

# CONSTRAINTS ON RESTRAINTS: A SIGNAL DETECTION ANALYSIS OF THE USE OF MECHANICAL RESTRAINTS ON ADULT PSYCHIATRIC INPATIENTS

NICHOLAS SCURICH\* RICHARD S. JOHN<sup>†</sup>

## I. INTRODUCTION

Few practices in contemporary psychiatry are as controversial as the use of restraints to restrict bodily movement. Mechanical restraints involve the application of devices, such as leather straps, body nets or camisoles to the wrists, the waist, or even the whole body, to prevent virtually any movement.<sup>1</sup> Interestingly, in almost no other context may a person may be

---

\* Doctoral Candidate, Department of Psychology, University of Southern California.

† Associate Professor, Department of Psychology, University of Southern California.

This research was conducted under the auspice of Professor Elyn Saks while the first author was a mental health law fellow at the Saks Institute for Mental Health Law, Ethics and Policy at the Gould School of Law. Deepest gratitude is owed to Professor Saks for all of her insightful comments and guidance. Additionally, thanks to Dr. John Jimenez for help with data collection and all participants who shared their experiences.

<sup>1</sup> Ann Fenley Minnick et al., *Prevalence and Patterns of Physical Restraint Use in Acute Care Setting*, 28 J. NURSING. ADMIN. 19, 20, 22 (1998). The Health Care Financing Administration (HCFA) defines restraints as a “manual method, physical or mechanical device, material, or equipment that immobilizes or reduces the ability of a patient to move his or her arms, legs, body, or head freely.” 42 C.F.R. § 482.13 (2011). Restraints do come in other forms, for instance, chemical restraints, in which medication (pharmacological drugs) is used to restrict bodily movement through sedation. Some object to the term “chemical restraint,” calling it both a “misnomer” and “pejorative.” Kenneth Tardiff & John Lion, *Seclusion and Restraint*, in ROBERT I. SIMON & KENNETH TARDIFF, *TEXTBOOK OF VIOLENCE ASSESSMENT AND MANAGEMENT* 351, 351 (2008). One source of contention regarding the use of the term “chemical restraints” stems from the lack of clarity as to whether pharmaceuticals are administered to merely sedate a patient or to actually alleviate the pathological state causing the agitation. Glen Currier & Michael Allen, *Physical and Chemical Restraint in Psychiatric Emergency Service*,

shackled to the point of absolute immobility. Indeed, even the most dangerous criminals cannot have their bodily movement completely restricted,<sup>2</sup> and many individuals would not condone this practice on animals. Yet, in the context of psychiatry, restraints are ubiquitous. They are commonly and exclusively applied to persons with mental illness, including