Lord Scott labeled it a co-victim of Mr. Stojevic’s conduct. Id. at [1], [115], [126].

464 See id. at [1]–[87].
465 See id. at [39]–[56].
466 See id. at [54].
which is the usual type of veil piercing.\textsuperscript{467} It instead involved reverse veil piercing. The company was deemed responsible for its shareholder’s conduct.\textsuperscript{468} Reverse piercing is generally considered more controversial than forward piercing because the company’s stakeholders must share the company’s assets with the shareholder’s creditors.\textsuperscript{469} Given that in most cases a company has more stakeholders—including shareholders and creditors—than an individual has creditors, and that a company does not control its shareholders’ conduct, reverse veil piercing affects more innocent parties and raises more fairness issues than forward piercing.\textsuperscript{470} The fact that the House of Lords countenanced reverse piercing in this case suggests a renewed readiness to set aside separate corporate personality when the circumstances so warrant.

b. \textit{EU Competition Law Cases}

EU courts have regularly applied the EU’s equivalent of the single economic unit theory in competition law (antitrust) cases to pierce the veil against corporate groups. In a recent decision, \textit{Akzo Nobel NV v. Commission}, the European Court of Justice affirmed a line of previous cases, including \textit{Stora Kopparbergs Bergslags AB v. Commission},\textsuperscript{471} which held that complete share ownership establishes a rebuttable presumption that the parent exercise control over the subsidiary.\textsuperscript{472} Absent a rebuttal of this presumption, the parent and the subsidiary will be treated as a single economic entity.\textsuperscript{473} In \textit{Akzo Nobel}, subsidiaries in the Akzo Nobel group, a Dutch chemical group, were fined for their participation in an international cartel concerning the chemical choline chloride.\textsuperscript{474} Liability of the subsidiaries was not in doubt as they had directly participated in the cartel. The European Commission sought to impose liability on the parent as well.\textsuperscript{475} In \textit{Michelin v. Commission}, the

\textsuperscript{467} See id.
\textsuperscript{468} See \textit{Stone & Rolls}, 1 A.C. 1391 (H.L.) at [54]. Recovery was highly probable had the House of Lords not barred the claims, given Moore Stephens’ admission of negligence in its audit. \textit{See id.} at [6].
\textsuperscript{469} See Henry Hansmann & Reinier Kraakman, \textit{The Essential Role of Organizational Law}, 110 \textit{Yale L.J.} 387, 398–405 (2000); \textit{see also} Youabian, \textit{supra} note 297, at 592–95.
\textsuperscript{470} See Youabian, \textit{supra} note 297, at 593–95.
\textsuperscript{473} \textit{Id.} at paras. 60–61.
\textsuperscript{474} \textit{Id.} at paras. 12–17.
\textsuperscript{475} \textit{Id.}
then-European Court of First Instance\textsuperscript{476} emphasized that “Community competition law recognises that different companies belonging to the same group form an economic unit and therefore an undertaking within the meaning of Article 81 EC and 82 EC if the companies do not determine independently their own conduct on the market.”\textsuperscript{477} In \textit{Akzo Nobel}, the European Court of Justice echoed this statement and asserted that “the concept of an undertaking, in the same context, must be understood as designating an economic unit even if in law that economic unit consists of several persons, natural or legal.”\textsuperscript{478}

The support from these EU cases for the single economic unit theory is unmistakable. Similar to Lord Denning’s reasoning in \textit{D.H.N.}, the European Court of Justice reaffirmed that a parent and its subsidiaries are deemed one economic unit by virtue of the parent’s complete control of the subsidiaries.\textsuperscript{479} In fact, the single economic entity approach under EU competition law goes one step further than its counterpart under English company law in two important ways. First, under the EU approach, complete share ownership establishes a rebuttable presumption that the parent exercises control over the subsidiary. The burden then falls on the parent to demonstrate that the subsidiary possesses commercial autonomy.\textsuperscript{480} No such presumption exists under the English single economic unit theory.\textsuperscript{481} Second, under the EU law, not only does the parent become liable for the subsidiary’s fine, its own revenue will also be taken into account for the calculation of the fine.\textsuperscript{482} In other words, the EU single economic entity approach not

\textsuperscript{476} The Court of First Instance was renamed the General Court. See Consolidated Version of the Treaty on the Functioning of the European Union art. 256, Sep. 5, 2008, [2008] O.J. 47.


\textsuperscript{478} \textit{Akzo Nobel}, 2009 E.C.R. I-8237 at para. 55.

\textsuperscript{479} \textit{Id.} at paras. 58–62.

\textsuperscript{480} \textit{Id.} at paras. 60–61.

\textsuperscript{481} In fact, in \textit{Gramophone and Typewriter v. Stanley}, the notion that complete share ownership of a subsidiary confers control over the parent was specifically rejected. [1908] 2 K.B. at 93–94, 95–96, 98–100.

\textsuperscript{482} Council Regulation 17, First Regulation Implementing Articles 85 and 86 of the Treaty, art. 15, 1962 J.O. (13) 204 (EC), available at http://eur-lex.europa.eu/RECH_reference_pub.do (type in year as “1962,” OJ number as “13,” and page number as “204”; search for “all OJ series”; click hyperlink to only result). Assume that the subsidiary’s turnover was one hundred million Euros in the relevant time period, and the parent’s turnover was the same. Further assume that the Commission determines the fine to be ten percent of the undertaking’s turnover. After the application of the single economic entity approach, the fine for the group increases from ten million Euros to twenty million Euros as the undertaking’s revenue is now deemed to be two hundred million Euros.
only shifts liability, it actually expands it. As a result, not only is limited liability abrogated, the liability of the group also increases.

C. Economic Support for a Revival of the Theory

Economic analysis of limited liability only pertains to involuntary piercing, and specifically to the imposition of shareholder liability. This should be obvious because limited liability, with which economic analysis is concerned, is only implicated in involuntary piercing cases. Economic analysis therefore has little relevance for voluntary piercing. The discussion below focuses on the transaction cost implications of the single economic unit theory. It analyzes how application of the theory will affect monitoring costs and risk allocation between the corporate parent of a wholly owned subsidiary and the subsidiary’s creditors. The discussion begins by focusing on the imposition of a subsidiary’s liabilities on the corporate parent, or forward piercing. It then extends to a possible full-blown enterprise liability theory that covers forward piercing, reverse piercing, and lateral piercing between affiliated corporations.

1. Corporate Parent Liability for Wholly Owned Subsidiaries

From an economic perspective, the optimal corporate liability rule is determined by reference to a number of attributes of the various stakeholder or claimant groups within a firm. These claimants include the shareholders and the creditors. The creditors can be categorized into involuntary creditors, who are mainly tort victims, and voluntary creditors. The voluntary creditors can in turn be subdivided into financial creditors, trade creditors, and employees. The relevant attributes are their monitoring costs and risk-bearing abilities.

---


484 See Easterbrook & Fischel, supra note 234, at 99–101, 103–07 (describing the varying costs to creditors, shareholders, and other members associated with different liability regimes); Halpern et al., supra note 35, at 133–36 (noting the liability regime’s impact on the magnitude of information costs, which in turn affects creditors’ expected rate of return); Leebron, supra note 234, at 1588–1605 (discussing risk allocation among corporate members under different liability regimes); Posner, supra note 219, at 507–09 (analyzing risk allocation under different liability regimes).

485 See Easterbrook & Fischel, supra note 234, at 99.

486 See id. at 104, 107.

487 Id. at 104.

488 See id. at 104–06.
Monitoring costs refer to the costs incurred by a corporate claimant to monitor the management. Monitoring is necessary to guard against fraud, self-interested transactions, and general mismanagement. Risk-bearing abilities refer to a corporate claimant’s ability to diversify away potential exposure to corporate insolvency through investments in other financial instruments. An efficient corporate liability rule is one that achieves an optimal level of firm-wide monitoring at the lowest aggregate monitoring costs, and that allocates risks to the corporate claimant group with the highest risk-bearing abilities.

To determine the optimal corporate liability rule for monitoring costs, the analysis must compare the information costs of the various groups of creditors on the one hand and the shareholders on the other hand, and also compare the aggregate monitoring costs of the firm under alternative rules of liability. Information costs refer to the costs incurred by a party to obtain and analyze information about the firm. A claimant group’s monitoring costs can be determined by taking into account its information costs and monitoring level. With respect to information costs, if the shareholders have higher information costs than the creditors as a whole, then it would be more economically efficient to place the burden of monitoring on the creditors. If, on the other hand, the creditors have higher information costs than the shareholders, the shareholders should monitor instead. The optimal liability rule shifts losses in the event of bankruptcy to the claimant group that has lower information costs so that it has the incentive to undertake more monitoring.

With respect to the monitoring level, the choice of liability rule may affect firm-wide monitoring costs in a number of ways. First—and related to information costs—shifting potential losses between shareholders and creditors alters their incentives to monitor. Second, it affects their scope of monitoring. If, following a change of liability rules, a

489 See Leebron, supra note 234, at 1605–08; see also Posner, supra note 219, at 507–09 (describing the costs associated with obtaining information and supervising a corporation’s internal affairs).
490 See Leebron, supra note 234, at 1606.
491 See id. at 1588–1605.
492 See Posner, supra note 219, at 507–09.
496 See Leebron, supra note 234, at 1605–07.
497 See Easterbrook & Fischel, supra note 234, at 115.
claimant group has to monitor a wider range of issues, its overall monitoring level will be higher.\textsuperscript{498} Third, shifting losses changes the information costs of the various corporate claimants. A change of liability rule may alter information costs by shifting the relative bargaining powers between the outside creditors and the firm.\textsuperscript{499} Under a rule of unlimited liability, for example, the outside creditors have the liability rule in their favor. If the shareholders want to contract out of unlimited liability, the creditors will be in a position to demand concessions, such as greater access to information. This reduces the creditors’ information costs.

Applying the above concepts and analytical framework, the corporate parent clearly has lower information costs than any creditor. The case of involuntary creditors is easy to dispose of as these creditors rarely have the opportunity to monitor the corporation.\textsuperscript{500} In most tort cases, the victim has no prior knowledge of the identity of the tortfeasor.\textsuperscript{501} Even for voluntary creditors—including the most sophisticated ones such as institutional investors and lenders—their access to information is likely to be more limited and costly than that of a corporate parent, which will have practically unfettered access to information from the subsidiary.\textsuperscript{502} In fact, it makes little sense even to speak of monitoring because the parent controls the subsidiary. The same entity is both monitoring and being monitored. Therefore, the optimal liability rule from a monitoring cost perspective is one that imposes liability on the corporate parent for the subsidiary’s liabilities.

Many commentators have observed that unlimited liability will increase monitoring costs because it will cause shareholders to monitor the wealth of fellow shareholders.\textsuperscript{503} This is important information to shareholders because their share of liability will be contingent on fellow shareholders’ wealth, should unlimited liability be joint and several.\textsuperscript{504} This observation has little relevance for the single economic unit theory for two reasons. First, many commentators have pointed out that the need to monitor shareholder wealth will be obviated under a pro

\begin{itemize}
  \item \textsuperscript{498} See id. at 90–91.
  \item \textsuperscript{499} See id. at 99–101.
  \item \textsuperscript{500} See Halpern et al., supra note 35, at 145–47.
  \item \textsuperscript{501} See Leebron, supra note 234, at 1601 n.114.
  \item \textsuperscript{502} See Easterbrook & Fischel, supra note 234, at 99–101, 105–06.
  \item \textsuperscript{503} See id. at 99–100; Halpern et al., supra note 35, at 130–31.
  \item \textsuperscript{504} See Easterbrook & Fischel, supra note 234, at 93–97; Halpern et al., supra note 35, at 136–38.
\end{itemize}
rata rule. Second, and more importantly, given that the single economic unit theory only applies to wholly owned subsidiaries, there is only one shareholder and no one else’s wealth that the corporate parent needs to monitor.

Some commentators argue that unlimited liability will cause the subsidiary’s creditors to monitor the shareholders’ wealth—or in the case of the single economic unit theory, the corporate parent’s credit worthiness. Although that may be true, a wholly owned subsidiary’s creditors only need to monitor the wealth of one shareholder: the corporate parent. The increase in monitoring costs would not be substantial. Therefore, from a monitoring cost perspective, the single economic unit theory would be an efficient rule because it shifts incentives to monitor to the corporate parent, which has the lowest information costs. It thus reduces the aggregate monitoring costs of the subsidiary.

To determine the optimal corporate liability rule from the perspective of efficient allocation of risks, the analysis entails identifying the corporate claimant group with the greatest ability to diversify away the non-systematic risks of its investment in the firm. Under the Capital Asset Pricing Model, each investment entails systematic risks, which arise from fluctuations in the general economy. They also entail non-systematic risks, which are peculiar to a particular firm or business sector. Non-systematic risks of a particular investment can be diversified away through the purchase of another investment that tends to move in the opposite direction from the investment at issue. Systematic risks, meanwhile, cannot be eliminated through diversification. A corporate claimant’s ability to diversify away the non-systematic risks of his or her investment in the firm depends on a number of factors, most notably the proportion of investment in his or her overall asset portfolio. If an individual has personal assets of one million dollars and invests fifty percent of them in a firm, it is unlikely that diversifi-

506 See Leebron, supra note 234, at 1605–08.
507 See Halpern et al., supra note 35, at 134; Posner, supra note 219, at 516–17.
508 See Halpern et al., supra note 35, at 133–35.
509 See Easterbrook & Fischel, supra note 234, at 96–97; Leebron, supra note 234, at 1595–1600.
511 Id. at 269–70.
512 Id.
513 See id. at 269–70, 293–94.
tion will eliminate the non-systematic risks of that investment. If that individual invests one percent of his or her assets to buy the shares of a listed corporation, he should be able to diversify away the non-systematic risks through purchases of other financial assets.

It may seem that the sole shareholder of a corporation is least able to diversify away risks. This is because its investment in the firm is likely to account for a greater proportion of its overall asset portfolio in comparison to the exposure of financial lenders or other creditors to this firm. The logical conclusion would be that the optimal rule from a risk allocation perspective would be to provide limited liability, which may undermine the premise of the single economic unit theory. This argument, however, is invalid for two reasons. First, while the above reasoning may be true for an individual shareholder, it is probably untrue for the parent of a *bona fide* corporate group, which has considerable ability to diversify its portfolio. In fact, in some instances, the raison d’être of corporate groups or conglomerates is diversification. Therefore, the corporate parent of a subsidiary is likely to have no less ability to diversify than any outside creditors. Second—and more importantly—is that in the context of the single economic unit theory, the shareholder at issue controls the firm. It has the ability to manage its risk exposure to the subsidiary. If a particular corporate claimant can control its own risk exposure, it would be economically efficient to shift corporate insolvency risks to the corporate claimant. This means that a corporate parent should be liable for its subsidiary’s debts, which supports the use of the single economic unit theory as a basis for involuntary piercing within corporate groups.

Another consideration raised in the economic literature is the role of limited liability in attracting corporate investment. It has been argued that one of the most persuasive justifications for limited liability is that without it, investors would shy away from corporate investments for fear of exposing themselves to massive liability payments. This would cause corporate investments to dry up, which would be highly detrimental to businesses. These scholars have in mind passive individual

---

515 See id.
516 See id. at 1629–30.
517 See id.
519 See Leebron, supra note 234, at 1629–30.
investors in publicly traded corporations.\textsuperscript{521} This concern is unlikely to matter in the situations covered by the single economic unit theory. First, the firms covered by the theory are wholly owned subsidiaries. They are not publicly traded.\textsuperscript{522} Second, the only investor at issue in these companies is the corporate parent, which is anything but passive. If its subsidiary may expose it to shareholder liability, it has the power to prevent such liability from materializing given its complete control over the subsidiary’s operations.\textsuperscript{523}

In sum, transaction cost analysis of limited liability lends strong support to the use of the single economic unit theory to hold a corporate parent liable for its wholly owned subsidiaries’ debts. Other commentators have made similar suggestions in the past.\textsuperscript{524} What remains to be seen is whether economic analysis also supports the introduction of full enterprise liability under the single economic unit theory.

2. Enterprise Liability

The single economic unit theory need not be confined to the imposition of subsidiary liability on a corporate parent. One commentator famously proposed single enterprise liability for a corporate parent and all of its wholly owned subsidiaries.\textsuperscript{525} The important question is whether the single economic unit theory should be similarly extended to reverse piercing and lateral piercing, in addition to the forward piercing considered above.

Other commentators have argued that separate corporate personality\textsuperscript{526} is important because it allows economically efficient asset partitioning, which they define as “the designation of a separate pool of assets that are associated with the firm, and that are distinct from the personal assets of the firm’s owners and managers . . . [and] the assignment to creditors of priorities in the distinct pools of assets that re-

\textsuperscript{521} See Easterbrook & Fischel, supra note 234, at 95.
\textsuperscript{522} Cf. Akzo Nobel, 2009 E.C.R. I-8237, at paras. 59–61 (describing the relationship of a parent company to its subsidiary, distribution of shares, and extent of control).
\textsuperscript{523} See id. at para. 60; Leebron, supra note 234, at 1627.
\textsuperscript{524} See, e.g., Blumberg, supra note 234, at 623–34; see also Hansmann & Kraakman, supra note 234, at 1882–90 (arguing for use of single economic unit theory in tort claims); Leebron, supra note 234, at 1588–95.
\textsuperscript{525} See Landers, supra note 96, at 590.
\textsuperscript{526} See Hansmann & Kraakman, supra note 469, at 393. In fact, they use a different terminology than “separate corporate personality.” For the sake of clarity, however, this Article will continue to use “separate corporate personality,” the terminology familiar to most readers.
result from the formation of a legal entity.”527 Asset partitioning is efficient because it reduces monitoring costs.528 There are two types of asset partitioning: affirmative and defensive.529 Affirmative asset partitioning refers to the assignment “to the firm’s creditors a claim on the assets associated with the firm’s operations that is prior to the claims of the personal creditors of the firm’s owners,” and defensive asset partitioning refers to “granting to the owners’ personal creditors a claim on the owners’ separate personal assets that is prior to the claims of the firm’s creditors.”530 As a result, forward piercing overrides defensive asset partitioning, and reverse piercing trumps affirmative asset partitioning.531

According to these commentators, affirmative asset partitioning is the most important attribute of corporation law.532 Without affirmative asset partitioning, the creditor of any shareholder can proceed against the corporate assets and petition for the firm’s liquidation.533 Therefore, the firm’s creditors will need to monitor each shareholder’s creditworthiness along with the firm’s.534 To make matters worse, the firm’s creditors will also need to monitor the investments made by each shareholder because the firm’s assets may be called upon as a result of a failed investment by any shareholder.535 The firm’s creditors will not be the only ones with incentives to conduct such monitoring. The shareholders will also monitor each other.536 Reverse piercing, which overrides affirmative asset partitioning, will hence result in a substantial increase in monitoring costs. In contrast, defensive asset partitioning, which implicates limited liability, is less essential and the argument for it less conclusive.537 This explains why there are different variations of defensive asset partitioning, while affirmative asset partitioning is fully preserved under practically all circumstances.538

527 Id.
528 Id. at 398.
529 Id. at 393.
530 Id.
531 See id.; Youabian, supra note 297, at 574–76 (explaining that in forward piercing cases the stockholders of a corporation may be held personally liable for corporate obligations, whereas in reverse piercing cases the corporation may be held liable for the debts of its shareholders).
532 See Hansmann & Kraakman, supra note 469, at 406.
533 Id. at 402.
534 Id.
535 Id.
536 Id. at 402–03.
537 See id. at 423, 428–32.
538 See Hansmann & Kraakman, supra note 469, at 423.
The implication of the foregoing discussion for the single economic unit theory and the corporate veil doctrine generally is that forward piercing is economically more defensible than reverse piercing. Nonetheless, the impact of reverse piercing under the single economic unit theory on the benefits of affirmative asset partitioning may not be as serious as has been suggested. The increase in monitoring costs will be much less substantial in the case of a wholly owned subsidiary. The commentators referred to above have acknowledged this. Given that there is only one shareholder, there will be no cross-shareholder monitoring. Subsidiary creditors’ monitoring of the parent will be more manageable, even though they will still need to monitor the corporate parent’s risk exposure, including the finances of affiliated corporations within the group. A separate justification for allowing reverse piercing, at least under some limited circumstances, is that otherwise corporate parents will simply move their assets to their subsidiaries to avoid liability. The parent will itself contract liabilities with outside creditors while keeping its assets in the subsidiaries. A categorical refusal to sanction reverse piercing would be tantamount to an open invitation for manipulation.

From an economic perspective, whether lateral piercing between two affiliated corporations should be allowed depends on the degree of integration between the two corporations’ businesses. With respect to affiliated corporations, one can align them by the degree of integration of their businesses. On one end of the spectrum are corporations that operate in the same line of business, perhaps as producers of different components of the same product. These may be called integrated subsidiaries. On the other end of the spectrum are members of a conglomerate whose businesses are unrelated to each other. These corporations may be called conglomerate or nonintegrated subsidiaries.

---

539 See id. at 398–405. Hansmann and Kraakman identified additional benefits for defensive asset partitioning, namely decision-making economies, enhanced creditor monitoring, economies of transfer, and risk-bearing economies. Id. at 424–27. These benefits are likely to be negligible, however, for a wholly owned subsidiary and its corporate parent. Id. at 402.

540 See id. (“The advantages of asset partitioning, and particularly of affirmative asset partitioning, are far more obvious in the case of a business firm that has numerous individuals as owners.”).

541 See id.

542 See Hansmann & Kraakman, supra note 469, at 402.

543 See Leebron, supra note 234, at 1623.

544 See id. at 1613–14, 1623.

545 See id. at 1616–17.

546 See Leebron, supra note 234, at 1616–17.
General Electric is the quintessential conglomerate.\(^{547}\) The disaggregation of a single line of business into distinct legal entities yields few economic benefits in terms of diversification.\(^{548}\) Hence, there is a strong argument for lateral piercing between integrated subsidiaries. The argument for limited liability is greater for conglomerate subsidiaries, no less because it will be easier for creditors to monitor affiliated corporations with distinct businesses than for them to monitor one subsidiary operating two businesses.\(^{549}\) In the former scenario, the two creditor groups need not spend time to familiarize themselves with another line of business.\(^{550}\) As an illustration, the argument for lateral piercing between the two subsidiaries in *D.H.N.* would be strong, given that they had no independent business of their own and merely owned different parts of the same grocery business.\(^{551}\)

In summary, the economic justifications for forward piercing under the single economic unit theory are strong. The theory should be made readily available for the imposition of subsidiary liabilities on a corporate parent. The argument is less conclusive for reverse piercing, in light of the impact it has on the benefits of affirmative asset partitioning.\(^{552}\) Still, the need to forestall circumvention means that reverse piercing must be permissible at least under some circumstances. Whether lateral piercing should be permitted depends on the degree of integration between the affiliated corporations.\(^{553}\) Separate corporate personalities should be upheld for affiliated corporations with dis-

\(^{547}\) See id. For information about General Electric’s diverse range of business, which includes aircraft engines, finance, healthcare, security, oil and gas, media, and entertainment, see http://www.ge.com/products_services/index.html.

\(^{548}\) See Leebron, supra note 234, at 1616–17 (“From a financial and commercial perspective, such entities are not truly independent. Investments in them do not represent a diversification of risk, and thus there is no reason to recognize a regime of limited liability for them, at least with respect to tort claimants.”). Professor Leebron’s proposal was confined to tort claims, hence the caveat at the end of his statement. Nevertheless, his reasoning is equally valid for contractual claims. He also argued that upholding limited liability for contractual claims between a corporate parent and its wholly owned subsidiary serves the purpose of isolation of assets for the purpose of secured borrowing. Therefore, limited liability should continue to apply to the claims of financial creditors. *Id.* at 1614. Although this argument may be valid for financial creditors, it has little applicability for trade creditors and employees. Taking these latter two contractual creditors into account, it is not so clear that limited liability protection for a corporate parent against its wholly owned subsidiary’s contractual claims is economically efficient.

\(^{549}\) See Hansmann & Kraakman, supra note 469, at 399–401.

\(^{550}\) See id.


\(^{552}\) See supra notes 532–538 and accompanying text.

\(^{553}\) See Leebron, supra note 234, at 1616–17.
tinct lines of business. Veil piercing should be allowed against integrated subsidiaries.

D. Reformulation of the Theory

1. Involuntary Piercing

Thus far, the discussion of the single economic unit theory has proceeded on the premise that it only applies between a corporate parent and its wholly owned subsidiaries. D.H.N. and Beckett both featured such groups. One problem with this narrow scope of application is that it exposes the theory to easy circumvention. There was no concern about circumvention in those two cases because they both involved voluntary piercing. In an involuntary piercing context, however, a corporate parent bent on avoiding shareholder liability could do so simply by adding a nominal shareholder to the subsidiary. To preempt such circumvention, the theory needs to be modified so that it continues to apply to cases where the additional shareholders are merely nominal. To escape the reach of the theory, the shareholders other than the corporate parent must evince meaningful independence from the corporate parent.

Complete, or virtually complete, share ownership is only a necessary—but not sufficient—condition for invoking the theory. Only when the corporate parent actually controls the subsidiary can it be fairly deemed responsible for the subsidiary’s conduct and liabilities. A requirement of control is consistent with economic analysis. Conclusions about the corporate parent’s lower information costs and superior risk-bearing abilities are both premised on its control over the subsidiary, as is the argument that shareholder liability will not deter a corporate parent from investing in its subsidiary. Without actual control, a corporate parent may no longer be the least-cost monitor or optimal risk-bearer. The economic justifications for applying the single economic unit theory will be weaker. Control must entail more than the mere

554 See Hansmann & Kraakman, supra note 469, at 399–401.
557 See Youabian, supra note 297, at 594.
558 See Easterbrook & Fischel, supra note 234, at 99–101; Leebron, supra note 234, at 1574–78, 1630.
power to elect directors.\textsuperscript{559} A corporate parent is only deemed to exercise control over a subsidiary if it can command its daily operations.\textsuperscript{560} The mere ability to elect directors may not allow the parent to lower its information costs and control its risk exposure to the subsidiary effectively. Moreover, every corporate parent with complete or close to complete share ownership of its subsidiaries has the power to elect directors.\textsuperscript{561} If mere ability to elect directors qualifies as control, it would practically eviscerate the control requirement.

Nevertheless, it remains to be determined which party should bear the burden of proving or negating control. From an evidentiary perspective, the burden of proof should rest on the corporate parent because it will have greater access to evidence that is needed for disproving its control of the subsidiary.\textsuperscript{562} This may be done by showing that the board exercises independent judgment in the daily operations of the subsidiary or that the subsidiary has acted against the parent’s wish in the past.\textsuperscript{563} To invoke the single economic unit theory in involuntary piercing cases according to this theory, then, the plaintiff will first be required to demonstrate the absence of independent shareholders in the subsidiary, which establishes a presumption of actual control by the corporate parent. The burden of proof then shifts to the parent to rebut the presumption of control.

The elements of the theory that have been discussed so far pertain to the independence and integrity of the subsidiary, or to use the terminology of the instrumentality doctrine—the economic substance of the subsidiary. As a variation of the corporate veil doctrine, the single economic unit theory should require an improper conduct or purpose, at least in the involuntary piercing context. This parallels the impropriety prong of the instrumentality doctrine.\textsuperscript{564} Examples of improper conduct include fraud, evasion of existing legal obligations, and cir-

\textsuperscript{559} Cf. Gramophone, [1908] 2 K.B.at 95–96 (noting that the power to elect directors alone does not confer corporate control).

\textsuperscript{560} See United States v. Bestfoods, 524 U.S. 51, 55 (1998) (holding that a corporate parent that actively exercised control over the operations of a subsidiary’s facility may be directly liable as an operator of that facility).

\textsuperscript{561} See Douglas & Shanks, supra note 1, at 196.

\textsuperscript{562} Cf. Akzo Nobel, 2009 E.C.R. I-8237, at paras. 60–61 (holding that the parent company has the burden of rebutting a presumption of liability for fines imposed on its subsidiaries).

\textsuperscript{563} One example of such independence is found in Craig v. Lake Asbestos of Quebec Ltd., where the parent could not prevent the subsidiary from making acquisitions without the parent’s consent. 843 F.2d 145, 152 (3d Cir. 1988).

\textsuperscript{564} See supra notes 347–350 and accompanying text.
cumvention of statutes.\textsuperscript{565} The impropriety requirement distinguishes the single economic unit theory from the kind of general enterprise liability proposed by some commentators.\textsuperscript{566} Although the economic analysis presented earlier may lend support to general enterprise liability, imposition of shareholder liability absent impropriety would be a dramatic extension of the English corporate veil doctrine.\textsuperscript{567} Despite some support for this approach in \textit{Rainham Chemical Works}, in which shareholder liability was imposed for a corporate tort in the absence of impropriety,\textsuperscript{568} the weight of English case law clearly indicates that impropriety must be present before shareholder liability can be imposed.\textsuperscript{569} Therefore, once control on the part of the corporate parent is established, the burden of proof shifts back to the plaintiff to demonstrate the presence of impropriety. Lastly, in keeping with an earlier suggestion of incorporating a proximate causation requirement in an English instrumentality doctrine,\textsuperscript{570} the plaintiff should additionally be required to show that the improper conduct by the parent contributes to its loss. Once all these elements are met, the single economic unit theory applies to impose liability on the corporate parent.

Both reverse piercing and lateral piercing should only be permitted under limited circumstances. In particular, they should only be allowed if there is evidence of asset stripping to evade liability. \textit{Yukong Line Ltd. of Korea v. Rendsburg Investments Corp. of Liberia} illustrates how lateral piercing can be crucial for preventing liability evasion.\textsuperscript{571} In that case, Rendsburg repudiated a charter-party with Yukong after market conditions had deteriorated.\textsuperscript{572} On the same day, Rendsburg’s shareholder transferred its assets to another company controlled by him.\textsuperscript{573} Although the shareholder at issue was an individual,\textsuperscript{574} one can easily imagine the same sequence of events transpiring with a corporate parent. Without lateral piercing, a corporate parent could do exactly what

\begin{itemize}
  \item \textsuperscript{566} See, \textit{e.g.}, Blumberg, \textit{supra} note 234, at 621–26; Landers, \textit{supra} note 96, at 651–52.
  \item \textsuperscript{567} See \textit{Stone & Rolls Ltd.}, 1 A.C. 1391 (H.L.) at [104]—[112]. It would also be a dramatic extension of the U.S. doctrine. See Thompson, \textit{supra} note 7, at 1041.
  \item \textsuperscript{568} Rainham Chem. Works, Ltd. v. Belvedere Fish Guano Co., [1921] 2 A.C. 465 (H.L.) at 470–79 (appeal taken from Eng.).
  \item \textsuperscript{569} See \textit{supra} notes 389–392 and accompanying text.
  \item \textsuperscript{570} See \textit{supra} text accompanying notes 400–406.
  \item \textsuperscript{571} [1998] 1 W.L.R. 294 (Q.B.) at 306. The shareholder in the case, however, was not a company, but an individual.
  \item \textsuperscript{572} \textit{Id.} at 297–98.
  \item \textsuperscript{573} \textit{Id.} at 298.
  \item \textsuperscript{574} \textit{Id.} at 302.
\end{itemize}
was done in *Yukong* to evade liability. One may question why lateral piercing should be confined to asset stripping situations and not be extended to all integrated subsidiaries. Although economic analysis suggests that lateral piercing should be available against integrated subsidiaries, determining the degree of economic integration of affiliated corporations is fraught with difficulty. It requires substantial business expertise on the part of the courts.\(^{575}\) Moreover, given English courts’ formalistic approach to judicial decision-making and aversion to non-doctrinal considerations,\(^{576}\) it is highly unlikely that they would embark on such an inquiry. Therefore, lateral piercing is appropriately confined to asset stripping situations.

2. Voluntary Piercing

One further issue is whether the single economic unit theory should be available for voluntary piercing at all. Given the arguments against voluntary piercing generally, the theory should arguably not be available in voluntary piercing cases. *Beckett*, however, signaled the English courts’ continual willingness to entertain voluntary piercing claims.\(^{577}\) If the theory is to be made available to voluntary piercing, it should only be under exceptional circumstances. Few extenuating circumstances should be recognized to ensure that voluntary piercing under the single economic unit theory is kept to a minimum. After a finding of an extenuating circumstance, it is then up to the courts to weigh the equity and policy considerations of the situation to decide whether voluntary piercing should be allowed.

Such an open-ended approach seems best suited for voluntary piercing for a number of reasons. First, the burden-shifting framework proposed for involuntary piercing is clearly ill-suited for voluntary piercing. Issues of control, impropriety, and causation linked with plaintiff loss have little to do with whether voluntary piercing should be permitted.\(^{578}\) Second, while many involuntary piercing cases implicate the same underlying issue of limited liability, voluntary piercing cases feature such a wide variety of facts that it is probably impossible to

---

\(^{575}\) See Leebron, [*supra* note 234], at 1615–17 (remarking on the difficulty a court can encounter when attempting to determine the degree of economic integration of affiliated corporations).

\(^{576}\) See Adams, [*supra*] Ch. at 536–37.

\(^{577}\) See *Beckett*, [*supra*] I.C.R. at 1552.

\(^{578}\) See *Henn & Alexander*, [*supra* note 217], at 357–58.
come up with a general framework for deciding them. Third, every voluntary piercing case will require the court to confront the question of whether the corporate parent should be allowed to set aside its own subsidiary’s separate personality after enjoying the benefits of incorporation. As such, a totality of circumstances approach would likely provide the best solution. Although this will mean that there is little predictability in voluntary piercing cases, the nature of voluntary piercing is such that lack of legal certainty seems inevitable.

**Conclusion**

By undertaking a comparative study of the corporate veil doctrine in the United States and United Kingdom, this Article aims to provide a better understanding of the English doctrine. A comparison between the United Kingdom and the United States is particularly illuminating because although the corporation law of the two countries shares a common lineage, their approaches to the corporate veil doctrine and their modes of judicial reasoning show considerable divergences. It is hoped that some useful insights have been gleaned about judicial treatment of issues of policy and justice and about judicial decision making on both sides of the Atlantic.

As a standard-based exception to the general rule of limited liability, the corporate veil doctrine involves the fundamental choice facing every area of law and every legal system: the choice between legal certainty and the attainment of justice in particular cases. Limited liability promotes legal certainty and encourages investment in business activities. But as many courts have recognized, unyielding protection of it will lead to injustices on particular occasions. It is the responsibility of the corporate veil doctrine to ensure that injustices do not result. Given the United States’ emphasis on economic freedom and entrepreneur-

---

579 See, e.g., *Macaura*, [1925] A.C. at 619–21 (involving an insurance claim for assets owned by a corporation but insured under the name of the sole shareholder and creditor as an individual); *Beckett*, [2007] I.C.R. 1541–45 (involving voluntary piercing of veil between group members to give effect to employment contract that prohibited competition with parent company after employment, although the subsidiary provided the group’s services); *D.H.N.*, [1976] 1 W.L.R. at 857 (involving voluntary piercing of a corporate parent operating a separate company which was owned by the parent’s subsidiary, a third company, where a provision for compensation upon government purchase of company property was involved).

580 Cf. *Tate Access Floors Inc. v. Boswell*, [1991] Ch. 512 at 531 (Eng.) (discussing how those who benefit by utilizing the corporation as a medium to conduct affairs should not then be able to set aside corporate personality to their benefit when the corporation is under scrutiny).
ship, and the United Kingdom’s relatively more liberal reputation, it is perhaps surprising that U.S. courts have taken a more interventionist approach to veil piercing. What explains this apparent mismatch, however, is beyond the scope of this Article and will probably require a comprehensive study of the judiciaries in these two countries.

This Article highlights a number of crucial differences between the corporate veil doctrines in the United Kingdom and the United States, from jurisprudential approaches to treatment of specific cases. The English formalistic approach to judicial decision making and strict adherence to separate corporate personality means that there is little need to incorporate notions of justice and policy considerations when deciding corporate veil claims. This de-emphasis on justice and policy in turn obviates the need to fashion new analytical frameworks and allows the English courts to rely on common law concepts, many of which are ill-suited for the purpose. Part III argues that despite all the differences between the United Kingdom and United States highlighted in Part II, there are in fact elements in the English corporate veil cases that share substantial commonality with U.S. cases. Revealingly, one can in fact construct an instrumentality doctrine out of the English cases. Having demonstrated the surprising similarities between the two corporate veil doctrines, this Article goes one step further by proposing a revival of the single economic unit theory that will put the English corporate veil doctrine ahead of its U.S. counterpart. It will allow the English courts to craft a more uniform and systematic approach to corporate veil cases arising in the context of corporate groups.
Abstract: On December 25, 2009, Umar Farouk Abdulmutallab attempted to detonate an explosive device on a plane landing in Detroit. The attack was unsuccessful, but it spurred an important domestic debate regarding U.S. anti-terrorist programs and policies. In particular, the event fueled an argument over the proper forum for the interrogation and prosecution of terrorist suspects captured in the United States. Focusing on national security issues, some contended that treating Abdulmutallab as a criminal defendant in an Article III court, rather than subjecting him to a military commission, was imprudent and dangerous, while others insisted that it was entirely appropriate and responsible. This Note will probe this debate by comparing the two tribunals as each relates to the legal protections for suspects during interrogation. The Note argues that although some differences do exist, it is quite plausible that treating Abdulmutallab and other captured terrorist suspects as criminal defendants in Article III courts does not adversely impact intelligence gathering and national security.

Introduction

At approximately noon on Christmas Day in 2009, federal officials were notified about a passenger on Northwest Flight 253 who had attempted to detonate an explosive device while on board the aircraft. Just prior to landing in Detroit, Umar Farouk Abdulmutallab, the so-called “Christmas Day Bomber,” lit a device concealed in his clothing, igniting a fire on the plane. The device failed to fully explode, and passengers and crew restrained the Nigerian national until officials could board the plane and take him into custody.

* Ian Kennedy is an Articles Editor for the Boston College International & Comparative Law Review. The author would like to pay tribute to his father, the late Christopher Kennedy, for his many years of support and sacrifice.

3 Barrett, supra note 1.
The event understandably touched a nerve in the American public, as some observed that the series of intelligence failures which culminated in allowing Abdulmutallab to board the plane were frighteningly similar to the failures that led to 9/11. In addition to the intelligence failures, there was also a firestorm of political controversy over the treatment of Abdulmutallab in the time immediately after the failed terrorist attack. The argument generally divided along party lines, with Republicans contending that it was a mistake to bring criminal charges in an Article III court against Abdulmutallab rather than holding him as an unprivileged enemy belligerent before a military commission; Democrats, meanwhile, argued that criminal charges in civilian court were appropriate and customary.

This Note analyzes the primary premise underlying the Republican criticism; namely, that putting Abdulmutallab in an Article III court grants him rights during interrogation that undermine American national security. This Note concludes that, indeed, there are some differences between the two tribunals in terms of the right to counsel and rights during interrogation, but the differences are not as significant as Republicans suggest. In fact, in the case of Abdulmutallab, it is quite possible that putting him in an Article III court does not adversely impact intelligence gathering at all.

Part I of this Note provides information about the events surrounding the interrogation of Abdulmutallab and an explanation of the Republican criticism of the process. It also provides a brief history of and the current state of military commissions. Part II evaluates features of both Article III criminal courts and military commissions. In particular, it analyzes a suspect’s right to counsel and limitations on interrogation as conferred by domestic and international law. Finally, Part III of this Note compares the differences between the two tribunals and concludes that the differences are quite likely irrelevant to intelligence gathering in the case of Abdulmutallab and other similar situations where terrorists are captured in the United States.

---

I. Background

A. Miranda and Political Controversy

1. Miranda Warnings

Federal and local authorities apprehended Abdulmutallab when the plane landed. While in their custody, he made incriminating statements regarding his efforts to bomb the plane. Following initial treatment for his injuries, FBI and other federal agents questioned the terrorist suspect for fifty minutes without administering Miranda warnings, relying on an exception which allows custodial interrogation before Miranda warnings in cases where it is reasonable to believe that there is a threat to public safety. After the fifty minutes, investigators took a five-hour break from questioning because Abdulmutallab’s “deteriorating” medical condition required surgery. Later, when a second team of agents started to question him, approximately ten hours after being arrested, he refused to speak and investigators subsequently read him his Miranda rights. In the month that followed, after federal agents indicted Abdulmutallab in Article III criminal court and brought his family to the United States, reports indicate that he changed course and substantially cooperated with investigators at the urging of his relatives.

2. Political Controversy

In the aftermath of the failed attack, Republican politicians and others forcefully questioned the prudence of trying Abdulmutallab and terrorists like him in a civilian court. Sarah Palin led the charge from

7 Barrett, supra note 1.
8 Id.
9 Walter Pincus, Christmas Day Bomb Suspect Was Read Miranda Rights Nine Hours After Arrest, Wash. Post, Feb. 15, 2010, at A6. Miranda warnings were first established in the seminal case of Miranda v. Arizona, 384 U.S. 436 (1966). Briefly stated, Miranda requires that during custodial interrogation and before any questioning occurs, law enforcement must notify the suspect of his right to remain silent, right to an attorney, and inform him that any statements made could be used in his prosecution. Miranda, 384 U.S. at 444.
10 Pincus, supra note 9.
11 Id. Initial reports suggested that he stopped speaking after the Miranda warnings, but a subsequent account indicated he received Miranda warnings only after refusing to speak. See Barrett, supra note 1; Pincus, supra note 9.
her Facebook account, proclaiming that “it simply makes no sense to treat an al-Qaeda trained operative willing to die in the course of massacring hundreds of people as a common criminal.”

Although the criticism has not always been focused, the main tenet of the Republican argument is that interrogating the accused as a criminal defendant and providing a defense attorney impedes intelligence gathering and threatens national security. As an alternative, Republicans advocate for the use of military commissions, which were constituted to provide hearings for a number of the Guantánamo Bay detainees as a way for them to challenge their detention. In a letter to the Attorney General, high-ranking Senate Republicans blamed the administration for misguidedly treating Abdulmutallab as a criminal defendant, “rather than... an intelligence resource to be thoroughly interrogated in order to obtain potentially life-saving information.”

The letter charged that the administration sacrificed valuable opportunities because of a “preoccupation with reading the Christmas Day bomber his Miranda rights.” Senator Lindsey Graham separately leveled a slightly different criticism, focusing on the legal counsel argu-

Action=PressRoom.PressReleases&ContentRecord_id=94b699e5–802a-23ad-4641–1a34b485c7dd.

14 Perez, supra note 5.


16 See Charlie Savage & Scott Shane, Experts Urge Keeping Two Options for Terror Trials, N.Y. Times, Mar. 9, 2010, at A15; Hatch, supra note 15. In addressing the closure of Guantánamo Bay, President Obama highlighted another possible category for some of the alleged terrorists held at Guantánamo: continued detention without trial before a military commission. Press Release, Barack Obama, President, Remarks by the President on National Security (May 21, 2009) [hereinafter Presidential Remarks], available at http://www.whitehouse.gov/the_press_office/Remarks-by-the-President-On-National-Security-5–21–09. Such detainees would receive “periodic review” under “an appropriate legal regime... that is consistent with our values and our Constitution.” Id. President Obama explained that for certain detainees a trial before a military commission would not be prudent because, while the government has proof that these individuals are dangerous due to their terrorist connections, it could not prove these specific charges in a military commission considering the deficiencies in the cases such as tainted evidence. See id. An analysis of this second system of review is beyond the scope of this Note, as Republicans have generally called for trying the suspect before military commissions. See Savage & Shane, supra.


18 Id.
ment and asserting that there is a “reason why we have never given unlawful enemy combatants legal counsel at the time of capture . . . it was a mistake to read him Miranda rights after he was apprehended and to suggest otherwise is just political spin.”

Attorney General Holder responded to these various criticisms with a pointed letter of his own, underscoring that after 9/11, without exception, the practice has been to criminally charge any suspected terrorist captured in the United States. The Attorney General further disputed the notion that valuable information could not be acquired while charging a terrorist criminally; he argued that “history shows that the federal justice system is an extremely effective tool for gathering intelligence.”

B. Establishing Military Commissions

Since the 9/11 attacks, the U.S. government has grappled with the question of the proper forum in which to hold terrorists accountable for their actions. Proposals of what should be done have ranged from trying such individuals in traditional criminal courts to fashioning a brand new international tribunal. Ultimately, the Bush administration favored creating a unique military tribunal now known as military commissions.

The first iteration of these commissions was established by executive order. The order limited the jurisdiction of the commissions to non-citizen members or associates of Al-Qaeda and delegated the duty of promulgating rules and procedures to the Secretary of Defense. Notably, the order explicitly rejected any form of judicial review in a domestic court or international tribunal for verdicts rendered by the commissions.

The first military commission hearing commenced in 2004, approximately thirty months after President Bush’s order. That hearing and others did not go smoothly, and critics at home and abroad criti-

---

19 Graham, supra note 13.
20 See Holder, supra note 6, at 2.
21 Id. at 4.
23 Id. at 318.
24 See id.
27 Id.
28 See Glazier, supra note 25, at 161.
cized the ostensibly ad hoc procedures that suggested the U.S. government was not fully prepared to provide an adequate adjudicatory body. In fact, the United Kingdom successfully removed two of its own citizens from facing the commission, insisting that they must be given “fair trials.”

All of the military commissions were abruptly halted in 2006, when a 5–3 ruling by the Supreme Court invalidated the rules of the commission in *Hamdan v. Rumsfeld*, holding that their “structure and procedures violate both the [Uniform Code of Military Justice] and the Geneva Conventions.”

The Bush Administration and Congress responded quickly in 2006, passing the Military Commissions Act (MCA), which sought to adhere to the guidance of the Supreme Court in *Hamdan*, thus establishing a system consistent with international and U.S. legal standards. The military commissions began to take the shape of the military’s court-martial system. The MCA implemented many changes to the military commissions; included among these changes are that the MCA: 1) created an appellate review board of military judges for rulings made by the trial panel presiding over the commission; 2) permitted further judicial review of these rulings by the Court of Appeals for the District of Columbia and the Supreme Court of the United States; 3) rendered secret evidence inadmissible, only allowing the adjudicator to consider evidence that the accused also could hear; and 4) assured the accused the right to be present at trial.

Though distinct from the original commissions, the new commissions established by the MCA would nonetheless fail under the judicial scrutiny of the Supreme Court in 2008. In *Boumediene v. Bush*, the Court struck down § 7 of the statute, which sought to strip Guantánamo Bay detainees of their habeas corpus rights, thus recognizing that such detainees could challenge their ongoing detention with a proper hearing in an Article III court or a sufficiently similar forum. While the decision focused on detention and left the overall

---

29 See id.
30 Id. at 157.
33 *See* Glazier, *supra* note 25, at 176.
34 *See* id. at 176–77.
36 *See* id.
system of military commissions intact, questions persisted as to its viability and constitutionality. In 2009, Barack Obama assumed the presidency, promising changes to military commissions. On January 22, 2009, he signed an executive order suspending all commissions, while underscoring that they had only led to three convictions in seven years of existence. He called for Congress to amend the MCA, and improve upon the tribunal’s faults. Accordingly, Congress amended the Military Commissions Act (2009 MCA), which, among other important changes, bans the use of evidence obtained by “cruel, inhuman, or degrading treatment” and reverses the burden of proof with regards to the admissibility of hearsay evidence so that the accused no longer has to prove such evidence is unreliable. The jurisdiction of the 2009 MCA extends to “unprivileged enemy belligerents.” Essentially, unprivileged belligerents encompass any person who is a member of al-Qaeda or an individual that engages or materially assists in hostilities against the United States. In observing these changes, at least some scholars noted that military commissions evolved quite a bit since their first iteration and are substantially more fair and effective.

---

38 See Presidential Remarks, supra note 16.
40 Presidential Remarks, supra note 16.
41 Id.
44 Id. § 948a(7). The 2009 MCA also excludes “privileged enemy belligerents” from the category of “unprivileged enemy belligerents.” Id. §§ 948a(6), 948a(7), 948b(a). In order for a captured detainee to be considered a “privileged enemy belligerent” they must fall under one of the eight categories of protected individuals in Article IV of the Geneva Conventions Relative to the Treatment of Prisoners of War. Id. § 948a(6).
45 See, e.g., Silliman, supra note 42, at 291.
II. DISCUSSION

A. Right to Counsel

1. Right to Counsel in Military Commissions

Some, perhaps even all, unprivileged belligerents being tried before a military commission enjoy a constitutional right to counsel.\(^46\) In 2004, after the Fourth Circuit Court of Appeals ruled that Yaser Hamdi did not have a right to immediately meet with an attorney, the Supreme Court asserted that Hamdi, a U.S. citizen detainee captured on the battlefield of Afghanistan and held in the United States, “unquestionably has the right to access counsel” with respect to proceedings on remand.\(^47\)

In addition, the 2009 MCA provides that unprivileged belligerents “shall be represented” by defense counsel.\(^48\) The statute allows the detainee to select any military or civilian counsel for representation in the commission; however, the civilian counsel, among other requirements, must be a U.S. citizen and obtain permission to handle sensitive national security information.\(^49\) Moreover, if civilian counsel is selected, then military counsel, also known as Judge Advocates Generals (JAG), must assist as part of the defense team.\(^50\)

With the integral role that JAGs play in the defense of the accused, obvious questions arise about their independence because they are lawyers employed by the military.\(^51\) Such questions are merely speculative, however, because JAGs have been zealous advocates and “stubborn rule of law defenders” in the war on terrorism, even to the point that pro-

\(^{46}\) See Hamdi v. Rumsfeld, 542 U.S. 507, 539 (2004). The reach of due process rights and the right to counsel for unprivileged belligerents mandated by Hamdi is still an unsettled legal issue. David Cole, Out of the Shadows: Preventive Detention, Suspected Terrorists, and War, 97 CAL. L. REV. 693, 743 n.227 (2009). Particularly complicated in the case of Abdulmutallab is that he had not established significant ties to the United States before being apprehended within U.S. borders. See Kwong Hai Chew v. Colding, 344 U.S. 590, 596 n.5 (1953) (noting that “once an alien lawfully enters and resides in this country he becomes invested with the rights guaranteed by the Constitution to all people within our borders. Such rights include those protected by the First and the Fifth Amendments and by the due process clause of the Fourteenth Amendment”).

\(^{47}\) Hamdi, 542 U.S. at 539. The Court avoided the question of whether Hamdi was immediately entitled to see an attorney after he was detained, only confirming that he should continue to have access going forward. See id.

\(^{48}\) 10 U.S.C. § 948c(b) (Supp. III 2009).

\(^{49}\) See id.

\(^{50}\) Id.

ponents of the unitary executive theory in the Bush Administration viewed JAGs’ penchant for independence as a major challenge to increasing Department of Defense control of the military.\textsuperscript{52}

As a practical matter, even though the Constitution and the statutory rules affirm a detainee’s right to counsel, some advocates have complained about the significant logistical impediments to meeting with individuals located at Guantánamo Bay.\textsuperscript{53} Complaints include difficulty traveling to the island, two-week notice requirements before a meeting, and a limited number of scheduling slots due to a lack of facilities.\textsuperscript{54} The continued relevancy of these concerns is called into question, however, by a recent review which found that in 2008, the detention facility hosted approximately 1800 attorney visits on behalf of approximately 200 detainees being held there.\textsuperscript{55} Moreover, in the case of future unprivileged belligerents who are not detained at Guantánamo Bay, the logistical concerns related to traveling will presumably not be as cumbersome, where the detention occurs off of the isolated island.\textsuperscript{56}

2. Right to Counsel for a Civilian Criminal

In relevant part, the Sixth Amendment of the U.S. Constitution guarantees that “[i]n all criminal prosecutions, the accused shall enjoy the right . . . to have the Assistance of Counsel.”\textsuperscript{57} The Supreme Court interprets this right to not only require access to counsel, but to bar the government from eliciting statements from the accused in the absence of an attorney once criminal proceedings commence.\textsuperscript{58} The right to counsel under the Sixth Amendment does not attach, however, until

\begin{itemize}
  \item \textsuperscript{52} Id. at 2000, 2001–02.
  \item \textsuperscript{53} E.g., David Frakt, \textit{The Difficulty of Defending Detainees}, 48 Washburn L. J. 381, 393–94 (2009).
  \item \textsuperscript{54} Id.
  \item \textsuperscript{55} See Presidential Remarks, \textit{supra} note 16 (mentioning the approximate number of detainees in Guantánamo); Rev. of Dep’t Compliance with President’s Executive Order on Detainee Conditions of Confinement, U.S. Department of Defense 66 (2010) [hereinafter Detainee Report], http://www.defense.gov/pubs/pdfs/REVIEW_OF_DEPARTMENT_COMPLIANCE_WITH_PRESIDENTS_EXECUTIVE_ORDER_ON_DETAINEE_CONDITIONS_OF_CONFINEMENTa.pdf (stating that there were 1800 attorney-detainee visits in 2008).
  \item \textsuperscript{56} See Frakt, \textit{supra} note 53, at 393–94 (describing some of the logistical difficulties of traveling to Guantánamo); Charlie Savage, \textit{Closing Guantánamo Fades as a Priority}, N.Y. Times, June 26, 2010, at A13 (reporting that the Guantánamo prison is set to be closed but progress has stalled along with President Obama’s proposal for alternative detainment camp locations).
  \item \textsuperscript{57} U.S. Const. amend. VI.
  \item \textsuperscript{58} Massiah v. United States, 377 U.S. 201, 206 (1964).
\end{itemize}
the government initiates adversarial judicial proceedings, “whether by way of formal charge, preliminary hearing, indictment, information, or arraignment.”

Even though there are no specific Sixth Amendment obligations, a defendant is still protected before adversarial judicial proceedings commence because the Fifth Amendment demands the government provide the accused certain protections known as *Miranda* warnings. Generally, under *Miranda*, once a defendant is taken into custody and before any interrogation can take place, law enforcement must advise the individual of their right to remain silent and their right to retain counsel. If the accused asserts his right to counsel, all questioning must cease until an attorney is present.

There are also significant exceptions to the *Miranda* doctrine. Pertinent to this Note is the public safety exception, where the mandate for law enforcement to issue *Miranda* warnings before interrogation is suspended if there is a substantial threat to public safety. The public safety exception provides that law enforcement may temporarily elect not to administer the warnings if the “need for answers to questions in a situation posing a threat to public safety outweighs the need for the prophylactic rule protecting the Fifth Amendment’s privilege against self-incrimination.” In evaluating the exigency of each situation in which the exception is invoked, the judicial branch will defer to the decisions of the officials on the ground.

**B. Interrogation Limits**

1. Limitations on Interrogations in Military Commissions

In 2006, the Supreme Court of the United States issued a decision in *Hamdan v. Rumsfeld* that struck down the Bush Administration’s procedures for military commissions because, among other requirements, the commissions failed to comport with Article III of the 1949 Geneva Conventions (known as “Common Article III”). The majority seized

---

61 Id. at 444.
62 Id. at 444–45.
64 Id. at 657.
65 Id.
66 See id. at 659.
on the language in Common Article III that affords combatants “all the judicial guarantees which are recognized as indispensable by civilized peoples,” and then reasoned that the MCA of 2006 did not meet that standard.68

The Hamdan ruling set the stage for President Obama’s executive order, issued soon after taking office in 2009, in which he pronounced Common Article III as a minimum standard for all interrogations of individuals in U.S. custody.69 The order further stipulated that interrogations would be consistent with the nation’s other international obligations, such as the Convention Against Torture (CAT), as well as domestic laws like the Detainee Treatment Act (DTA).70 President Obama specifically identified Army Field Manual 2–22.3 as the appropriate guide for interrogators in the field, as it was deemed to be in compliance with all of the aforementioned legal obligations.71

a. Common Article III Limitations

In relevant part, Article III of the 1949 Geneva Conventions requires parties to refrain from “violence to life and person” including “torture” when dealing with detainees, and it prohibits “outrages upon personal dignity, in particular, humiliating and degrading treatment.”72 Early in his presidency, President Obama ordered a Department of Defense commission to undertake the question of the specific application of Common Article III to detainee treatment.73 As part of evaluating the treatment of individuals at Guantánamo, the specially-convened commission elaborated on the more specific obligations that flowed from Common Article III’s broad principles.74

Specifically, the commission noted that in accordance with Common Article III, any form of sensory deprivation is prohibited.75 They further explained that the broad mandate prohibits humiliation including rape, sexual assault, and subjecting detainees to “public curiosity”; moreover, it barred violence, threats of violence, or any physical force upon detainees except when in self-defense or other narrowly defined

---

70 Id.
71 See id.
72 Geneva, supra note 68, art. 3.
74 See id.
75 See id. at 29.
emergency situations. As a general rule, the commission announced that if all interrogators adhered to the instructions of Army Field Manual 2–22.3, the methods used would be in concert with Common Article III.77

b. Army Field Manual 2–22.3 Limitations

Army Field Manual 2–22.3, written and published in 2006, established a minimum standard for the treatment of any detainee in military custody and it identifies interrogation techniques consistent with international law.78 The Manual very deliberately lays out the many forms of permissible interrogation methods in extended detail,79 and its general overview of forbidden techniques tracks the language of Common Article III by prohibiting “cruel, inhuman or degrading treatment.”80 Specifically, among other more extreme techniques, it bans: forced nudity or sexually suggestive poses; placing tape or a hood over the eyes of a detainee; using military dogs in interrogation; and the deprivation of food, water, or medical care.81

The overall policy of the Manual is further underscored by its proposed method of evaluation for interrogators that question whether a particular technique crosses the line into a banned procedure.82 It instructs soldiers to ask, “if the . . . technique were used by the enemy against one of your fellow soldiers, would you believe the soldier had been abused?”83 Consequently, the Army Field Manual not only prohibits certain techniques in accordance with international law, but it also drives home the broader notion that enemies should only be subjected to tactics which soldiers would be comfortable having their comrades experience.84

---

76 See id. at 40, 49.
77 See id. at 61.
80 Id. at 5–74.
81 Id. at 5–75.
82 See id. at 5–76.
83 Id.
84 See id.
c. Convention Against Torture and Detainee Treatment Act Limitations

In addition to Common Article III and the Army Field Manual, the CAT and the DTA were also identified by President Obama in Executive Order 13491 as establishing the lawful parameters of detainee interrogation. While the CAT is an international agreement and the DTA is domestic law, both documents ban torture as well as any “cruel, inhuman or degrading treatment or punishment.” What is more, in both the CAT and DTA, the United States interprets “cruel, inhuman or degrading treatment or punishment” according to the Fifth, Eighth, and Fourteenth Amendments of the U.S. Constitution. That is, the interrogation or punishment that would contravene those constitutional limitations would also violate the government’s international obligations under the CAT and its domestic law obligations under the DTA. In this way, two of the laws which President Obama highlights as constraints on interrogation methods are coextensive with the legal protections of the U.S. Constitution on interrogation.

2. Limitations on Interrogation for a Civilian Criminal

Just as the decision by the Supreme Court in *Miranda v. Arizona* expanded the opportunity of a criminal defendant to access counsel, the case similarly impacted a defendant’s protections against law enforcement interrogation. Once in custody, under *Miranda* and its progeny, a criminal defendant invoking his right to remain silent cuts off all questioning unless he freely chooses to subsequently waive that right and speak with authorities.

In addition to the accused’s right to end questioning as articulated in *Miranda*, the Due Process Clause also limits the conduct of law enforcement officials during the interrogation process. Due process demands that any confession be voluntary, while taking into account

---

88 See 42 U.S.C. §§ 2000dd(a), 2000dd(d); Scharas, supra note 87, at 403.
89 See 42 U.S.C. §§ 2000dd(a), 2000dd(d); Scharas, supra note 87, at 403.
90 See *Miranda*, 384 U.S. at 444.
characteristics of the suspect as well as the degree of misconduct and coercion by law enforcement. The case law holds that physical force during an interrogation makes any confessions or evidence obtained from those tactics involuntary and thus inadmissible at trial. Similarly, threats of physical violence also are considered involuntary and violate the Due Process Clause.

In establishing the contours of due process outside of the categorical prohibitions against violence and threats of violence, an inquiry into what is permissible by law enforcement is very fact sensitive. For instance, the Supreme Court of Minnesota ruled that police threatening to charge a defendant with more crimes if he lied is involuntary, but the Second Circuit Court of Appeals upheld a confession when law enforcement referenced the electric chair as punishment and blatantly misrepresented evidence. As another example, where police threatened to cut off state aid to a suspect’s children if she did not confess, the Supreme Court found such actions “impellingly coercive” and thus involuntary, but the Eighth Circuit Court of Appeals ruled that police bursting into a suspect’s room at 6:30 in the morning, holding him naked, and asserting that he would go to jail for life if he did not cooperate did not render his subsequent confession involuntary.

In another line of cases and constitutional restraints, the Supreme Court held that the substantive due process rights conferred by the Fifth and Fourteenth Amendments bars executive abuse of power that “shocks the conscience.” This inquiry is also fact sensitive as “[d]eliberate indifference that shocks in one environment may not be so patently egregious in another.” As a constitutional matter, considering this standard as well as the aforementioned voluntariness requirement, courts have found the following conduct violates the Constitution: hand-cuffing a detainee to a post while standing up for a substantial period of time and beyond the time needed to ensure order; keeping the temperature at unreasonable levels in detention facilities; and questioning that lasts an entire day over the course of multiple

93 See Bloom & Brodin, supra note 91, at 232.
96 See Bloom & Brodin, supra note 91, at 232.
97 See Green v. Scully, 850 F.2d 894, 903, 904 (2d Cir. 1988); State v. Garner, 294 N.W.2d 725, 727 (Minn. 1980).
101 Id. at 850.
days.\textsuperscript{102} In contrast, courts have not struck down practices such as extended solitary confinement or constant lighting during detention so long as authorities are doing so in order to promote order or safety.\textsuperscript{103}

\section*{III. Analysis}

The provisions for access to counsel and the legal limitations on interrogations for detainees tried before military commissions are substantially similar to those involving an Article III criminal suspect.\textsuperscript{104} The principal ground for any differentiation is that, unlike Article III suspects, military commission detainees are not entitled to \textit{Miranda} warnings.\textsuperscript{105} Nevertheless, any adverse impact on national security caused by administering or not administering \textit{Miranda} warnings is tenably refuted by scholarship which concludes that \textit{Miranda} does not negatively affect law enforcement efforts in any empirically demonstrable way.\textsuperscript{106} Moreover, as a matter of anecdotal evidence, the case of Abdulmutallab does not undermine this argument.\textsuperscript{107}

\subsection*{A. Right to Counsel Comparison}

In certain respects, the right to counsel under military commissions is commensurate with the right to counsel enjoyed by criminal defen-

\textsuperscript{102} See \textit{Leyta v. Denno}, 347 U.S. 556, 561 (1954); \textit{Michael John Garcia, Interroga-


\textsuperscript{103} \textit{Garcia}, supra note 102, at 4–5.

\textsuperscript{104} See 10 U.S.C. § 949c(b) (Supp. III 2009) (stating that unprivileged belligerents shall be represented by counsel before a military commission); 42 U.S.C. §§ 2000dd(a), 2000dd(d) (2006) (asserting that no individual in U.S. custody shall be subject to cruel, inhuman, or degrading treatment as defined by the Fifth and Fourteenth Amendments of the U.S. Constitution); \textit{Massiah v. United States}, 377 U.S. 201, 206 (1964) (holding that a criminal defendant must be represented by counsel after indictment); \textit{Bloom & Brodin, supra} note 91, at 231–32 (discussing the Due Process Clause limitations on interrogations for Article III criminal suspects).


\textsuperscript{107} See \textit{Pincus, supra} note 9 (noting that Abdulmutallab stopped speaking with authorities before being read his \textit{Miranda} rights).
The 2009 MCA states the accused “shall be represented . . . before a military commission,” and the Constitution similarly requires that the right of the accused to counsel attaches at the commencement of adversarial judicial proceedings. What is more, there is no evidence to suggest the caliber and zeal of counsel representing an unprivileged belligerent—whether it is JAG or a team consisting of civilian attorneys and JAG—would necessarily be less than that of an attorney representing an Article III criminal. Despite some past problems with attorney access at Guantánamo, a recent Department of Defense report concludes the attorneys of the most recent detainees who were charged under the MCA have “access to their detainee clients and the means to seek redress in the event they believe their access is unreasonably curtailed.”

In the interests of not being too sanguine, it should be underscored that there are a range of challenges still faced by attorneys representing military commission detainees, but as a comparative matter, many defense lawyers in Article III courts also operate in an imperfect criminal justice system with glaring inequities and difficulties.

Despite the aforementioned similarities, a critical difference between military commissions and Article III courts is that no constitutional or legal authority asserts that unprivileged belligerents are entitled to the warnings and protections of Miranda. At the beginning of any custodial interrogation, Miranda requires, among other things, that a suspect be explicitly informed of his right to counsel and that questioning cannot proceed in the absence of an attorney without an affirmative waiver of the right. Thus, although unprivileged belligerents possess the right to an attorney for the military commission proceedings once they are charged, there is no legal obligation that they be offered access to counsel. Such an attitude toward non-citizens, however, is unprecedented in American criminal law. The two cases that are closest to the MCA—those involving Gideon v. Wainwright, 372 U.S. 335 (1963), and Escobedo v. Illinois, 378 U.S. 478 (1964)—are dramatically different.

---

108 See 10 U.S.C. § 949c(b); Massiah, 377 U.S. at 206.
109 See 10 U.S.C. § 949c(b); Massiah, 377 U.S. at 206.
109 Luban, supra note 51, at 2000 (suggesting that JAGs may even be more effective rule of law defenders than civilian attorneys).
111 See Matthew Ivey, Challenges Presented to Military Lawyers Representing Detainees in the War on Terrorism, 66 N.Y.U. ANN. SURV. AM. L. 211 passim (2010) (surveying the many challenges faced by military lawyers representing military commission detainees).
112 See Mary Sue Backus & Paul Marcus, The Right to Counsel in Criminal Cases, a National Crisis, 57 HASTINGS L.J. 1031 passim (2006) (highlighting the myriad difficulties for indigent defendants in the criminal justice system).
113 Compare 10 U.S.C. §§ 948, 949 (containing no reference to Miranda rights or any analogous protections in military commissions), and Sievert, supra note 105, at 92 (asserting that detainees do not enjoy Miranda rights in military commissions), with Miranda, 384 U.S. at 444 (holding that criminal defendants are entitled to Miranda warnings).
114 Miranda, 384 U.S. at 444.
access to an attorney before that point.\textsuperscript{116} Moreover, a very important legal question remains in situations where an unprivileged belligerent requests counsel at the beginning of custodial interrogation. This precise question has not yet been resolved by the courts.\textsuperscript{117}

Another important difference between military commissions and Article III proceedings is that, once formal charges are filed in an Article III court, the attorney becomes the intermediary between the state and the accused, and the Sixth Amendment bars the government from eliciting statements outside the presence of counsel.\textsuperscript{118} No such constitutional right exists with military commissions, and the 2009 MCA does not prohibit custodial interrogation of the suspect once the military commission commences.\textsuperscript{119}

\textbf{B. Comparison of Limitations on Interrogation Techniques}

Myriad sources of law are now interpreted to apply to terrorists held as unprivileged belligerents.\textsuperscript{120} In fact, the standards of treatment in some of those sources explicitly peg the protections of detainees to what is permitted by the U.S. Constitution under the Fifth, Eighth, and Fourteenth Amendments.\textsuperscript{121} Both the DTA and CAT use the Constitution to mark the outward bounds of permissible techniques.\textsuperscript{122}

Still, the fact that the Constitution is invoked as a marker of sensible treatment of unprivileged belligerents does not mean that the permissible interrogation techniques of an alleged purse-snatcher will be the same as that for a terrorist suspect.\textsuperscript{123} Justice Jackson famously remarked that the Constitution is not a “suicide pact,”\textsuperscript{124} and Supreme

\textsuperscript{116} See 10 U.S.C. § 949c(b); Sievert, supra note 105, at 92.

\textsuperscript{117} See Hamdi v. Rumsfeld, 542 U.S. 507, 539 (2004) (observing that Hamdi “unquestionably has the right to access counsel” regarding proceedings on remand, but not addressing whether he had the right to counsel immediately after he was detained); Cole, supra note 46, at 742 n.227 (noting that the extent of the right to counsel is still being debated by advocates).

\textsuperscript{118} Massiah, 377 U.S. at 206.


\textsuperscript{121} See 42 U.S.C. §§ 2000dd(a), 2000dd(d); Scharas, supra note 87 (indicating that the obligations of CAT are coextensive with the Fifth and Fourteenth Amendments).

\textsuperscript{122} See 42 U.S.C. §§ 2000dd(a), 2000dd(d); Scharas, supra note 87.

\textsuperscript{123} See County of Sacramento v. Lewis, 523 U.S. 833, 850 (1998).

\textsuperscript{124} Terminello v. City of Chicago, 337 U.S. 1, 37 (1948) (Jackson, J. dissenting).
Court jurisprudence quite reasonably holds that due process is not subject to a “mechanical application” and law enforcement conduct that “shocks in one environment may not be so patently egregious in another.”

To this end, commentators note that permissible techniques for those accused of more pedestrian crimes may be different than those permitted for use on suspected terrorists; however, this would be the case irrespective of whether the accused is held as an unprivileged belligerent or an Article III criminal.

The notion that the protection for alleged terrorists during interrogation is equivalent to the constitutional safeguards for accused criminals in the Article III context is further bolstered by recent interpretations of Common Article III in the Geneva Conventions. In a 2009 executive order, President Obama identified Common Article III as a “minimum baseline” for treatment of all detainees under authority of the United States.

Common Article III explicitly bans physical violence, and the Department of Defense interprets the provision to also ban threats of physical violence. Similarly, the constitutional requirement of due process also prohibits violence and threats of violence in the interrogation process.

Conversely, while international and domestic legal limitations on the interrogation of detainees are equivalent to the protection of Article III criminal suspects under the Constitution, unprivileged belligerents do not enjoy the prophylactic constitutional protection of Miranda. Just like the safeguards of the right to counsel, law enforcement is not required to advise belligerents of their right to remain silent as they would with an Article III criminal suspect. Of course, the significance of this difference is somewhat mitigated by the current practice at Guantánamo Bay, which makes all of its interrogations voluntary and reports that one-third of its interrogations are initiated by detainees.

---

125 Lewis, 523 U.S. at 850.
126 Garcia, supra note 102.
127 See id.
128 See Geneva, supra note 68, art. 3; Detainee Report, supra note 55, at 40.
130 See Geneva, supra note 68, art. 5; Detainee Report, supra note 55, at 40.
132 Sievert, supra note 105, at 92.
133 See Miranda, 384 U.S. at 444; Sievert, supra note 105, at 92.
134 See Detainee Report, supra note 55, at 61.
C. Are the Republican Criticisms Merited?

Republicans are right to assert that the accused in an Article III criminal court enjoys different rights with respect to counsel and interrogation as compared to a detainee prosecuted in a military commission.\(^{135}\) Although unprivileged belligerents and criminal suspects are entitled to the same protection under the Due Process Clause of the Fifth and Fourteenth Amendments as it relates to physical and mental abuse, belligerents do not enjoy the prophylactic constitutional protection of \textit{Miranda}.\(^{136}\) Additionally, although there is a right to counsel in military commissions as well as Article III criminal courts,\(^{137}\) there is no constitutional mandate requiring belligerents to be advised of this right and no Sixth Amendment obligation that limits authorities from speaking with the belligerent unless his lawyer is present, once adjudicatory proceedings commence.\(^{138}\) Importantly, then, the question becomes whether these differences actually undermine intelligence gathering and national security. And, more specifically, whether the treatment of Abdulmutallab as an Article III criminal defendant was sensible in terms of national security, as the Obama Administration insists, or if this argument is “ridiculous” and the “whole process of criminalizing the war is misguided,” as one leading Republican asserted.\(^{139}\)

1. Practical Effects of the Differences in Procedural Rights

Undermining the Republican argument and perhaps contrary to popular perceptions, it is likely that the issuing of \textit{Miranda} warnings to terrorist suspects does not make a significant difference as far as intelligence gathering goes.\(^{140}\) Among law enforcement officials, the “pervasively shared” view is that “\textit{Miranda} safeguards do not pose any serious

\(^{135}\) \textit{Compare} 10 U.S.C. §§ 948, 949 (containing no reference to \textit{Miranda} rights or any analogous protections in military commissions), \textit{and Sievert, supra note 105, at 92 (asserting that detainees do not enjoy \textit{Miranda} rights in military commissions)}), \textit{with Miranda}, 384 U.S. at 444 (holding that criminal defendants are entitled to \textit{Miranda} warnings).

\(^{136}\) \textit{Compare} 10 U.S.C. §§ 948, 949, \textit{and Sievert, supra note 105, at 92, with Miranda, 384 U.S. at 444. See Rogers v. Richmond, 365 U.S. 534, 541 (1936) (holding that due process limits interrogation methods by law enforcement); Scharas, supra note 87, at 403 (indicating that the obligations of CAT are coextensive with the Fifth and Fourteenth Amendments).

\(^{137}\) \textit{See} 10 U.S.C. § 949c(b); \textit{Massiah}, 377 U.S. at 206.

\(^{138}\) \textit{See} 10 U.S.C. §§ 948, 949 (containing no reference to \textit{Miranda} rights or any analogous protections); \textit{Massiah}, 377 U.S. at 206.

\(^{139}\) \textit{See} Holder, \textit{supra note 6}; Graham, \textit{supra note 13}.

\(^{140}\) \textit{See} Schulhofer, \textit{supra note 106, at 506}.
impediment to effective law enforcement." Generally, academics concur with this assessment. In attempting to quantify the effects of *Miranda*, one important study maintained that even accepting an opponent’s questionable methodology, the adverse impacts of *Miranda* could only be estimated to have lost convictions 0.78% of the time. Moreover, the study concludes that, “[f]or practical purposes, *Miranda*’s empirically demonstrable harm to law enforcement is essentially nil.”

A likely rebuttal to this argument is that the metric behind the *Miranda* data focuses on convictions, whereas the argument of Republicans is not that the United States will experience a decline in convictions, but that issuing *Miranda* rights decreases the likelihood of obtaining important intelligence. This line of argument is suspect, however, because the explanation for why so little convictions are lost is that the same information can be obtained using alternative interrogation techniques while nonetheless employing the prophylactic procedure of *Miranda*. That is, experience shows that *Miranda* warnings do not protect against investigators finding the truth, but rather, the warnings protect criminals from investigators’ use of coercive methods of getting to the truth. If one accepts this notion, then it becomes difficult to make the argument that *Miranda* is somehow unduly restrictive of intelligence gathering and should never be issued to a terrorist suspect who could possibly possess national security information. Indeed, scholars and law enforcement generally believe that *Miranda*’s adverse impact on law enforcement is minimal.

Even if one does not accept this argument, two more related but distinct points still undercut the assertion that putting alleged terrorists

---

141 Id. at 501.
143 Schulhofer, supra note 106, at 506.
144 Id.; see also George C. Thomas III, Plain Talk About the *Miranda* Empirical Debate: A “Steady-State” Theory of Confessions, 43 UCLA L. REV. 933, 935 (1996) (concluding that according to available data, *Miranda* “has had no effect on the overall confession rate, using ‘confession’ to include all incriminating statements”). One of the leading studies pointed to by opponents of *Miranda* concludes that *Miranda* has led to 3.8% lost convictions. Cassell, supra note 142, at 438. However, the methodology leading to this figure has been vigorously disputed. Schulhofer, supra 106, at 505–06.
145 See Schulhofer, supra note 106, at 506 (analyzing the social impacts of *Miranda* in terms of lost convictions).
146 See id. at 561, 562 (the fundamental purpose of *Miranda* is “not to eliminate confessions, but to eliminate compelling pressure in the interrogation process”).
147 See id.
148 See id.
149 Id. at 501; Cassell, supra note 142.
in Article III criminal courts, and thus following Miranda and its progeny, impedes intelligence gathering. First, only one out of every five criminal defendants will actually exercise their Miranda rights. Moreover, those with felony records are four times as likely to invoke their rights as those without because of their past experience with the criminal justice system. Thus, even if one believed Miranda significantly impacted an interrogator’s ability to collect national security information, presumably most terrorist suspects would not have previous experience with the U.S. criminal justice system, and thus would be less likely to invoke their Miranda rights.

Second, and more importantly, Miranda jurisprudence factors in exigencies such as critical national security matters, and makes exceptions in instances where the “need for answers to questions in a situation posing a threat to public safety outweighs the need for the prophylactic rule protecting the Fifth Amendment’s privilege against self-incrimination.” Therefore, even if one accepts the argument that Miranda warnings restrict intelligence gathering, the public safety exception of Miranda mitigates the potential effect of this point because it allows law enforcement to not issue the warnings when there is an immediate need for answers. The fact that Miranda entails a balancing test which weighs the potential threat suggests that courts and law enforcement would apply the exception even more broadly in the context of terrorist interrogations. Attorney General Eric Holder has acknowledged the value of this exception to Miranda for law enforcement in the case of terrorist plots, and he recently asked Congress for more clarity and guidance in applying the decades-old doctrine in the modern age.

---

151 See Leo, supra note 150, at 654.
152 Id. at 654–55.
153 See id.
154 See Quarles, 467 U.S. at 657.
155 See id.
156 See id.; see also John M. Allen, Note, Expanding Law Enforcement Discretion: How the Supreme Court’s Post-September 11th Decisions Reflect Necessary Prudence, 41 Suffolk U. L. Rev. 587, 595 (2008) (positing that the public safety exception of Quarles is “necessary to [a] successful antiterrorism investigation”).
157 See The Justice Department: Hearing of the H. Comm. on the Judiciary, 111th Cong. (2010) (response to questioning by Eric Holder, Att’y Gen. of the United States). In the hearing, Attorney General Holder explained his prior remarks to major press outlets about the Department of Justice seeking guidance with respect to applying the public safety exception of Miranda: “[W]e think that with regard to that small sliver—only terrorism-related matters, not in any other way just terrorism cases—that modernizing, clarify-
Moving beyond *Miranda*, another difference between the interrogations and right to counsel of an unprivileged belligerent in a military commission versus the accused in an Article III criminal court is that a belligerent possesses no Sixth Amendment right barring questioning outside the presence of counsel once judicial proceedings commence. If an alleged terrorist were tried in an Article III court, the required presence of an attorney during interrogations could render interrogations less effective, or simply create a logistical impediment. Thus, it is certainly plausible that this difference could hamper intelligence gathering, as the Sixth Amendment attaches as early as the prosecutor filing an indictment or arraigning the accused.

Nevertheless, the Sixth Amendment right to have counsel present during interrogations may be waived by a defendant at any time during the judicial proceedings. Moreover, once the military commission commences, although there is no right to have counsel present during interrogations, the unprivileged belligerent still has access to at least one attorney. Thus, whatever supposed interference results from an attorney’s presence during interrogations in an Article III proceeding could nonetheless occur during pre-interrogation detainee-counsel meetings; the belligerent can even condition his speaking with authorities on his attorney being present. In this way, although there is a difference between Article III courts and military commissions in terms of legal rights, these differences become much less apparent when one examines their practical effects.

---

158 Compare 10 U.S.C. §§ 948, 949 (containing no reference to *Miranda* rights or any analogous protections), and *Padilla*, 233 F. Supp. 2d at 600 (finding no Sixth Amendment rights for detained enemy combatant), with *Massiah*, 377 U.S. at 206 (holding that a criminal defendant enjoys a constitutional right to counsel after he is indicted).


160 See *id*.


162 See 10 U.S.C. § 949c(b).

163 See *id*.

164 See *id*. 
2. Application to Abdulmutallab

Applying the aforementioned analysis to the case of Abdulmutallab underscores why it is quite plausible that treating the Christmas Day Bomber as an Article III criminal suspect did not make a significant difference in terms of national security intelligence. Abdulmutallab would have received a lawyer regardless of whether he was charged as an Article III criminal or an unprivileged belligerent. In addition, the Fifth and Fourteenth Amendments would have limited the interrogation practices of authorities irrespective of his status.

Of course, if Abdulmutallab were held as an unprivileged belligerent, he would not have been read Miranda warnings, and early reports indicate that Abdulmutallab stopped talking to authorities after receiving Miranda warnings. Later reports, however, asserted that authorities did not issue Miranda warnings until after he refused to speak, and that he was legally questioned by authorities under Miranda’s public safety exception in the time before the warnings were issued. Indeed, the event which seemed to trigger his reluctance to speak was not the Miranda warnings, but the fact that he spent about four hours in surgery improving what was a “deteriorating” condition over the course of the initial interrogations.

Whatever the exact course of events, it is undisputed that Abdulmutallab resumed answering questions about a month later, in January 2010. The Federal Bureau of Investigation enlisted the help of his family, and reports indicate that Abdulmutallab subsequently cooperated for a number of days and provided significant intelligence to authorities.

Thus, the case of Abdulmutallab illustrates precisely why treating him as a criminal defendant did not significantly undermine intelligence gathering efforts. The Miranda warnings issued to him do not

---

165 See id.; Massiah, 377 U.S. at 205; Rogers, 365 U.S. at 541; Scharas, supra note 87 (indicating that the obligations of CAT are coextensive with the Fifth and Fourteenth Amendments).
166 See 10 U.S.C. § 949c(b); Massiah, 377 U.S. at 205.
167 See Rogers, 365 U.S. at 541; Scharas, supra note 87 (indicating that the obligations of CAT are coextensive with the Fifth and Fourteenth Amendments).
168 See Sievert, supra note 102, at 92; Barrett, supra note 1.
169 Pincus, supra note 9.
170 Holder, supra note 6, at 4.
171 Pincus, supra note 9.
172 See Zeleny & Savage, supra note 12.
173 Id.
174 See id.
appear to have altered his decision to stop talking to authorities, and even if they had, shrewd and resourceful interrogation tactics managed to get him to speak again. 175 Moreover, the apparent success with cajoling him to speak a second time suggests that the attached right to counsel was not a major impediment to interrogation efforts. 176

Abdulmutallab’s case is not an aberration. 177 There are many cases in recent history where law enforcement treated suspected terrorists with the same procedures as Article III criminal defendants and got similar results. 178 Richard Reid, the infamous shoe bomber, received Miranda warnings multiple times over a two-day period without exercising his rights or discontinuing cooperation as a result. 179 Similarly, alleged terrorists L’Houssaine Kherchtou and Nuradin Abdi both provided important counter-terrorism intelligence after receiving Miranda warnings. 180 In this way, practical experience does not suggest that putting a terrorist suspect in an Article III court necessarily undermines intelligence collection. 181

D. Are Military Commissions Useless?

In spite of the similarities highlighted in the previous discussion between the rights of detainees in military commissions to those of an Article III criminal suspect, military commissions are not useless. 182 Indeed, there are other concerns which support the use of military commissions. 183 For one, military commissions are “portable” and can be held on any U.S. military base throughout the world; thus, municipal

175 See Pincus, supra note 9; Zeleny & Savage, supra note 12. This course of events fits one of the central arguments made by those who argue the social costs of Miranda are low. See, e.g., Schulhofer, supra note 106, at 561–62. The reason that Miranda does not sacrifice confessions is because law enforcement can get a confession even when a person is aware of his constitutional rights. See id.

176 See Zeleny & Savage, supra note 12.


178 See, e.g., id.; Holder, supra note 6, at 4.

179 Allen, supra note 177.

180 Holder, supra note 6, at 4.

181 See id.

182 See 10 U.S.C. § 949c(b) (stating that unprivileged belligerents shall be represented by counsel before a military commission); 42 U.S.C. §§ 2000dd(a), 2000dd(d) (asserting that no individual in U.S. custody shall be subject to cruel, inhuman, or degrading treatment as defined by the Fifth and Fourteenth Amendments of the U.S. Constitution); Messiah, 377 U.S. at 206 (holding that a criminal defendant must be represented by counsel after indictment); Bloom & Brodin, supra note 91, at 231–32 (discussing the Due Process Clause limitations on interrogations for Article III criminal suspects).

183 See Silliman, supra note 42, at 294.
and state governments would not shoulder the considerable expense of providing security for a terrorist trial.\textsuperscript{184} In addition, military commissions are typically considered more efficient and may deliver verdicts more quickly.\textsuperscript{185}

Another very significant difference between Article III courts and military commissions is that the latter offer the possibility of more flexibility with evidentiary rules.\textsuperscript{186} For instance, at a time of military conflict, it may be too onerous for the government to adhere to the strict rules of hearsay, and evidentiary requirements may be adjusted to reflect this reality.\textsuperscript{187}

Finally, there is some force to the argument that international terrorists are committing acts of war, and thus ought not to be treated as domestic criminals simply as a matter of principle.\textsuperscript{188}

Yet, while these concerns (and more) are certainly legitimate and may lead one to conclude that military commissions are the most appropriate forum for international terrorists, it goes too far to say that treating the Christmas Day Bomber and other terrorist suspects as Article III criminal suspects inevitably puts the United States at risk.\textsuperscript{189} The legal protections provided to suspects in Article III courts in terms of access to counsel and interrogation limits are largely the same as in

\textsuperscript{184} Id.

\textsuperscript{185} Id.


\textsuperscript{187} See Hamdi, 542 U.S. at 533–34.

\textsuperscript{188} See Fred Kaplan, \textit{Cheney’s War}, Slate (Feb. 18, 2010), http://www.slate.com/id/2245172/pagenu... Of course, this sort of position cuts both ways, as some have argued al-Qaeda terrorists represent no State and recognizing their attacks as acts of war confers an undeserved “mark of respect” upon them. Wesley K. Clark & Kal Raustiala, Op-Ed., \textit{Why Terrorists Aren’t Soldiers}, N.Y. Times, Aug. 8, 2007, at A19. The inimitable Judge William Young more forcefully asserted this notion when presiding over the sentencing hearing of the Shoe Bomber Richard Reid. See Hon. William G. Young, \textit{A Lament for What Was Once and Yet Can Be}, 32 B.C. Int’l & Comp. L. Rev. 305 app. at 328 (2009). In response to Reid’s declaration that “I am at war with your country,” Judge Young replied:

\begin{quote}
You are not an enemy combatant. You are a terrorist. You are not a soldier in any war. You are a terrorist. To give you that reference, to call you a soldier gives you far too much stature. Whether it is the officers of government who do it or your attorney who does it, or that happens to be your view, you are a terrorist. And we do not negotiate with terrorists. We do not treat with terrorists. We do not sign documents with terrorists. We hunt them down one by one and bring them to justice.
\end{quote}


\textsuperscript{190} See Schulhofer, \textit{supra} note 106, at 506.
military commissions, and it is quite plausible that the type of tribunal used to handle these types of criminals does not make a difference in terms of national security.190

Conclusion

The issue of terrorism and how the United States treats captured terrorist suspects evokes significant criticism and debate from academic, legal, and political circles as well as everyday citizens. There is no doubt that people passionately disagree on many of the central issues including what to do with terrorists such as Abdulmutallab. Those who argue that terrorists should be treated as unprivileged enemy belligerents and thus tried before military commissions rightfully deserve a place at the table to advance their ideas; however, there is no place for debating without facts, and mischaracterizing the obligations of the United States under domestic and international law. There are some differences between the way Abdulmutallab and terrorists like him would be treated in Article III criminal court as compared to military commissions, but it is far from certain that such differences would adversely impact intelligence gathering efforts. While there are other good reasons to support military commissions, politicians, administration officials, and citizens who passionately care about this issue must not lose sight of this point.

190 See 10 U.S.C. § 949c(b) (stating that unprivileged belligerents shall be represented by counsel before a military commission); 42 U.S.C. §§ 2000dd(a), 2000dd(d) (asserting that no individual in U.S. custody shall be subject to cruel, inhuman, or degrading treatment as defined by the Fifth and Fourteenth Amendments of the U.S. Constitution); Mas-siah, 377 U.S. at 206 (holding that a criminal defendant must be represented by counsel after indictment); Bloom & Bronin, supra note 91, at 231–32 (discussing the Due Process Clause limitations on interrogations for Article III criminal suspects).
GERMAN HOMESCHOOLERS AS “PARTICULAR SOCIAL GROUP”: EVALUATION UNDER CURRENT U.S. ASYLUM JURISPRUDENCE

Miki Kawashima Matrician*

Abstract: Thirty years after the enactment of the Refugee Act of 1980, the Board of Immigration Appeals and U.S. courts and have not reached consensus on a uniform definition for the protected category of “particular social group.” The lack of consensus has created much confusion and inconsistent results for applicants seeking asylum in the United States. This Note examines one family’s grant of asylum as a vehicle for analyzing the two main approaches to “particular social group” and argues that the current treatment of the two standards as mutually exclusive by the BIA and the federal courts is inconsistent with the U.N. Guidelines. The Note concludes that U.S. jurisprudence on “particular social group” should mirror the approach of the U.N. Guidelines, which envisions broader protection under that category.

INTRODUCTION

On January 26, 2010, U.S. Immigration Judge Lawrence Burman granted political asylum to the Romeikes, a German family who fled their native country to escape government persecution for homeschooling their children.1 In 2006, Uwe and Hannelore Romeike, concerned that the national school curriculum did not comport with their religious beliefs as evangelical Christians, withdrew their children from a public

---

* Miki Kawashima Matrician is Senior Articles Editor for the *Boston College International & Comparative Law Review*. The author would like to thank her husband, Jeffrey Matrician, and the rest of her family for their encouragement and support.

school in Bissengen, Germany and began educating them at home. They considered the public school curriculum to be contrary to their religious beliefs, in part because language and images contained in textbooks conflicted with their moral views. Concerned that their children were being bullied as a result of their religious beliefs, they chose to withdraw their children from public school and begin educating them at home.

One morning in October 2006, German police officers entered the Romeike home without a written court order, forcibly removed the Romeike children, and escorted them to public school. A few days later, the police returned to the home but were prevented from taking the children by a citizens’ group protesting outside. The parents received several notices from the school principal and the chief law enforcement official in Bissengen ordering them to send their children to school or face legal consequences.

From the time they removed their children from public school until they left for the United States, the Romeikes accumulated approximately $10,000 in fines for refusing to send them to school. They unsuccessfully petitioned the authorities and filed complaints in the courts. After court decisions in 2006 and 2007 paved the way for the German government to take custody of home-schooled children, the Romeikes fled Germany out of fear of losing their children.

The Romeikes entered the United States as tourists in August 2008 and filed for asylum within several months of their arrival. They successfully persuaded the immigration court in Memphis, Tennessee to

---

2 German Family Seeks U.S. Asylum to Homeschool Kids, FOX NEWS (Mar. 31, 2009), http://www.foxnews.com/story/0,2933,511825,00.html; see Moore, supra note 1.
3 See id.
4 See Moore, supra note 1.
6 Id.
7 Id.
9 Respondent’s Brief, supra note 5, at 6.
10 See Moore, supra note 1; Loller, supra note 8; Homeschooling Family Granted Political Asylum, supra note 1.
11 See Moore, supra note 1; Loller, supra note 8; Homeschooling Family Granted Political Asylum, supra note 1.
12 Respondent’s Brief, supra note 5, at 7.
grant asylum on the grounds that the German government persecuted
the family and violated their basic human rights. The court’s decision
was the first grant of political asylum for persecution based on violation
of compulsory education laws.

In light of the U.S. government’s appeal of the immigration
court’s decision to the Board of Immigration Appeals (BIA), this
Note examines the current standard for grant of asylum on the basis of
the protected ground of “particular social group.” Part I of this Note
describes the compulsory education law in Germany under which the
Romeikes claim persecution, explores the practice of homeschooling in
Germany, and provides a brief procedural overview for gaining “refu-
gee” status in the United States. Part II discusses two distinct
approaches adopted by the BIA and federal circuit courts of appeal in
defining “particular social group.” Part III applies those approaches to
the Romeike case and identifies inconsistent applications in the various
circuits and advocates for a uniform standard to comport with the aims
of international law. This Note argues that, on appeal, the BIA should
find that all German homeschoolers comprise a “particular social
group,” regardless of whether the Romeike family successfully estab-
lished a claim of “well-founded fear of persecution.”

I. Background

A. The German Education Law Regime Prohibits Homeschooling

The unfavorable treatment of homeschoolers in Germany is not a
unique phenomenon; indeed, there is a robust debate in the United
States and elsewhere regarding the validity of homeschooling as a
means of education. For instance, in February 2008 the Second Dis-
trict Court of Appeals in Los Angeles handed down a surprising deci-
sion upholding the constitutionality of a state statute prohibiting ho-
meschooling for children between the ages of six and eighteen unless

15 See Homeschooling Family Granted Political Asylum, supra note 1.
16 See Moore, supra note 1.
17 Telephone interview with Michael Donnelly, Staff Att’y, HSLDA (Mar. 12, 2010).
18 At the time of this writing, the BIA had not yet issued a decision in the case. Id.
19 See Aaron T. Martin, Homeschooling in Germany and the United States, 27 Ariz. J. Int’l
& Comp. L. 1, 29–30 (2010); Amanda Petrie, Home Education in Europe and the Implementa-
their parents possess teaching credentials.\textsuperscript{18} Within six months, however, the same court reversed, holding that as long as parents declare their home to be a private school, they may teach their children even without teaching credentials.\textsuperscript{19}

Moreover, there is a movement to end homeschooling in United Nations member countries on the theory that a child’s right to education may be vindicated only by compulsory education in traditional schools outside the home.\textsuperscript{20} By the same token, there is also a movement to recognize the right of parents to choose the appropriate form of schooling for their children.\textsuperscript{21} The issue has also concerned U.S. policymakers; the Georgia and Tennessee state legislatures have passed resolutions expressing disapproval of the German compulsory education laws.\textsuperscript{22}

In Germany, compulsory education laws require that children not only receive formal education from ages six or seven for a period of nine years, they also require all children to attend either a public school or state-approved private school.\textsuperscript{23} In addition, German law generally does not recognize correspondence education for children living within Germany.\textsuperscript{24} Before World War II, the state recognized homeschooling as


\textsuperscript{19} See In re Jonathan L., 81 Cal. Rptr. 3d 571, 590 (Dist. Ct. App. 2008).

\textsuperscript{20} See Martin, supra note 17, at 56; Petrie, supra note 17, at 480.

\textsuperscript{21} See e.g., HOME SCH. LEGAL DEF. ASS’N, http://www.hslda.org (last visited Apr. 16, 2011) (American homeschooling organization that advocates for parents’ freedom of choice over the direction of their children’s education); NETZWERK BILDUNGSFREIHEIT, http://www.netzwerk-bildungsfreiheit.de (last visited Apr. 16, 2011) (a German lobbying organization advancing educational freedom in Germany and providing support for homeschooling families).


\textsuperscript{23} See Thomas Spiegler, Home Education in Germany: An Overview of the Contemporary Situation, 17 EVALUATION & RES. IN EDUC. 179, 180 (2003). By comparison, in the United States, although the education of children is compulsory under the laws of every state, parents are not required to send their children to a state-approved school. See Marilyn Grady et al., Compulsory Education: A Policy Analysis 15 (Apr. 25, 1994), available at http://www.eric.ed.gov/PDFS/ED577556.pdf. Compulsory education statutes differ state by state and impose varying minimum and maximum ages. Id. Courts have recognized parents’ rights to direct their children’s upbringing under the First and Fourteenth Amendments to the U.S. Constitution. See id. at 8.

\textsuperscript{24} Id. Correspondence education for children living abroad is accepted. Id. Correspondence education, or distance learning, is a method of providing education for students who
a valid exception to laws mandating compulsory education. In 1938, however, those exceptions were eliminated, and violations of compulsory education laws triggered criminal penalties.

Such violations constitute an civil offense in the first instance, and can result in a fine of several thousand Euros for subsequent violations. Continued contravention may incur forcible enforcement, and other significant penalties. In extreme cases, courts may revoke custody of the children and even imprison the parents for up to six months. Alternatively, courts may impose fines that accrue daily for as many as 180 days.

In 2003, Germany’s Federal Constitutional Court, its highest court, reaffirmed the government’s authority to compel attendance in state-run schools and held that the state’s interest in ensuring access to adequate education outweighed the parents’ interest in choosing how to educate their children. The court recognized an impracticability exemption for children whose parents’ occupations required extensive travel. It did not, however, create an exemption for homeschooling on the basis of religion or conscience.

B. Homeschooling Movement

Notwithstanding the threat of punishment, the parents of approximately 500 German children choose to teach their children at home. Teaching children outside of the public school setting is commonly known as “homeschooling,” but many parents join together in “learning groups” in a place other than a home to educate their chil-

---

25 See Martin, supra note 17, at 7-8.
26 See id. at 8. For instance, in one case a father who homeschooled his children for religious reasons faced a five-day prison sentence and the loss of child custody. Id.
27 See Spiegler, supra note 23, at 180-81.
28 See id at 181.
29 See id.
30 See id.
31 See Martin, supra note 17, at 19, 22; Homeschooling Family Granted Political Asylum, supra note 1.
32 See Homeschooling Family Granted Political Asylum, supra note 1.
33 See id.
34 Spiegler, supra note 23, at 179.
Some parents even opt to join homeschooling organizations for support, to exchange ideas, and to find legal representation.\textsuperscript{36} German parents have chosen to homeschool for a variety of reasons, including religious concerns or because a child’s medical condition precludes conventional schooling.\textsuperscript{37} Regardless of the impetus for the decision, homeschooling parents generally share one outlook: that they have the right to direct their child’s education, tailored to the individual needs and abilities of the child.\textsuperscript{38} They believe this is consistent with internationally accepted principles of human rights.\textsuperscript{39} The Universal Declaration of Human Rights provides that “[e]veryone has the right to education[;]” “elementary education shall be compulsory[;]” and, “parents have a prior right to choose the kind of education that shall be given to their children.”\textsuperscript{40} Furthermore, Article 13 of the International Covenant on Economic, Social, and Cultural Rights (ICESCR) recognizes:

[T]he liberty of parents . . . to choose for their children schools, other than those established by the public authorities, which conform to such minimum educational standards as may be laid down or approved by the State and to ensure the religious and moral education of their children in conformity with their own convictions.\textsuperscript{41}

That right is also recognized by the European Convention on Human Rights, which states that “the State shall respect the right of parents to

\textsuperscript{35}See Petrie, supra note 17, at 479; Spiegler, supra note 23, at 184.
\textsuperscript{36}See id. For instance, Stuttgart Area Home Schoolers maintains a website that provides resources for curriculum providers, educational requirements, and social gatherings. Stuttgart Area Home Schoolers, http://www.stuttgarthomeschoolers.com (last visited Apr. 16, 2011).
\textsuperscript{37}See Spiegler, supra note 23, at 182, 183. The decision to homeschool is often motivated by religious beliefs, but not always. See Petrie, supra note 17, at 479–80. Parents might have any number of reasons for choosing homeschooling, including a child’s specialized needs, practical reasons that hinder a child’s attendance, or philosophical outlook, among others. See Martin, supra note 17, at 7.
\textsuperscript{38}See Spiegler, supra note 23, at 183.
\textsuperscript{40}Petrie, supra note 17, at 480 (citing Universal Declaration of Human Rights, supra note 39, art. 5).
\textsuperscript{41}ICESCR, supra note 39, art. 13, ¶ 3.
ensure such education and teaching in conformity with their own religions and philosophical convictions.”

In reference to these rights of parents, and relying specifically on the ICESCR, the U.N. Special Rapporteur to the Human Rights Council of the General Assembly has urged the German government to respect a parent’s choice for his or her children. The report stated that “education may not be reduced to mere school attendance,” noted that “[d]istance learning methods and home schooling represent valid options,” and urged states to refrain from restricting “forms of education that do not require attendance at a school.”

German education law, diverging from international conventions, prioritizes the interest of the state in requiring attendance at public schools, with narrow exceptions. Having received penalties and threats from the German authorities, the Romeikes feared losing custody of their children. Ultimately, this fear led them to flee Germany and seek asylum in the United States.

II. Discussion

As previously discussed, to successfully obtain asylum in the United States, a refugee must prove that, on account of at least one of five protected grounds—race, nationality, religion, political opinion, or membership in a particular social group—he or she either suffered persecution, or has a well-founded fear of future persecution. In the

---

42 European Convention on Human Rights, supra note 39, art. 2.
44 See Report by U.N. Rapporteur, supra note 43, ¶ 62. Some countries have confused the right of a child to education with a right to education in a public, state-approved school. See Petrie, supra note 17, at 480.
45 See Martin, supra note 17, at 10; see also Petrie, supra note 17, at 480.
46 See Moore, supra note 1; Loller, supra note 8.
47 See Moore, supra note 1; Loller, supra note 8.
48 An applicant must show that the past persecution or fear of future persecution was “on account of” one of the protected grounds. See Immigration and Nationality Act (INA) § 1101(a)(42), 8 U.S.C. § 1158 (2009). Courts have interpreted the phrase to signify that a “nexus” must exist between the persecution and the protected grounds. See Deborah E. Anker, LAW OF ASYLUM IN THE UNITED STATES 268 (3d ed. 1999).
49 Under the INA, the applicant can either establish a well-founded fear of future persecution upon returning to his or her home country to satisfy that prong, or show past persecution. See Ira J. Kurzman, IMMIGRATION LAW SOURCEBOOK 347, 360 (10th ed. 2006). In INS v. Cardoza-Fonseca, the Supreme Court ruled that applicants are required to establish that “persecution is a reasonable possibility” to satisfy the burden of proof for “a well-
Romeike case, Immigration Judge Burman found that the family had a “well-founded fear of persecution” on account of membership in the “particular social group” of homeschoolers.51

A. Seeking Asylum in the United States—Procedural Overview

Upon arriving in the United States, the Romeikes sought safe haven as refugees fleeing persecution.52 The process begins with an application to the U.S. Citizenship and Immigration Services, to be filed within one year of arrival in the United States.53 Asylum officers process asylum applications, hold “non-adversarial” interviews and make determinations as to the applicant’s eligibility for asylum.54 The officer may consider the applicant’s testimony, information presented at the

---


51 Homeschooling Family Granted Political Asylum, supra note 1.

52 See Respondent’s Brief, supra note 5, at 7; Moore, supra note 1.

53 See INA § 208(a)(2)(B). Applications by refugees may be divided into three categories: affirmative applications, defensive applications, and expedited removal. Aleinikoff, supra note 50, at 849–50. Affirmative asylum applications are those filed by applicants who are not in removal proceedings. See id. at 850. Defensive asylum applications are cases filed by applicants whose removal proceedings are underway and require a higher burden of proof. See id. at 851; Anker, supra note 48, at 17. Expedited removal procedures apply to applicants who enter the United States without an inspection, and must first establish a “credible fear.” Aleinikoff, supra note 50, at 852. After clearing that initial hurdle, those cases are treated as defensive applications. See id. The Romeikes filed an affirmative application a few months after an inspected entry as tourists. See Respondent’s Brief, supra note 5, at 7.

54 See 8 C.F.R. § 1208.9 (2010); Aleinikoff, supra note 50, at 851.
interview, and information from the U.S. State Department and “other credible sources,” including international organizations.\textsuperscript{55} Given the difficulties of proof in asylum cases, the asylum interview is designed to elicit all relevant information, with little limitation on the types of evidence that may be considered.\textsuperscript{56}

The asylum officer decides whether the applicant is eligible for asylum.\textsuperscript{57} Officers may deny a grant of asylum either because the officer does not believe that the applicant satisfied his or her burden of proof in establishing the necessary statutory definition of “refugee,” or because the officer believes the applicant falls outside of the scope of that definition.\textsuperscript{58} Where the asylum officer feels the applicant failed to adequately establish his or her case, the officer refers the case to the immigration court for a hearing.\textsuperscript{59}

An applicant may appeal an unfavorable decision by the immigration judge to the BIA.\textsuperscript{60} Similarly, the Department of Homeland Security (DHS) may also appeal a decision by the immigration judge to the BIA.\textsuperscript{61} In the Romeike case, the government has appealed the immigration judge’s decision in favor of the applicants.\textsuperscript{62} If the BIA denies the appeal, applicants may file a further appeal to the federal Court of Appeals in the circuit in which the case originated.\textsuperscript{63} A circuit court is required to afford great deference to the BIA’s findings of fact and may reverse a BIA decision only if it finds that the BIA committed legal error or abused its discretion.\textsuperscript{64} Generally, when a court of appeals re-

\textsuperscript{55} See 8 C.F.R. §§ 1208.9(e), 1208.11, 1208.12.
\textsuperscript{56} See Anker, supra note 48, at 88.
\textsuperscript{57} See id. Withholding of deportation is unlike an affirmative grant of asylum; it is a mandatory form of relief, and the Attorney General may not return the applicant to his or her home country. An order to withhold deportation, however, is not the same as a grant of permanent residence. See id.
\textsuperscript{58} See Jaya Ramji-Nogales et al., Refugee Roulette: Disparities in Asylum Adjudication and Proposals for Reform 13 (2009).
\textsuperscript{59} See 8 C.F.R. § 1208.14(c)(1).
\textsuperscript{60} See id.
\textsuperscript{61} See id.
\textsuperscript{62} Telephone interview with Michael Donnelly, supra note 15; see also Moore, supra note 1.
\textsuperscript{64} Ramji-Nogales, supra note 58, at 14.
verses a decision of the BIA, it may only remand to the BIA and cannot
grant asylum.\textsuperscript{65}

B. Sources of U.S. Asylum Law

U.S. asylum law derives from international law, particularly the 1951
United Nations Convention relating to the Status of Refugees (Convention)
and the 1967 United Nations Protocol Relating to the Status of
Refugees (Protocol).\textsuperscript{66} In 1980, Congress enacted the Refugee Act,
which aimed to tailor its asylum laws to uphold international treaty obli-
gations under the Protocol.\textsuperscript{67} The Act also standardized the procedure
and requirements for granting asylum to refugees, repealed restrictions
limiting asylum to refugees from certain countries, and extended asy-
rum to people fleeing from “friendly” governments.\textsuperscript{68} Although Con-
gress intended for the definition of “refugee” to mirror the definition in
the Protocol, the Protocol is not binding on the BIA or U.S. courts in
construing legal requirements.\textsuperscript{69}

The definition of “refugee” set forth in § 101(a)(42) of the INA is
identical to the Protocol definition:

\begin{quote}
[A]ny person . . . who is unable to or unwilling to return to,
and is unable or unwilling to avail himself or herself of the
protection of, that country because of persecution or a well-
founded fear of persecution on account of race, religion, na-
tionality, membership in a particular social group, or political
opinion.
\end{quote}

When Congress initially ratified the Protocol, however, it omitted the
“particular social group” basis, despite modeling the remainder of the
definition of “refugee” after the Convention and Protocol.\textsuperscript{70} The Refu-

\begin{footnotesize}
\textsuperscript{65} Id.
\textsuperscript{66} Germain, \textit{supra} note 1, at 1.
\textsuperscript{67} See Cardoza-Fonseca, 480 U.S. at 436–37; Germain, \textit{supra} note 1, at 2–3.
\textsuperscript{68} Maureen Graves, \textit{From Definition to Exploration: Social Groups and Political Asylum Eligibility}, 26 San Diego L. Rev. 739, 744 (1989).
\textsuperscript{69} See Medellin v. Texas, 552 U.S. 491, 505–06 (2008); \textit{In re Acosta}, 19 I\&N Dec. 211, 220 (BIA 1985).
\textsuperscript{70} See INA § 101(a)(42); U.N. Protocol, \textit{supra} note 50, art. I(2). Although Congress intended the definition of “refugee” to mirror the definition in the Protocol, the Protocol is not binding on the BIA or U.S. courts. See \textit{Medellin}, 552 U.S. at 505–06; \textit{Acosta}, 19 I\&N Dec. at 220.
\textsuperscript{71} Graves, \textit{supra} note 68, at 746–47.
\end{footnotesize}
The Refugee Act did not introduce “particular social group” as an additional ground on which asylum could be granted until 1980.72 Congress did not provide guidance on the meaning or the requirements of “particular social group;”73 for this reason, the Supreme Court sought guidance in the Handbook on Procedures and Criteria on Determining Refugee Status published by the Office of United Nations High Commissioner for Refugees (UNHCR) (U.N. Handbook), which emphasizes conformity with the Protocol.74 Accordingly, immigration judges, the BIA, and federal courts frequently cite to the U.N. Handbook in their decisions.75 To supplement the U.N. Handbook, the UNHCR issued Guidelines on International Protection: “Membership of a Particular Social Group,” which provide further guidance on the interpretation of “particular social group.”76

C. Defining a “Particular Social Group”

Of the five protected categories, “particular social group” may be the most difficult to define,77 and neither the Protocol nor Congress provide any clear guidance.78 Due to the lack of a clear definition, conflicting interpretations of the drafters’ intended purpose have spawned confusion on the proper construction of “particular social group.”79 The competing positions may be divided into two camps.80

One view is that Congress intended to meet international human rights standards and to adequately respond to humanitarian needs.81 Proponents of that view cite to the Swedish delegate’s observation at the Conference of Plenipotentiaries that “certain refugees have been persecuted because they belonged to particular social groups,” which led to the addition of the “particular social group” category to the Con-

---

72 See id.
73 See Acosta, 19 I&N Dec. at 232.
74 See Cardoza-Fonseca, 480 U.S. at 436–37.
75 See id.; Poradisova v. Gonzalez, 420 F.3d 70, 79–81 (2d Cir. 2005); Acosta, 19 I&N Dec. at 232.
77 ALEINKOFF, supra note 50, at 897.
78 Acosta, 19 I&N Dec. at 232.
79 See Aleinkoff, supra note 63, at 265.
80 See Anker, supra note 48, at 379; Aleinkoff, supra note 63, at 266.
81 See Graves, supra note 68, at 750.
vention near the end of the deliberations. Some commentators speculate that the impetus behind this late addition was to establish an expansive category to extend protections to those people not covered by the other four grounds.

The other competing view derives from a fear that a broad construction of “particular social group” will result in a flood of asylum seekers. Under this view, limits are necessary to prevent an influx of a large number of people fleeing civil war and ethnic strife. Another theory underpinning this approach is that expanding the category beyond that intended by the drafters would effectively impose obligations on the signatory states to which they did not consent.

The BIA and courts of appeals have developed general standards for interpreting “particular social group” that generally follow two approaches: one focuses on the immutable characteristics common to the group; the other considers external perceptions of the group. The following case illustrations explore the tests, their application, and their ramifications.

1. The Internally Defined Approach

a. Immutable Characteristics Test

The BIA first introduced this approach in Matter of Acosta. To determine the scope of protection, the BIA employed the doctrine of ejusdem generis to identify connections between race, religion, nationality, and political opinion. The BIA identified immutability as the common thread between race and nationality. In recognition of the fact that immutability cannot be a basis for the two other protected categories of
religion and political opinion, the BIA added a second factor to the immutability test, explaining that characteristics which "ought not to be required to be changed" are also protected. The Acosta definition of "particular social group," as a group sharing "a characteristic that either is beyond the power of an individual to change or is so fundamental to individual identity or conscience that it ought not be required to be changed," relies on the internal unifying characteristic of group members. The BIA referred to the U.N. Handbook's suggestion that "a particular social group" connotes "persons of similar background, habits, or social status" in support of the formulation of its standard. Examples of such groups are those characterized by gender, clan membership, sexual orientation, family, shared past experiences, and "matter[s] of conscience."

The BIA applied the standard to determine whether Acosta was a member of a "particular social group" consisting of taxi drivers in San Salvador who "refus[ed] to participate in guerrilla-sponsored efforts to destabilize the government, such as work-stoppages." The BIA ruled that neither being a taxi driver nor refusing to participate in guerrilla-sponsored activities qualified as immutable characteristics, because the members of the group could avoid the threats either by changing their jobs or by cooperating in work stoppages. In upholding the immigration judge's denial of asylum, the BIA noted that, although it would be unfortunate for taxi drivers to have to change their jobs or to have to cooperate with guerrillas, international law does not guarantee an individual's choice of work.

In Lwin v. INS, the Seventh Circuit Court of Appeals applied the Acosta test for "immutable characteristics," noting that it best "preserve[d] the concept that refugee status is restricted to 'individuals who are either unable by their own actions, or as a matter of conscience should not be required, to avoid persecution.'" Lwin, a Burmese citizen, was the father of a student dissident who fled the Burmese military regime's crackdown on protests. After being interrogated twice by Burmese police, he agreed to report any future contact with his son.

91 See Aleinikoff, supra note 63, at 276.
92 See Acosta, 19 I&N Dec. at 233; Anker, supra note 48, at 378.
94 See id.; Aleinikoff, supra note 63, at 276.
95 See Acosta, 19 I&N Dec. at 234.
96 Id.
97 Id. at 213, 234, 236.
98 144 F.3d 505, 512 (7th Cir. 1998).
99 Id. at 507.
Ultimately, he refused to do so and his home was searched on three occasions.\textsuperscript{100} On a visit to see his son in the United States, Lwin filed for asylum on grounds of persecution for his membership in a “particular social group,” which he defined as parents of student dissidents.\textsuperscript{101} In support, he provided evidence that a parent of another dissident had been sentenced to twelve years in prison.\textsuperscript{102} The court reversed the BIA’s denial of asylum, holding that Lwin had established that he was a member of the social group comprised of parents of Burmese student dissidents who have received punishment by the government because of their contact with their dissident children.\textsuperscript{103} The BIA continues to rely on the “immutable characteristics” test as the primary standard by which it determines whether an applicant is a member of a “particular social group.”\textsuperscript{104}

The strength of this internally defined “immutable characteristics” approach is that it provides protection for traits that are fundamental to human dignity, and civil and human rights.\textsuperscript{105} A weakness, however, is that in the absence of clear guidance, the subjective nature of the inquiry of whether characteristics are fundamental to human dignity and “ought not be required to be changed” might improperly result in value judgments rather than legal judgments.\textsuperscript{106} Moreover, critics have argued that the framework denies protection to groups—like students, unions, refugee camp workers, or homeless children—who may be persecuted for an identity that is widely recognized by society, but that nevertheless lacks a sufficient basis in civil or political rights.\textsuperscript{107}

b. \textit{Voluntary Associational Relationship Test}

The Ninth Circuit Court of Appeals diverged from the BIA’s “immutable characteristics” approach and set forth a different definition of “particular social group” in \textit{Sanchez-Trujillo v. INS.}\textsuperscript{108} Although both ap-

\begin{flushright}
\textsuperscript{100} Id. at 508.
\textsuperscript{101} See id.
\textsuperscript{102} See id.
\textsuperscript{103} See id. at 510, 512. Ultimately, the case was remanded on the ground that Lwin had failed to establish that he had a well-founded fear of persecution. See id. at 508.
\textsuperscript{104} In re C-A-, 23 I&N Dec. 951, 956 (BIA 2006) (explaining that the BIA will “continue to adhere to the Acosta formulation” after “[h]aving reviewed the range of approaches to defining particular social group”).
\textsuperscript{105} See Inlender, supra note 85, at 686.
\textsuperscript{106} See Acosta, 19 I&N Dec. at 233; Inlender, supra note 85, at 693.
\textsuperscript{107} See Aleinikoff, supra note 63, at 295; see also Inlender, supra note 85, at 686 (citing James Hathaway, \textit{The Law of Refugee Status} 8 (1991)).
\textsuperscript{108} 801 F.2d 1571, 1576 (9th Cir. 1986); Thomas, supra note 82, at 805.
\end{flushright}
proaches focus on immutable characteristics common to the group, the Ninth Circuit further narrowed the category by requiring that the group result from a “voluntary associational relationship.” Construing the statutory phrase “particular social group,” the court determined that the words “particular” and “social” indicate that the term does not encompass large demographic segments of the population. By requiring that individuals take some affirmative action to affiliate with others, the standard excludes those who have chosen not to associate with others with similar characteristics.

Applying this standard, the Sanchez-Trujillo court determined that young, working class urban males of military age did not constitute a “particular social group” because such a definition was impermissibly broad. In an effort to avoid granting asylum to unacceptably large demographic groups, the court cited the lack of cohesiveness and homogeneity within the group as the basis for denying the group protected status.

The court used contrasting examples to illustrate the scope of the test. It posited hypothetically that a group of males taller than six feet would fall outside the scope of “particular social group,” despite the immutability of a person’s physical features. By contrast, the court explained that the family unit is a paradigmatic example of a “particular social group” and emphasized its fundamental affiliational concerns, common interests, and size. This comparison, however, revealed potential inconsistencies in the Ninth Circuit’s test, given that families are not necessarily voluntary associations.

In Hernandez-Montiel v. INS, the court retreated from the Sanchez-Trujillo test, apparently in an effort to resolve the tension between the BIA’s Acosta ruling and the voluntary associations test in Sanchez-

---

109 See Sanchez-Trujillo, 801 F.2d at 1576; Thomas, supra note 82, at 806.
110 See Sanchez-Trujillo, 801 F.2d at 1576.
111 See Aleinikoff, supra note 63, at 278 (explaining that classes of gays and lesbians are unlikely to be cohesive or display close affiliation among members); Thomas, supra note 82, at 807 (giving as an example Nazi persecution of non-religious Jews despite the victims’ lack of interest in associating with each other and complete assimilation into German society).
112 See Sanchez-Trujillo, 801 F.2d at 1576.
113 See id. at 1577; see also Perdomo v. Holder, 611 F.3d 662, 666 (9th Cir. 2010) (remanding case where BIA denied asylum and declined to recognize “particular social group” consisting of women from Guatemala).
114 See Sanchez-Trujillo, 801 F.2d at 1576.
116 See Sanchez-Trujillo, 801 F.2d at 1576.
117 See Thomas, supra note 82, at 806.
Ostensibly, the court combined the Acosta and Sanchez-Trujillo standards, but it did not provide any guidance for application of this new test in future cases. Holding that the test is whether one is "united by a voluntary association, including a former association, or by an innate characteristic that is so fundamental to the identities or consciences of its members that members either cannot or should not be required to change it," the court determined that Mexican "gay men with female sexual identities" qualified as a particular social group. In its application of the new standard, however, the court implicitly abandoned the "voluntary association" test it purported to be upholding, and instead relied solely on Acosta’s "immutable characteristics" test.

2. Externally Oriented Social Perceptions Approach

In stark contrast to the internally defined "immutable characteristics test," the U.S. Court of Appeals for the Second Circuit interpreted "particular social group" from an external perspective. In Gomez v. INS, the court referred to the Ninth Circuit’s holding in Sanchez-Trujillo, but imposed an additional requirement: the persecutor or society in general must perceive the common characteristic as recognizable and discrete. The court’s emphasis on a common characteristic that “serves to distinguish [a victim] in the eyes of a persecutor—or in the eyes of the outside world in general” departs from precedent. Despite expressing approval of the Sanchez-Trujillo voluntary associations test, the court ultimately ruled based on whether there is an objective perception that a cognizable group exists, rather than on immutable characteristics.

After evaluating whether the attributes of a group comprising El Salvadoran women who had been brutalized by guerrillas were recognizable and discrete from the perspective of the persecutor, the Gomez court denied asylum to a woman who suffered similar abuse. The court dismissed the merits of Gomez’s claim without considering the

118 225 F.3d 1084, 1093–94 (9th Cir. 2000); see Aleinikoff, supra note 63, at 278.
119 See id.
120 Hernandez-Montiel, 225 F.3d at 1093.
121 See id. at 1093–96.
122 Thomas, supra note 82, at 807.
123 947 F.2d 660, 664 (2d Cir. 1991); Thomas, supra note 82, at 807.
124 Thomas, supra note 82, at 807 (citing Gomez, 947 F.2d at 664).
125 See Gomez, 947 F.2d at 664; Inlender, supra note 85, at 696.
126 Gomez, 947 F.2d at 663–64.
group in the social context of El Salvador. Despite having announced its test as one that evaluates a group based on the perspective of the persecutor or that of society in general, the court imposed without explanation a limiting principle, stating that “broadly-based characteristics such as youth and gender will not by itself endow individuals with membership in a particular social group.” Under this new, two-pronged requirement, asylum was denied.

In spite of the ruling in Gomez, commentators have praised the adaptability of the social perceptions approach to evolving social contexts, so that a common trait may set a group apart in one particular social context, but not in another. One commentator argues that whereas some characteristics are immutable because people are born with them or because of past experience, other changeable characteristics such as conduct, manner of dress, or expression of belief are used by persecutors to identify targets. Another commentator, while bemoaning the ruling in Gomez, nevertheless believes that the approach, if applied correctly, is actually broader than the immutable characteristics approach; under this view, Gomez would give fair opportunity for recognition to groups of people that may not share immutable characteristics under the current doctrine and thus would be ineligible for protection.

In contrast, in Gatimi v. Holder, Judge Richard Posner of the Seventh Circuit was highly critical of the BIA’s use of the social perceptions standard. Gatimi argued that she was subject to persecution for her membership in the Kikuyu tribe in Kenya, which forces women to undergo genital mutilation. Reversing the BIA’s denial of asylum, Judge Posner stated that the BIA’s approach was illogical, and concluded that:

If you are a member of a group that has been targeted for assassination or torture or some other mode of persecution, you will take pains to avoid being socially visible; and to the extent that the members of the target group are successful in re-

---

127 See Inlender, supra note 85, at 705.
128 Gomez, 947 F.2d at 664; see Inlender, supra note 85, at 701.
129 Gomez, 947 F.2d at 664; see Inlender, supra note 85, at 701.
130 See Aleinikoff, supra note 63, at 300; Inlender, supra note 85, at 697.
131 See Thomas, supra note 82, at 820.
132 See Inlender, supra note 85, at 705, 709.
133 See Gatimi v. Holder, 578 F.3d 611, 615 (7th Cir. 2009).
134 See id. at 613–14.
maining invisible, they will not be “seen” by other people in the society “as a segment of the population.”

Thus, by Judge Posner’s reasoning, the BIA’s reliance on the externally defined test of “visibility” resulted in the denial of “particular social group” status to women who have not yet suffered genital mutilation and, thus, are not visibly distinct from the rest of society. Such a result would contravene the goal of the Convention to protect targeted individuals from persecution.

Despite Judge Posner’s criticisms, the social perceptions test could be applied in a manner consistent with the humanitarian goals of the Convention. This result may be achieved if the BIA and the courts consider not only whether the common characteristic of the group “serves to distinguish them in the eyes of a persecutor—or in the eyes of the outside world in general,” but also the entire social context, by hearing evidence such as expert testimony, country conditions reports, and media coverage.

3. Recent Cases Illustrating the BIA’s Approach

In the cases mentioned above, the BIA and the courts applied either the internally-defined immutable characteristics approach or the externally-oriented social perceptions approach, thereby treating the two, effectively, as mutually exclusive standards. More recently, however, the courts and the BIA have issued decisions that discuss both approaches. In Matter of C-A-, the BIA recognized that social perceptions are a “relevant factor,” but ultimately relied more heavily on the immutable characteristics test. Because it concluded that the respondent failed both the immutable characteristics and social perceptions test, however, the BIA did not address whether the immutable characteristics and social perceptions constitute a two-pronged requirement.

---

135 See id. at 615.
136 Id.
137 See Id., supra note 85, at 703.
138 See id. at 615.
139 See supra note 85, at 708.
140 See Gomez, 947 F.2d at 663–64; Inlender, supra note 85, at 708.
141 See id. at 663–64; Inlender, supra note 85, at 707.
142 See Gatimi, 578 F.3d at 614–16; Hernandez-Montiel, 225 F.3d at 1093; Lavin, 144 F.3d at 512; Gomez, 947 F.2d at 663–64; Sanchez-Trujillo, 801 F.2d at 1576; Acosta, 19 I&N Dec. at 232.
143 See Santos-Lemus v. Mukasey, 542 F.3d 738, 744 (9th Cir. 2008) (stating that “social visibility” and “particularity” are factors to consider in determining whether a group constitutes a particular social group under the INA); In re S-E-G-, 24 I&N Dec. 579 (BIA 2008); C-A-, 23 I&N Dec. at 956–57.
144 See supra note 85, at 708.
145 See id.
Two years later, in Matter of S-E-G-R, the BIA rendered a decision exhibiting stronger reliance on the social perceptions test than the immutable characteristics test.\footnote{See S-E-G-R, 24 I&N Dec. at 587.} Examining a social group of “Salvadoran youths who have resisted gang recruitment,” the BIA briefly mentioned that age is a mutable characteristic due to its temporary nature.\footnote{See id. at 583 (noting the ages of respondents as 18 and 21 years respectively and that they were 16 and 19 years of age at the time of the hearing).} It then introduced a new test of “particularity” to be satisfied in addition to the social perceptions test.\footnote{See id. at 584.} The stated purpose of the “particularity test” is to limit “particular social group” designation to cases where “the proposed group can accurately be described in a manner sufficiently distinct that the group would be recognized, in the society in question, as a discrete class of persons.”\footnote{See id. at 585 (declining to recognize social group because it makes up a “potentially large and diffuse segment of society”).}

In a companion case, the BIA held that the “analysis must focus on fundamental characteristics and social visibility within the country in question.”\footnote{In re E-A-G-, 24 I&N Dec. 591, 595 (BIA 2008).} The court declined to address issues of “particularity,” but focused instead on “the existence and visibility of the group in the society in question.”\footnote{Id.} From this line of cases, it remains unclear how the two primary tests will be applied in the future.\footnote{See E-A-G-, 24 I&N Dec. at 595; S-E-G-R, 24 I&N Dec. at 584–88; C-A-, 23 I&N Dec. at 956–57.}

### III. Analysis

#### A. The United Nations Guidelines

The BIA and the courts have relied on the U.N. Handbook for guidance in formulating asylum standards.\footnote{See INS v. Cardoza-Fonseca, 480 U.S. 421, 439 (1987); Sanchez-Trujillo v. INS, 801 F.2d 1571, 1576 (9th Cir. 1986); In re Acosta, 19 I&N Dec. 211, 233 (BIA 1985).} That document defines “particular social group” as “persons of similar background, habits or social status;” moreover, suggesting a focus on internal characteristics common to individual members, it notes that claims may overlap with other grounds such as race, religion, or nationality.\footnote{U.N. Handbook, supra note 50, ¶ 77.} The next paragraph, however, implicates the perspective of the persecutor, noting that certain groups may be targeted because of a perceived lack of loy-
alty to the government. No provision in the U.N. Handbook suggests that the approaches are mutually exclusive.

On the contrary, the U.N. Guidelines, issued in 2002, urge the reconciliation of the two standards. The U.N. Guidelines affirm both the immutable characteristics approach and the social perceptions approach. Referring to the immutable characteristics approach as the “protected characteristics approach,” the U.N. Guidelines echo the Acosta standard: “an immutable characteristic may be innate (such as sex or ethnicity) or unalterable for other reasons (such as the historical fact of a past association, occupation or status).” To determine whether a group is defined by such a characteristic, the U.N. Guidelines suggest reference to “a past temporary or voluntary status that is unchangeable because of its historical permanence” or, to “a characteristic or association that is so fundamental to human dignity that group members should not be compelled to forsake it.”

Upon defining the two approaches, the U.N. Guidelines indicate that the two must be reconciled, to prevent gaps in coverage. In a significant move, they formulate a new, unified standard incorporating those two approaches, concluding that:

[A] particular social group is a group of persons who share a common characteristic other than their risk of being persecuted, or who are perceived as a group by society. The characteristic will often be one which is innate, unchangeable, or which is otherwise fundamental to identity, conscience or the exercise of one’s human rights.

The new standard would find a “particular social group” in cases where the internal characteristic standard would have excluded them, such as voluntary behavior, conduct, or expression of belief, which are funda-

---

153 See id., ¶ 78.
154 See generally id.
155 U.N. Guidelines, supra note 76, ¶¶ 7, 8, 10.
156 Id., ¶¶ 6, 7.
157 See id.; Acosta, 19 I&N Dec. at 233.
159 See id. ¶ 7.
160 Id., ¶¶ 6, 7, 10.
161 See id. ¶ 11 (emphasis added).
mental to a person’s identity but which courts might not have considered “immutable.”\textsuperscript{162} Furthermore, that standard would also permit a finding of “particular social group” in cases where, under the social perceptions standard, an applicant would have been excluded from the category for having successfully concealed their identity which would otherwise have been a target.\textsuperscript{163} The U.N. Guidelines suggest that the immutable characteristics and social perceptions test do not comprise two prongs of a test, both of which must be satisfied.\textsuperscript{164} Rather, the U.N. Guidelines indicate that an asylum seeker merits the protection as a member of a “particular social group” upon the satisfaction of one of two tests.\textsuperscript{165}

B. Applying the U.N. Guidelines to the Romeike Case

The BIA should not merely mention the U.N. Guidelines; it should apply both the immutable characteristics and the social perceptions tests to determine whether the Romeikes merit protection as members of a “particular social group.”\textsuperscript{166} In evaluating the family as part of a group of German parents who choose to homeschool their children, the BIA should assess whether the group shares a common characteristic that could be “innate, unchangeable, or which is otherwise fundamental to identity, conscience or the exercise of one’s human rights.”\textsuperscript{167} It should then thoroughly consider whether that group is perceived as a distinct group by German society.\textsuperscript{168} Under this approach, the Romeikes should qualify as a “particular social group” worthy of asylum.\textsuperscript{169}

First, the decision to homeschool one’s children is a characteristic that can be considered immutable.\textsuperscript{170} Characteristics that are “so fundamental to individual identity or conscience that [they] ought not be required to be changed” qualify as an immutable characteristic.\textsuperscript{171} Moreover, under the U.N. Guidelines, “human rights norms may help identify” such characteristics.\textsuperscript{172}

\begin{itemize}
  \item \textsuperscript{162} See Thomas, supra note 82, at 820.
  \item \textsuperscript{163} See Gatimi v. Holder, 578 F.3d 611, 615 (7th Cir. 2009).
  \item \textsuperscript{164} See U.N. Guidelines, supra note 76, ¶ 11.
  \item \textsuperscript{165} See id.; Homeschooling Family Granted Political Asylum, supra note 1.
  \item \textsuperscript{166} See id.
  \item \textsuperscript{167} See id.; Acosta, 19 I&N Dec. at 232.
  \item \textsuperscript{168} See Lwin v. INS, 144 F.3d 505, 512 (7th Cir. 1998); Gomez v. INS, 947 F.2d 660, 663–64 (2d Cir. 1991); Acosta, 19 I&N Dec. at 232; U.N. Guidelines, supra note 76, ¶ 6.
  \item \textsuperscript{169} See U.N. Guidelines, supra note 76, ¶ 6.
  \item \textsuperscript{170} See Acosta, 19 I&N Dec. at 233.
  \item \textsuperscript{171} See U.N. Guidelines, supra note 76, ¶ 6.
\end{itemize}
As “parents who for religious, political, social, academic, or conscientious reasons do not send their children to state-approved schools, but choose to educate them at home,” the Romeikes should qualify as members of a “particular social group.” They believe are “so fundamental to their identities or consciences that they ought not to be required to be changed.” This view is supported by the recognition in the Universal Declaration of Human Rights, the ICESCR, and the European Convention of a parent’s broad latitude to make decisions affecting his or her child’s education, including conscience-based decisions. In light of this broad, international support, the BIA should find that homeschoolers comprise a “particular social group.”

Second, German homeschoolers may properly be perceived as a group by German society in general; indeed, there is ample evidence of a broader homeschooling movement throughout Germany. Furthermore, the German Constitutional Court affirmed that compulsory education laws apply to homeschoolers. The court’s description of homeschoolers, as “religiously or philosophically motivated ‘parallel societies,’” evinces the state’s perception of homeschoolers as a recognizable group.

Moreover, the resolutions passed by American state legislatures strongly suggest that in the United States, German homeschoolers have already received recognition as a distinct group. For example, as previously mentioned, state legislatures in Georgia and Tennessee have passed resolutions urging Germany to allow homeschoolers to determine their children’s education. In order “to prevent the emergence of parallel societies based on separate philosophical convictions,” both resolutions denounce the policy of the German government against

---

173 Respondent’s Brief, supra note 5, at 9.
174 See id. at 11.
175 *See Universal Declaration of Human Rights, supra note 39, art. 5; ICESCR, supra note 39, art. 13, ¶ 3; European Convention for Human Rights, supra note 39, art. 2.*
176 *See Universal Declaration of Human Rights, supra note 39, art. 5; ICESCR, supra note 39, art. 13, ¶ 3; European Convention for Human Rights, supra note 39, art. 2.*
177 *See Inlender, supra note 85, at 707.*
178 *See Respondent’s Brief, supra note 5, at 2–3, 9–12; Petrie, supra note 17, at 490; Spiegler, supra note 23, at 184.*
180 See id.
181 See Respondent’s Brief, supra note 5, at 11; *supra* text accompanying notes 20–23.
182 *See supra text accompanying notes 20–23.*
homeschooling.\textsuperscript{183} The resolution passed by the Georgia legislature also refers to the high academic standards achieved by “home educated students,” reinforcing the perception of homeschoolers as not merely a recognizable group, but as a viable one as well.\textsuperscript{184} Thus, the Romeikes have a strong case for “particular social group” under the social perceptions test recognized by the U.N. Guidelines.\textsuperscript{185}

The BIA should follow the U.N. Guidelines and accord equal weight to both the immutable characteristics and the social perceptions tests and find that the Romeikes are members of a “particular social group” under either test.\textsuperscript{186} By considering the Romeikes’ claims in light of the U.N. Guidelines, the BIA would fulfill the humanitarian vision of the Convention Relating to the Status of Refugees.\textsuperscript{187} Moreover, the proposed approach would further Congress’ intent underlying the enactment of the Refugee Act of 1980—to standardize the procedure and requirement for “particular social group.”\textsuperscript{188}

**CONCLUSION**

The Romeikes made a conscientious choice to homeschool their children because the public school curriculum did not comport with their religious beliefs. By withdrawing their children from public school, they violated German compulsory education laws. They continued to homeschool their children despite considerable financial penalties, and ultimately fled Germany to avoid losing custody of their children. They came to the United States as tourists and sought asylum.

The requirements under current INA provisions on asylum pertaining to “particular social group,” based on the Convention Relating to the Status of Refugees and further implemented by the Refugee Act of 1980, are ill-defined and are the source of much confusion, especially with regard to the “particular social group” category. The BIA and the courts have evaluated “particular social group” under two approaches that have been applied inconsistently, leading to arbitrary denials of asylum. Commenting on the phenomenon, Judge Posner warned that “given the uncertainties in the law, the difficulties in the facts, [and] the seemingly arbitrary variance among the immigration judges, the court of appeals judges are also going to be falling back on

\textsuperscript{183} See Georgia Resolution, supra note 22; Tennessee Resolution, supra note 22.
\textsuperscript{184} See Georgia Resolution, supra note 22.
\textsuperscript{186} See U.N. Guidelines, supra note 76, ¶ 11.
\textsuperscript{187} See U.N. Convention, supra note 50, ¶ 62; Inlender, supra note 85, at 701.
\textsuperscript{188} See Graves, supra note 68, at 743–44.
personal reactions, intuitions, values, and so on. . . . This is supposed to be a uniform body of federal law.”

In the *Romeike* case, the BIA can and should find that the family is a member of a “particular social group” under either the immutable characteristics or the social perceptions approaches. German homeschoolers share characteristics, the validity of which are recognized by international law, that ought not to be changed. The U.N. Declaration of Human Rights, the ICESCR, and the European Convention for Human Rights all recognize a parent’s right to choose the appropriate educational venue for her child. Moreover, homeschoolers in Germany join organizations to provide support for each other, exchange ideas, and share legal representation. They are perceived as a recognizable group by their alleged persecutor, as well as by society at large—in Germany and abroad.

While some groups constitute a “particular social group” under one standard and not the other, the U.N. Guidelines address this discrepancy. In the future, the BIA and the courts must uniformly adopt the standard provided by the U.N. Guidelines. Under the guidelines, the two standards do not comprise a two-pronged test; they constitute two distinct tests, and the satisfaction of one test should lead to a finding of “particular social group.” This approach will realize the vision intended by Congress in its enactment of the Refugee Act of 1980. In formulating a cogent standard, DHS, BIA, and the courts must not be blinded by fears of a flood of applicants. They should strive to fulfill the humanitarian obligations required by the Convention, to provide a safe haven for those in dire straits.

---

189 Ramji-Nogales, supra note 58, at 79.
DEFINING INVESTOR CONFIDENCE: AVOIDING INTERPRETIVE UNCERTAINTY IN CHEVRON CORP. v. ECUADOR

Jason Burke*

Abstract: In an increasingly globalized world, foreign direct investment is becoming an incredibly important tool for investors in developed nations and the developed nations in which they are investing. Investors have increasingly been seeking protections for their investments in foreign nations. This is why approximately 2400 bilateral investment treaties were signed between various nations between 1994 and 2006. When conflicts arise, the job of interpreting these treaties often falls to investment arbitration tribunals. Indeed, in 2010, an arbitration tribunal (Tribunal) operating under the United Nations Commission on International Trade Law (UNCITRAL) rules adjudicated a dispute between Chevron and the Republic of Ecuador (Ecuador) and interpreted the bilateral investment treaty between the U.S. and Ecuador (BIT). This Comment argues that the Tribunal’s interpretation of the BIT was the most reflective of the investor’s expectations and thus encouraged further investment. As incentivizing this investment is the very purpose of the BIT, the Tribunal reached the best possible conclusion as to its meaning.

INTRODUCTION

On March 30, 2010, the Tribunal established under the arbitration rules of UNCITRAL decided Chevron Corp. v. Ecuador, which settled several long-standing disputes between Texaco—a subsidiary of Chevron1—and Ecuador.2 The arbitration concerned seven cases filed by Texaco against Ecuador when Texaco’s two-decade contract for certain oil-related rights in Ecuador expired in 1992.3 By the time Chevron commenced this arbitration, six of the seven cases in question had sat

* Jason Burke is a Staff Writer for the Boston College International & Comparative Law Review.

1 Texaco initially filed the cases at issue in this arbitration against Ecuador; Chevron subsequently acquired Texaco and is now claiming this arbitration through its subsidiary. Chevron Corp. v. Ecuador, UNCITRAL Arbitration, Partial Award on the Merits, ¶¶ 134, 342 (Mar. 30, 2010), http://ita.law.uvic.ca/documents/ChevronTexacoEcuadorPartialAward.PDF.

2 See id.

3 Id. ¶¶ 133–134.
dormant for between thirteen and fifteen years with little action by the
Ecuadorian judiciary. An Ecuadorian court dismissed the last case for
abandonment, but another Ecuadorian court subsequently overturned
the dismissal.

Against this background, the Tribunal found that, by the time the
action was commenced in December of 2006, Ecuador had violated the
BIT by causing undue delay in Texaco’s Ecuadorian legal proceedings.
The Tribunal went on to dismiss Ecuador’s contention that Chevron
was precluded from bringing this arbitration because Texaco and Che-
vron had failed to exhaust every domestic legal remedy. Ultimately, the
Tribunal awarded Chevron $698,621,904.84 in damages. The judgment
was met with swift criticism from Ecuador’s attorney general who
stated that “[t]his new effort to compromise the Ecuadorian state in
its firm commitment to respect the independence of its judicial system
. . . will not succeed.”

Part I of this Comment discusses the origins of the disputes that
led to the Chevron case and the Interim Award regarding the Tribunal’s
jurisdiction. Part II discusses the traditional formulation of denial of
justice under customary international law and compares it to the stan-
dard created by the Tribunal in this case. Part III argues that the Tri-
baln arrived at the correct decision by heavily utilizing the Vienna
Convention on the Law of Treaties (Vienna Convention) to craft a stan-
dard that was extremely faithful to the text of the BIT, thus avoiding
the type of uncertainty that chills investment.

I. Background

A. Texaco's Operations in Ecuador and Related Lawsuits

Relations between the parties in the case commenced in 1964
when Ecuador gave certain rights related to the exploration and pro-
duction of oil to Texaco. In 1973, the 1964 agreement was replaced
by a new agreement (1973 Agreement), which granted Texaco explora-

---

4 Id. ¶ 270.
5 Id.
6 Id.
8 See id. ¶¶ 546–550.
9 See Ecuador Rejects Ruling Awarding $700m to Chevron, Bloomberg Businessweek
10 Chevron Corp. v. Ecuador, UNCITRAL Arbitration, Partial Award on the Merits, ¶ 125
(Mar. 30, 2010), http://ita.law.uvic.ca/documents/ChevronTexacoEcuadorPartialAward.PDF.
tion and drilling rights in Ecuador’s Amazon region, so long as Texaco provided for Ecuador’s domestic oil needs at a lower price to be set by the nation. In 1987, an earthquake shook Ecuador and seriously disrupted Texaco’s ability to deliver enough oil to satisfy domestic needs, forcing Ecuador to seek out alternate sources of oil and pay more than the agreed-upon domestic price. Ecuador later would demand that Texaco provide it with many extra barrels of crude oil at the lower agreed-upon price to compensate its government for the oil bought at a higher price during the crisis.

Efforts to negotiate an extension of Texaco’s rights to explore and exploit oil in Ecuador failed and in 1992, the 1973 Agreement expired. In this time frame, between 1991 and 1993, Texaco filed the seven claims at issue in this case against Ecuador. Five of these cases claimed that Ecuador had overstated its domestic needs, taking more oil than it was entitled to at the low domestic price. Another case involved a force majeure issue related to the 1987 earthquake. The final claim regarded a violation of a refinancing agreement that had been signed between the two parties in 1986. Texaco filed all of these cases between 1991 and 1993. A decade later, Ecuadorian politics entered a period of crisis, during which the nation’s judicial system experienced profound instability. During this period, these seven cases sat mostly dormant despite numerous inquiries by Texaco (and Chevron, after it acquired Texaco).

B. Filing for Arbitration and the 2008 Interim Award

With little progress on its claims for over a decade, Chevron commenced arbitration in a UNCITRAL tribunal in December of 2006. The BIT, which entered into force in 1997, specifies the rights that each

11 Id. ¶¶ 127–128.
12 See id. ¶¶ 130–131.
13 Id. ¶ 132.
14 See id. ¶ 133.
15 See id. ¶ 134.
16 Chevron, UNCITRAL Arbitration, ¶ 135.
17 Force majeure is “[a]n event or effect that can be neither anticipated nor controlled. The term includes both acts of nature (e.g., floods and hurricanes) and acts of people (e.g., riots, strikes, and wars).” Black’s Law Dictionary 718 (9th ed. 2009).
18 Chevron, UNCITRAL Arbitration, ¶ 135.
19 Id.
20 See id. ¶ 140 (Table 1).
21 See id. ¶ 142–143.
22 See id. ¶¶ 255, 270.
23 See id. ¶ 145.
nation shall accord to foreign investors from the other. 24 Further, the BIT specifies that an action questioning a breach of the treaty could be commenced in an UNCITRAL tribunal. 25 At the jurisdictional phase of this arbitration, Ecuador argued that the Tribunal lacked jurisdiction to hear the claim for a variety of reasons. 26 First, Ecuador cited statements made by Chevron and Texaco in a U.S. court action, in which the corporations stated that Ecuador’s judiciary was fair and competent. 27 Ecuador argued that Chevron should not be allowed to contradict itself and claim that Ecuadorian courts were incompetent to decide the cases at issue in this arbitration. 28 Second, Ecuador claimed that the cases in dispute did not concern “investment” as defined by the BIT. 29 Third, Ecuador argued that this claim was not ripe for adjudication as Chevron failed to exhaust all domestic remedies. 30 Finally, Ecuador claimed that the events giving rise to the claims took place before May 11, 1997, the date on which the BIT came into effect, thus making the BIT inapplicable. 31 On December 1, 2008, the Tribunal rejected Ecuador’s contentions about the Tribunal’s lack of jurisdiction and allowed the case to proceed to the merits phase. 32

II. Discussion

A. Denial of Justice Under Customary International Law

That states have a “duty to provide decent justice to foreigners” is one of the oldest precepts of international law. 33 A claim of “denial of justice” has long existed as a way for foreigners to vindicate this right. 34 Though it is widely accepted as a cause of action in international law, the exact definition of the term has been debated throughout the last

---

25 See id. art. VI(3)(a).
26 Chevron, UNCITRAL Arbitration, ¶ 9.
27 See id. ¶ 10.
28 Id.
29 See id. ¶ 12.
30 See id. ¶ 13.
31 Id. ¶ 15.
32 Chevron, UNCITRAL Arbitration, ¶ 25.
33 Jan Paulsson, Denial of Justice in International Law 1 (2005).
34 See id.
Much of this debate is related to the fact that the term “denial of justice” has intentionally been kept a broad and abstract concept. Moving towards the other extreme, a bright-line rule would allow states to get away with treating foreigners poorly by conforming to legal formalities.

Nevertheless, the most basic formulation of a standard denial of justice claim has been articulated by scholars with relative consistency. Responsibility for denial of justice will attach to a state actor only where “the factual circumstances” are “egregious.” Many scholars have historically argued that state actors can be held accountable for denial of justice where the actions of the judiciary are “grossly unfair” or “manifestly unjust.” Yet the question of what constitutes a “manifestly unjust” judgment is subject to considerable debate. Two other categories of denial of justice—the refusal to allow foreigners to assert their rights before domestic courts and judicial delays that are “equivalent to refusal”—are, however, firmly accepted by a majority of international lawyers. Under customary international law, the foreigner seeking redress for denial of justice must also prove that she has exhausted all local remedies. Thus, a denial of justice claim will usually be based on an allegedly unjust final decision of the nation’s highest appeals court,

35 See Alwyn V. Freeman, The International Responsibility of States for Denial of Justice, 96–97 (Kraus Reprint Co. 1970) (1938) (listing six varying definitions of denial of justice); Paulsson, supra note 33, at 65 (stating that “two centuries of debate have focused on” what makes a decision sufficiently “manifestly unjust and one-sided” to amount to a denial of justice); Sir Gerald Fitzmaurice, The Meaning of the Term “Denial of Justice”, 13 Brit. Y.B. Int’l L. 93, 93 (1932) (noting that many possible definitions of denial of justice have been used).
36 See id. at 59 n.2.
37 See id. at 59–60.
38 See J.W. Garner, International Responsibility of States for Judgments of Courts and Verdicts of Juries Amounting to Denial of Justice, 10 Brit. Y.B. Int’l L. 181, 183–84 (1929) (“[A] draft formulated by an American Committee of jurists in 1928 defines denial of justice for which the state is responsible as including the denial, delay, or exceptional difficulty of access to the courts, gross deficiencies in the judicial or remedial process, the absence of those guarantees which are indispensable to the proper administration of justice, or a ‘manifestly unjust’ judgment.”); see also Paulsson, supra note 33, at 65 (citing Vattel with approval, who in 1758 proposed a three-tier definition of denial of justice, including: (1) “not admitting foreigners to establish their rights before the ordinary courts”; (2) “delays which are ruinous or otherwise equivalent to refusal”; and (3) “judgments [that are] ‘manifestly unjust and one-sided’”).
39 Paulsson, supra note 33, at 60.
40 See Garner, supra note 38, at 183.
41 See Paulsson, supra note 33, at 65.
42 Id.
43 See id. at 100.
which illustrates that the entire justice system has failed to adequately protect the rights of the foreigner.\(^{44}\)

Chevron initiated its claim partly upon this customary international law concept of denial of justice.\(^{45}\) As a primary matter, Chevron argued that Ecuador had violated customary international law and committed a denial of justice by means of an undue delay in the court cases it filed against Ecuador that were pending for over a decade.\(^{46}\) By proving that Ecuador violated customary international law, Chevron asserted that it also proved a violation of the BIT,\(^{47}\) which states that “investment . . . shall in no case be accorded treatment less than that required by international law.”\(^ {48}\) In the alternative, Chevron argued that the decisions rendered after the commencement of the action by Ecuador’s courts were so manifestly unjust as to amount to a denial of justice.\(^ {49}\) As a final matter, Chevron insisted that exhaustion of local remedies was not a requirement under customary international law,\(^{50}\) but even if it were a requirement, it would be fulfilled in this case.\(^{51}\)

Ecuador, for its part, emphasized the high burden that the plaintiff carries in an action for denial of justice.\(^{52}\) According to Ecuador, a denial of justice claim based upon undue delay must assert that the state’s delay amounted to a “refusal to judge,”\(^ {53}\) which Ecuador contended that Chevron had not done.\(^ {54}\) In response to Chevron’s alternative argument alleging a denial of justice based on manifestly unjust decisions, Ecuador claimed that the decisions were not unjust, and even if they were, they did not reflect a high enough level of judicial impropriety to breach international law and thereby invoke international arbitration.\(^ {55}\) Finally, Ecuador insisted that Chevron had not exhausted all domestic remedies, as most of the cases had not been appealed to the highest court.\(^ {56}\) Despite the robust corpus of customary international

\(^{44}\) See id. at 100–01.


\(^{46}\) See id. ¶¶ 167–171.

\(^{47}\) Id. ¶ 167.

\(^{48}\) Investment Treaty, supra note 24, art. II(3)(a).

\(^{49}\) Chevron, UNCITRAL Arbitration, ¶ 188.

\(^{50}\) Id. ¶ 277.

\(^{51}\) See id. ¶ 279.

\(^{52}\) Id. ¶ 178.

\(^{53}\) Id. ¶ 180.

\(^{54}\) See id. ¶ 181.

\(^{55}\) See Chevron, UNCITRAL Arbitration, ¶ 195.

\(^{56}\) See id. ¶¶ 295–297.
law concerning denial of justice that both parties cited as central to their arguments, the Tribunal’s opinion proceeded in quite a different direction.57

B. A Lower Standard for U.S.-Ecuador Relations:
The BIT’s Distinct Standard

Even though the customary international law framework discussed above informed the Tribunal’s decision, the Tribunal noted that its job was to interpret the BIT according to the rules set forth in the Vienna Convention.58 As such, the Tribunal used a great deal of the opinion to interpret the text of Article II(7) of the BIT,59 which requires that “[e]ach Party shall provide effective means of asserting claims and enforcing rights with respect to investment, investment agreements, and investment authorizations.”60

Though the Tribunal observed that the obligations imposed by Article II(7) are similar to those imposed by customary international law,61 it found that the standard set by the BIT was distinct from the standard for finding a denial of justice under customary international law.62 Indeed, the Tribunal observed that the standard necessary for finding a breach of the BIT could be lower than that set by customary international law.63

Thus, rather than having to prove that the state’s judicial shortcoming was “egregious,” as is the general standard under customary international law,64 the claimant must only prove that the state failed to provide an “effective means” for enforcing rights and bringing claims.65 Specifically, the Tribunal noted that the standard for finding a denial of justice under customary international law “requires the demonstration of ‘a particularly serious shortcoming’ and egregious conduct that ‘shocks, or at least surprises, a sense of judicial propriety.’”66 By contrast, the Tribunal held that “under Article II(7), a failure of domestic courts to enforce

---

57 See id. ¶ 242.
58 See id. ¶¶ 158–159, 244.
60 Investment Treaty, supra note 24, art. II(7).
61 See Chevron, UNCITRAL Arbitration, ¶ 244.
62 See id. ¶¶ 242–244.
63 See id. ¶ 244.
64 See Paulsson, supra note 33, at 60.
65 See Chevron, UNCITRAL Arbitration, ¶ 244.
66 Id. (quoting Opinion of Jan Paulsson, para. 10 (Nov. 2008) and Respondent’s Counter-Memorial on the Merits of Sept. 22, 2008).
rights ‘effectively’ will constitute a violation of Article II(7).” As to the relationship between the two standards, the Tribunal notes that judicial misconduct that violates Article II(7) “may not always be sufficient to find a denial of justice under customary international law.” The Tribunal’s decision, therefore, creates a standard that is distinct from—and seemingly easier to violate than—the traditional standard.

Moreover, the Tribunal decided that although a party must have exhausted local remedies in order to prove a denial of justice under customary international law, Article II(7) once again creates a distinct standard in this regard. The Tribunal was convinced that exhaustion of local remedies was a general requirement under customary international law. In contrast, the central issue under Article II(7) is whether the state has provided the claimant with an effective means for enforcing rights. The Tribunal found that only “a qualified requirement of exhaustion of local remedies applies under the ‘effective means’ standard of Article II(7).” If the respondent state can prove that valid local remedies did exist and were not pursued, a claimant’s contention that it was not provided with effective means for enforcing its rights may be defeated. Nevertheless, proof of exhaustion of local remedies is not necessary to make a prima facie case for a breach of Article II(7).

Proceeding with an analysis of the case under this legal framework, the Tribunal found that an undue delay serves to violate the new standard, as Article II(7) mandates that Ecuador “provide foreign investors with means of enforcing legitimate rights within a reasonable amount of time.” Although Ecuador highlighted various legal options not used by Chevron, the Tribunal was ultimately convinced that such options would have been ineffective and thus dismissed Ecuador’s defense. After considering and dismissing a variety of factors that could have justified a delay, the Tribunal ruled that Ecuador’s long period of inaction constituted an undue delay, in violation of Article II(7).
III. Analysis

The Tribunal was correct to use the Vienna Convention to carve out a standard that was distinct from the customary international law standard because such an interpretation of the BIT is the most predictable. Such predictability best promotes investment. Bilateral investment treaties have flourished as a way to give investors the certainty and security they seek when investing abroad. Interpretations of bilateral investment treaties that create uncertainty for investors may serve to chill investment, which such treaties were often meant to promote. Ultimately, in the instant case, Chevron’s expectations of its own rights under the BIT were based heavily on an interpretation informed by the Vienna Convention. Thus, by strictly applying the Vienna Convention to its analysis of the BIT, the Tribunal promoted certainty and security for American investors like Chevron in Ecuador. Any other interpretation would have created uncertainty about the rights guaranteed to investors by the BIT, thus chilling investment and contravening the paramount goal of the BIT itself.

80 See Chevron Corp. v. Ecuador, UNCITRAL Arbitration, Partial Award on the Merits, ¶¶ 159, 161–162, 242–244 (Mar. 30, 2010), http://ita.law.uvic.ca/documents/ChevronTexacoEcuadorPartialAward.PDF (stating that both parties relied on the Vienna Convention in making their arguments regarding the substantive provisions of the BIT, explaining that the Tribunal actually applied the Vienna Convention to interpret the BIT, and ultimately creating a standard distinct from the international law standard).

81 See Susan D. Franck, The Legitimacy Crisis in Investment Treaty Arbitration: Privatizing Public International Law Through Inconsistent Decisions, 73 Fordham L. Rev. 1521, 1538, 1558 (2005) (arguing that increasing investors’ confidence in their ability to enforce rights under bilateral investment treaties promotes foreign investment and asserting that inconsistency in the interpretation of bilateral investment treaties creates uncertainty about and damages expectations of the rights afforded by such a treaty).


83 See id. at 139 (“Allowing the host State to renege on its agreement in the BIT creates uncertainty in the global marketplace and can serve only to discourage foreign investment.”).

84 See id. at 135 (“The BIT serves to attract foreign investment by granting broad investment rights to investors and creating flexibility in the resolution of investment disputes.”).

85 See Chevron, UNCITRAL Arbitration, ¶ 159.

86 See id. ¶¶ 161–162.

87 Cf. Franck, supra note 81, at 1558 (claiming that inconsistent interpretations of bilateral investment treaties decrease certainty about the rights guaranteed by the treaty, thereby damaging investors’ expectations).

88 See id.

89 See Investment Treaty, supra note 24, pmbl. (stating that the United States and Ecuador entered into the BIT “desiring to promote greater economic cooperation between
In a globalized world with increasing amounts of investment flowing across national borders, parties looking to invest in foreign nations seek increasing amounts of security. To this end, many nations have sought to guarantee expansive protection to investors by entering into bilateral investment treaties. These treaties “govern[] the treatment of investments made in the territory of each state by individuals or companies from the other state.” These investment treaties have proliferated, particularly between 1994 and 2006, by which time there were approximately 2400 bilateral investment treaties between various nations.

Given that the increasingly popular use of bilateral investment treaties has at its core the goal of securing foreign investment, interpretations that create uncertainty about the rights enshrined in such treaties ultimately discourage investment. Jarrod Wong, a former legal adviser at the Iran-U.S. Claims Tribunal, in talking about the “umbrella clause” of certain bilateral investment treaties, highlights several arbitral decisions that he argues were under-protective of investment because they failed to faithfully interpret the text of the relevant treaties. Wong argues that a failure to interpret an “umbrella clause” in accordance with its text and history allows host states to renge on their responsibilities under bilateral investment treaties, creating uncertainty for investors from which foreign investment will suffer.

Similarly, in an article addressing a number of inconsistent arbitral decisions arising out of bilateral investment treaties, Susan D. Franck, a practitioner and professor of international arbitration law, also expresses concern about the deleterious effects of interpretive uncertainty

---

90 See Franck, supra note 81, at 1525.
91 See Wong, supra note 82, at 135–36.
92 Id. at 141.
93 See id. at 135–36.
94 See id. at 139.
95 An “umbrella clause” is a clause that “imposes a requirement on each Contracting State to observe all investment obligations entered into with investors from the other Contracting State.” Id. at 136. Liberally interpreted, such a clause allows an arbitration tribunal interpreting a bilateral investment treaty to exercise “jurisdiction over breach-of-contract claims since a breach of the investment contract is also a breach of the umbrella clause.” See Wong, supra note 82, at 137.
96 See id. at 139.

2011] Avoiding Interpretive Uncertainty in Chevron v. Ecuador 473

on foreign investment.\textsuperscript{99} She states that “[i]nconsistency creates uncertainty and damages the legitimate expectations of investors and Sovereigns.”\textsuperscript{100} Franck notes that one main effect of bilateral investment treaties is to increase investment incentives by granting rights to investors and thus reducing the risks associated with investment; it logically follows that throwing such rights into doubt will increase risk and thereby stifle investment.\textsuperscript{101}

The interpretation of the BIT at issue in \textit{Chevron} that best encourages investment is one that protects the expectations of investors by applying the source widely accepted as an accurate statement of customary international law on treaty interpretation—the Vienna Convention.\textsuperscript{102} Such an interpretation is the least likely to create the type of confusion that would chill the very investment that a bilateral investment treaty is designed to ensure.\textsuperscript{103} To that end, the Tribunal noted that \textit{Chevron} firmly accepted that the correct way to interpret the BIT was by using the Vienna Convention, which reflects customary international law with regard to treaty interpretation.\textsuperscript{104} Therefore, an interpretation of the BIT based firmly in the Vienna Convention would be most reflective of the investor’s expectations and therefore most likely to create the desired certainty and investment.\textsuperscript{105}

The Tribunal found that the relevant provisions of the Vienna Convention with regard to treaty interpretation are Articles 31 and 32.\textsuperscript{106} First, Article 31 states that a treaty should be “interpreted in good faith in accordance with the ordinary meaning to be given to the terms of the treaty in their context and in the light of its object and purpose.”\textsuperscript{107} As stated by the Tribunal, the explicit refusal to include customary international law as the standard for finding a breach of the BIT meant that there was a different standard.\textsuperscript{108} Indeed, the “effective means” standard created by the court is exceptionally consistent with this interpretive tenet in that it specifically uses the words of the BIT

\begin{footnotesize}
\begin{enumerate}
\item[99] See Franck, \textit{supra} note 81, at 1558.
\item[100] Id.
\item[101] Id. at 1538.
\item[102] See \textit{Chevron}, UNCITRAL Arbitration, ¶ 159; Franck, \textit{supra} note 81, at 1538, 1558.
\item[103] See Wong, \textit{supra} note 82, at 135, 139.
\item[104] See \textit{Chevron}, UNCITRAL Arbitration, ¶ 159.
\item[105] See id.; Franck, \textit{supra} note 81, at 1538.
\item[106] See \textit{Chevron}, UNCITRAL Arbitration, ¶¶ 161–162.
\item[108] \textit{Chevron}, UNCITRAL Arbitration, ¶ 242.
\end{enumerate}
\end{footnotesize}
Further, Article 32 of the Vienna Convention allows a tribunal to consider other factors—such as the “preparatory work of the treaty and the circumstances of its conclusion”—but only to confirm the meaning dictated by Article 31. Applying Article 32, the Tribunal noted that this BIT was drafted at a time when the United States deemed more specific protection for investors to be desirable, thus confirming the “effective means” standard mandated by Article 31. Ultimately, the Tribunal’s interpretation of the BIT was remarkably consistent with the Vienna Convention. As Chevron had based its arguments around the Vienna Convention, this interpretation most closely reflected its expectations. Ultimately, this interpretation promoted the most possible certainty for potential future investors.

**Conclusion**

The concept of denial of justice is not a new one under customary international law. Indeed, decades of scholarship have coalesced around a general theory that, to amount to a denial of justice, judicial misconduct must be egregious. Moreover, a party claiming to be a victim of denial of justice must prove that it has exhausted all local remedies. In *Chevron*, however, the Tribunal found that while the customary international law standard of denial of justice was influential, it was not determinative, as the job of the Tribunal was to interpret the BIT. As such, the Tribunal found that Article II(7) set forth a distinct “effective means” standard, which was less demanding than customary international law with regard to proving denial of justice and exhaustion of local remedies. This less demanding standard is preferential because the Tribunal strictly applied the Vienna Convention to create a faithful, text-based interpretation of the BIT. Because the Vienna Convention had in-

---

109 See Investment Treaty, supra note 24, art. II(7) (“Each party shall provide effective means of asserting claims and enforcing rights with respect to investment, investment agreements, and investment authorizations.”) (emphasis added); Vienna Convention, supra note 107, art. 31(1) (“A treaty shall be interpreted in good faith in accordance with the ordinary meaning to be given to the terms of the treaty in their context and in the light of its object and purpose.”); *Chevron*, UNCITRAL Arbitration, ¶ 242 (“The Tribunal thus finds that Article II(7), setting out an ‘effective means’ standard, constitutes a *lex specialis* and not a mere restatement of the law on denial of justice.”).

110 See Vienna Convention, supra note 107, art. 32.

111 See *Chevron*, UNCITRAL Arbitration, ¶ 243.

112 See Vienna Convention, supra note 107, arts. 31–32; *Chevron*, UNCITRAL Arbitration, ¶¶ 242–244.

113 See *Chevron*, UNCITRAL Arbitration, ¶ 159.

114 See Franck, supra note 81, at 1538.
formed Chevron’s interpretation of the BIT and its rights as an investor, the Tribunal ultimately confirmed Chevron’s expectations and avoided creating the type of interpretive uncertainty that can chill the very investment that the BIT was meant to foster. Such an approach for future Tribunals interpreting the BIT would further increase certainty for investors and incentivize U.S. investment in Ecuador.
THE INTERNET IS NOT A LAWLESS PRAIRIE: DATA PROTECTION AND PRIVACY IN ITALY

Noah C.N. Hampson*

Abstract: As the Internet has become more pervasive, so too have concerns about the security of personal data online. The rapid evolution of Internet technology has outpaced the legislative process, leaving courts to resolve complex and important questions of policy. Their answers to these questions can have dramatic implications for the future of the Internet as a platform for global communication. The judicial decisions in this area are frequently issued ad hoc by judges who may be unfamiliar with the technology at issue and unaware of the potential ramifications of their rulings. The February 2010 conviction in Italy of three Google executives for violations of data protection laws sparked widespread controversy and criticism on this basis. This Comment argues that the Italian court’s decision is a prominent example of the broader trend of inexpert judicial regulation of the Internet.

INTRODUCTION

On February 24, 2010, Italian Judge Oscar Magi convicted three Google executives for violations of Italian data protection laws after a video was uploaded to Google Video showing an autistic student being bullied by classmates.¹ Although the substance of the court’s decision was not released until April 12, the announcement immediately ignited controversy and debate over the decision’s ramifications for Internet regulation and freedom of expression.² Some commentators assailed

---

* Noah C.N. Hampson is a Staff Writer for the Boston College International & Comparative Law Review. He would like to thank Professor Mary-Rose Papandrea for providing the inspiration for this Comment, Steven Joseph for his invaluable editorial assistance, and Olga Laur for her support and encouragement.


the verdict as a threat to commercial Internet companies and to the freedom of the World Wide Web.³ Others welcomed it as a vindication of an individual’s right to privacy.⁴

Prosecutors originally charged four Google executives with defamation, and three with violating Italy’s Personal Data Protection Code.⁵ When the decision was finally released, it revealed that the court had convicted the executives charged with violation of data protection law but had acquitted them all of the defamation charge.⁶ Surprisingly, the court imposed liability despite Italian regulations,⁷ European Union directives,⁸ and a contemporaneous judgment by the European Court of Justice,⁹ all of which exempted certain categories of information society service providers (ISSPs) from liability for user-generated content.¹⁰ The court’s decision substantiates concerns that advances in technology vastly exceed the pace of formal Internet regulation.¹¹ The


¹¹ See, e.g., David R. Johnson & David Post, Law and Borders—The Rise of Law in Cyberspace, 48 STAN. L. REV. 1367, 1370 (1996) (arguing that the Internet undermines the link between geographical location and government power to regulate online behavior); Paul Krugman, Facing the Music, N.Y. TIMES, July 30, 2000, at WK15 (suggesting that technology
decision also raises serious questions about the data protection responsibilities of ISSPs and the viability of a truly “free” Internet in the international marketplace.12

Part I of this Comment provides background on the trial and the case against Google. Part II discusses existing European Union and Italian law pertaining to data protection and Internet privacy. It also explains a recent judgment by the European Court of Justice interpreting the responsibilities of ISSPs with regard to personal privacy. Part III analyzes the reasoning by which the court reconciled its decision to impose liability with seemingly contradictory precedent. This section also interprets the Google decision in the context of the broader international debate over Internet regulation; specifically, it focuses on the increasingly prominent role of courts in regulating the Internet. Finally, this Comment concludes that because the rate of technological advances far outpaces efforts at formal rulemaking, Internet regulation is occurring in courts around the world on an ad hoc basis, with significant ramifications for privacy rights, freedom of expression, and the future of the Internet itself.

I. Background

On September 8, 2006, an Internet user named Giulia Lisa uploaded a three-minute video to Google Video depicting an autistic boy named Francesco Giovanni De Leon being harassed and insulted by a group of schoolmates in Turin, Italy.13 Over the next two months, the video was viewed more than five thousand times and became one of the most popular videos in Google Video’s “Funny” category.14 During that

---


14 Sartor & Viola, supra note 5, at 357.
time, users commented on the video and some may even have flagged it for inappropriate content. On November 6, 2006 Google received an email from a user complaining about the video and requesting its removal. Although there is evidence that Google employees undertook to remove it within twenty-four hours, the video was not actually removed until Italian police contacted Google and demanded that the video be taken down. A two-year investigation ensued, and in early November 2008 four Google executives—David Drummond, senior vice-president and chief legal officer; Peter Fleischer, global privacy counsel; Arvind Desikan, senior product marketing manager; and George Reyes, former chief financial officer—were charged in Italian court with defamation and violation of Italy’s Personal Data Protection Code.

The defamation charge alleged that Google and its executives had contributed to the defamation of the student depicted in the video. Prosecutors argued that by failing to exercise any positive control to prevent defamatory content from appearing on Google Video, Google had violated Article 40 of the Italian Criminal Code. The data protection charge alleged that Google illegally processed sensitive personal data for the purpose of obtaining a gain in violation of Section 167 of Italy’s Personal Data Protection Code. After a bench trial presided over by Judge Magi, all four executives were ultimately acquitted of the defamation charge, but three were convicted of violating provisions of the Personal Data Protection Code and sentenced to six-month suspended prison terms.

---

15 Id.
17 See Caso Google, Com’è Andata, supra note 16.
18 See Sartor & Viola, supra note 5, at 357; Caso Google, Com’è Andata, supra note 16.
19 See Sartor & Viola, supra note 5, at 357.
20 Id.
21 See id. at 359 (explaining that Article 40 imposes criminal liability for “failing to prevent an event which one has the legal obligation to prevent,” and that such failure “amounts to causing” that event).
22 Id. at 361–62.
II. Discussion

A. The European Union Directives

In the past fifteen years, the European Parliament has issued several Directives to member states of the European Union concerning Internet regulations. In particular, Directive 2000/31 (Liability Directive) promulgates regulations intended to ensure “a high level of Community legal integration in order to establish a real area without internal borders for information society services.” Although the Directive provides that “the protection of individuals with regard to the processing of personal data is solely governed by Directive 95/46/EC of the European Parliament and of the Council . . . and Directive 97/66/EC of the European Parliament and of the Council,” Articles 12–14 establish certain liability exemptions for information society services. These Articles specifically exempt services that are limited to a “mere conduit” of data, “caching,” or “hosting.” In addition, Article 15 prevents member states from imposing a general obligation on providers of these services to monitor information they transmit or store. It also prevents member states from imposing a general obligation to actively investigate illegal activity.

With regard to personal data, Directive 95/46 (Privacy Directive) articulates principles for the protection of “fundamental rights and freedoms” of natural persons—in particular the right to privacy—in the processing of personal data. The Privacy Directive also prohibits interference with the free flow of personal data between member states.

25 Id. para. 14.
26 See id. arts. 12–14.
27 Id. art. 12.
28 Id. art. 13.
29 Id. art. 14.
30 Directive 2000/31, art. 15.
31 Id.
32 Council Directive 95/46, art. 1(1), 1995 O.J. (L 281) 31 (EC). “Personal data” is defined as “any information relating to an identified or identifiable natural person (‘data subject’); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.” Id. art. 2(a). “Processing” is defined as “any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction.” Id. art. 2(b).
except under certain circumstances. The overarching purpose is to coordinate regulation across member states, thereby ensuring uniform protection of individual rights, removing impediments to economic activity and law enforcement, and still leaving some latitude for states to tailor national data protection laws within a limited margin.

B. Italian Law

Italian Legislative Decree no. 196 of 30 June 2003, known as the Personal Data Protection Code (PDPC), regulates all personal data processing in Italy. Unlike the Privacy Directive, which addresses concerns relating both to protection of personal data and to preserving the free flow of information between EU member states, the purpose of the PDPC is to “ensure that personal data are processed by respecting data subjects’ rights, fundamental freedoms and dignity, particularly with regard to confidentiality, personal identity, and the right to personal data protection.” Pursuant to the Privacy Directive, Section 153 establishes the Italian personal data protection supervisory authority charged with enforcing data protection rules: the Garante Per La Protezione Dei Dati Personali (Garante). The PDPC requires the consent of the data subject before personal data may be processed. Additionally, when the data consist of “sensitive” information, the written consent of the data subject is required, as well as written permission from the Garante.

33 Id. art. 1(2).
34 See id. paras. 7–9.
36 Directive 95/46, art. 1.
37 D.Lgs. n. 196/2003, § 2. Paragraph 2 of Section 2 notes that “[t]he processing of personal data shall be regulated by affording a high level of protection for the rights and protections referred to in paragraph 1 in compliance with the principles . . . by which data subjects can exercise such rights and data controllers can fulfill the relevant obligations.” Id. § 1(2).
38 Id. §§ 153–154.
39 Id. § 23. “Processing” and “personal data” are both defined in the PDPC largely the same as they are defined in Article 2 of Directive 95/46. Compare id. § 4(1)(a)–(b), with Directive 95/46, art. 2(a)–(b).
40 D.Lgs. n. 196/2003, § 26(1). “Sensitive data” is defined in the PDPC as “personal data allowing the disclosure of racial or ethnic origin, religious, philosophical or other beliefs, political opinions, membership of parties, trade unions, associations or organizations of a religious, philosophical, political or trade-unionist character, as well as personal data disclosing health and sex life.” Id. § 4(1)(d). Moreover, Section 26(5) explicitly forbids the dissemination of health data. Id. § 26(5).
Section 167 of the PDPC—the provision under which the Google executives were sentenced—provides that persons who process personal data "with a view to gain for himself or another or with intent to cause harm to another"\(^\text{41}\) in breach of certain other provisions of the PDPC shall be punished by imprisonment for six to thirty-six months, depending on the nature of the violation.\(^\text{42}\) Thus, the Google executives were convicted of participating in the processing of De Leon’s sensitive personal data without his consent or that of his guardians, without the permission of the Garante, and with a view to profit.\(^\text{43}\)

Legislative Decree no. 70 of 9 April 2003 implemented the Liability Directive in Italy and governs certain aspects of electronic commerce.\(^\text{44}\) Article 16 of the Decree exempts ISSPs from liability for user-generated content, provided they have no actual knowledge of illegal content or activity, and that upon learning of illegal content or activity the provider acts “immediately to remove or to disable access” to the content.\(^\text{45}\) For ISSPs whose activities are limited to that of a “mere conduit,” “hosting,” or “caching,” Article 17 provides that there is no general duty to monitor content.\(^\text{46}\)

C. The European Court of Justice

Not long after the convictions were announced, but before the court’s decision was released, the European Court of Justice (ECJ) issued a judgment in a trademark infringement case in which it interpreted the exemption provisions of the Liability Directive as applied to Google’s AdWords marketing system.\(^\text{47}\) The plaintiffs, purveyors of luxury goods, sued Google for contributory trademark infringement after AdWords advertisements for counterfeits of the plaintiffs’ products appeared online.\(^\text{48}\) AdWords is a paid information referencing service which displays “sponsored links” alongside the natural search results of a user’s query in Google’s search engine.\(^\text{49}\) AdWords was also used to

\(^{41}\) Id. § 167(1).
\(^{42}\) See id. § 167(1)–(2).
\(^{43}\) Sartor & Viola, supra note 5, at 362.
\(^{44}\) D.Lgs. n. 70/2003, art. 16; Sartor & Viola, supra note 5, at 360.
\(^{45}\) See D.Lgs. n. 70/2003, art. 16; Sartor & Viola, supra note 5, at 360.
\(^{46}\) D.Lgs. n. 70/2003, art. 17.
\(^{48}\) Id. paras. 28–30.
\(^{49}\) Id. paras. 23–24.
place advertising on Google Video.\textsuperscript{50} Advertisers are able to reserve one or more keywords, and when a search is entered containing those keywords a link to the advertised website appears accompanied by a short commercial message.\textsuperscript{51} The process of choosing and reserving keywords, drafting the commercial message, and attaching a link, is performed exclusively by the advertiser; the advertisements are generated automatically by AdWords.\textsuperscript{52} Google charges a fee for this service based, among other factors, on the number of clicks the link receives and a maximum price per click selected by the advertiser.\textsuperscript{53}

The ECJ ruled that Article 14 of the Directive must be interpreted to mean that the exemption from liability applies to a service like AdWords if that service “has not played an active role of such a kind as to give it knowledge of, or control over, the data stored.”\textsuperscript{54} If it has not, the service “cannot be held liable for the data which it has stored at the request of an advertiser, unless, having obtained knowledge of the unlawful nature of those data or of that advertiser’s activities, it failed to act expeditiously to remove or to disable access to the data concerned.”\textsuperscript{55} The ECJ did not rule on whether Google played a sufficiently active role, nor if it failed to act expeditiously upon receiving information concerning the alleged infringement.\textsuperscript{56} The ECJ did note, however, that it was necessary to determine whether Google was a neutral participant, and to examine Google’s role in selecting keywords and drafting the commercial message that accompanied the sponsored links.\textsuperscript{57} The court also noted that the mere fact that Google charges a fee for AdWords, and that a relationship may exist between the appearance of an advertiser’s keyword and the terms in a user’s query, does not deprive Google of the exemption from liability provided for in the Liability Directive.\textsuperscript{58}

The ECJ’s judgment reflects the principle embodied in the Liability Directive that the free flow of information over the Internet should be protected.\textsuperscript{59} Because of the risk to free expression and electronic commerce, the Liability Directive suggests—and the ECJ confirmed—

\textsuperscript{50} Sentenza n. 1972/2010, at 17–18 n.30; see Sartor & Viola, supra note 5, at 363, 372–73.
\textsuperscript{51} Joined Cases C-236/08 to C-238/08, paras. 23–24.
\textsuperscript{52} See id. para. 27.
\textsuperscript{53} Id. para. 25.
\textsuperscript{54} Id. para. 120.
\textsuperscript{55} Id.
\textsuperscript{56} See id. para. 119.
\textsuperscript{57} Joined Cases C-236/08 to C-238/08, paras. 114, 118.
\textsuperscript{58} Id. paras. 116–117.
that liability for user-generated content should be limited to instances in which the ISSP has either participated in illegal activity, or failed to prevent or mitigate the damage caused by it.60

III. Analysis

Given the level of statutory protection afforded to ISSPs under both EU and Italian law, the conviction of the Google executives is surprising for several reasons. First, the court acknowledges that requiring ISSPs to monitor all user-generated content in order to prevent the upload of personal data would be impossible; indeed, the court states that ISSPs have no legal obligation to do so.61 Because the prohibition on legally required monitoring is derived from Article 15 of the Liability Directive, it is noteworthy that the court did not discuss the Directive or the corresponding Italian Legislative Decree that implemented the Liability Directive in Italy.62 This omission may be partially explained by the court’s conclusion that Google Video is not a “mere conduit,” a “host,” or limited to “caching” services, but is instead an “active” content provider and as such is ineligible for the exemption.63 Nevertheless, if it is impossible for Google to monitor all content, then imposing liability for failing to prevent uploading of illegal content would severely burden Google’s operations in Italy.64 After all, the decision implies that Italian law requires Google to do something that the court itself acknowledges is impossible.65

Second, several aspects of the court’s decision contradict the ECJ’s judgment concerning Google’s AdWords platform.66 The judgment suggests that to lose the liability exemption provided in the Liability Directive, an ISSP’s activity must cross a certain threshold of active par-

60 See id. paras. 1–3, 5, 8–10; Joined Cases C-236/08 to C-238/08, para. 120.
64 Cf. Sartor & Viola, supra note 5, at 360, 370 (noting that court’s approach “would have a broad impact on providers of platforms for user-generated contents,” and that “they would be in principle liable for all content”).
ticipation in the processing of data. \textsuperscript{67} That is, the ECJ construes the Article 14 exemption broadly, requiring that the ISSP’s participation give it actual knowledge of or control over the content. \textsuperscript{68} By contrast, the Italian court implicitly endorses the view that any activity that is more than mere storage and the simple facilitation of access is de facto active participation, regardless of whether there is knowledge or control. \textsuperscript{69} Such a narrow construction of the liability exemption may reflect the court’s failure to appreciate the complexity of current technology and modern web hosting models. \textsuperscript{70} Although the ECJ’s judgment preceded the release of the opinion by the Italian court, the ECJ case was decided after the conviction had been announced. \textsuperscript{71} Consequently, it is difficult to know whether, and to what extent, the ECJ’s judgment played a role in Judge Magi’s decision.

Nevertheless, the Italian court’s decision also contradicts other aspects of the ECJ’s ruling on ISSP exemption from liability. \textsuperscript{72} In classifying Google as an active content provider, the Italian court pointed to advertising revenues and actions taken by Google to encourage users to upload video, and reasoned that Google promoted the absence of content control as an inducement to attract users. \textsuperscript{73} This reasoning conflicts with the ECJ’s determination that liability depends on the relationship of the ISSP to the content, not on the commercial nature of the service, or even whether there is a clear link between keyword advertising and user queries. \textsuperscript{74} By construing the liability exemption so


\textsuperscript{68} See id.

\textsuperscript{69} See Sartor & Viola, supra note 5, at 360–71.

\textsuperscript{70} See id. at 360–70. When the liability exemptions were drafted, web hosting was a far simpler enterprise than it is today with the advent of Web 2.0 and the proliferation of user-generated content. Id. at 360–70, 374–75; see also Tim O’Reilly & John Battelle, Web Squared: Web 2.0 Five Years On 2, available at http://assets.en.oreilly.com/1/event/28/web2009_websquared-whitepaper.pdf [hereinafter Web 2.0 White Paper] (“The Web is no longer a collection of static pages of HTML that describe something in the world. Increasingly, the Web is the world—everything and everyone in the world casts an ‘information shadow,’ an aura of data which . . . offers extraordinary opportunity and mind-bending implications.”).

\textsuperscript{71} See Joined Cases C-236/08 to C-238/08, para. 109, 116–117; Sartor & Viola, supra note 5, at 371. The ECJ explicitly provided that the “mere fact” that a service provided by the ISSP was in exchange for payment does not deprive the ISSP of the liability exemptions under the Directive. Joined Cases C-236/08 to C-238/08, para. 116.
narrowly, the Italian court excludes ISSPs like Google, YouTube, and Facebook, such that operating sites featuring user-generated content may be far more difficult or even impossible.\textsuperscript{75}

The Italian decision is emblematic of a broader trend in which courts have emerged as a source of Internet regulation.\textsuperscript{76} Given the novel, transnational, and ever-evolving nature of the Web, questions of Internet governance are difficult and controversial; indeed, a major theme in the ongoing debate is whether it ought to be formally regulated at all.\textsuperscript{77} Additionally, despite the Internet’s growing ubiquity, questions of Internet governance are still embryonic in the popular consciousness.\textsuperscript{78} Although many people have cultivated Second Lives online,\textsuperscript{79} far fewer are likely to have considered what rules should govern the Internet “world.”\textsuperscript{80} Courts likely reflect this pattern: judges almost certainly use the Internet for personal purposes, and some even cite Internet sources in judicial opinions.\textsuperscript{81} But ordinary or even regular use of a thing does not an expert make.\textsuperscript{82}

Beyond the question of judicial familiarity with technology, however, the potential impact of governments or private interests seeking to

\textsuperscript{75} See Donadio, \textit{supra} note 2; Arthur, \textit{supra} note 3; Claburn, \textit{supra} note 2.


\textsuperscript{77} See \textit{e.g.}, \textit{Jack Goldsmith & Tim Wu}, \textit{Who Controls the Internet?} 179–81 (2006) (noting that despite its transformative potential as a global communication tool, the Internet cannot function effectively without underlying law and territorial government); \textit{Jonathan Zittrain}, \textit{The Future of the Internet—And How to Stop It} 127–30 (2008) (describing the merits of a standard-based approach to internet governance as compared to a rule-based approach).


\textsuperscript{80} See \textit{The Consensus Machine}, \textit{supra} note 78.


control or limit the flow of information is another major area of concern. Some commentators have suggested that the Italian decision may be an example of such influence; rumors abound that efforts to control new media enterprises like Google and YouTube are motivated by a desire to protect the business interests of media magnate and Italian Prime Minister Silvio Berlusconi.

The upshot is that this scenario results in inapt decisions, and it permits important questions of policy to be decided on an ad hoc basis by inexpert and possibly even biased judges. Inevitably, this pattern makes the rule of law unpredictable, and as ISSPs are forced to bear the burden of this uncertainty, many will attempt to limit the risk of liability by avoiding certain markets. Indeed, some Internet companies have declined to operate in up to half of the world’s fifty largest economies out of concern for legal uncertainty. If, as some say, the promise of the Internet lies in its ability to empower people by democratizing access to information and eliminating barriers to communication, a legal environment that inhibits Internet development through unpredictability threatens to retard the realization of that promise.

Conclusion

The conviction of the Google executives is a recent example of a ruling that contradicts existing laws and precedent governing ISSPs and the Internet generally. It reveals serious limitations in the ability of the existing regulatory edifice to address the complex and novel questions presented by the evolution of the Internet. Given the dynamic nature of Internet information and the global reach of Internet companies, legal uncertainty can have significant consequences. The unpredictability of legal outcomes can deter investments in new technologies and markets, limiting the potential benefits of the Internet.

See, e.g., Goldsmith & Wu, supra note 77, at 184 (noting that the influence of the Chinese government is changing the nature of the Internet); Zittrain, supra note 77, at 112–13 (describing the potential benefits non-democratic regimes might derive from exercising control over technology); Bruce Einhorn, How China Controls the Internet, Bloomberg: BusinessWeek (Jan. 13, 2006), http://www.businessweek.com/bwdaily/dnflash/jan2006/nf20060113_6735.db053.htm (arguing that by acceding to Chinese censorship requirements Internet companies facilitate political suppression).


Cf. Goldsmith & Wu, supra note 77, at 142 (noting that the judiciary can be “corrupt or incompetent, and can fail to follow the law or reflect the wishes of the people”).

See id. at 144–45.

See id.


See Goldsmith & Wu, supra note 77, at 144–45.
and rapid nature of technological development, and the comparatively static pace of legislation, it is nearly certain that courts will continue to have a central role in regulating the Internet for the foreseeable future. The likelihood of similarly dissonant rulings from other courts threatens innovation and technological development, which, in turn, limits the fundamental promise of the Internet as a force for good.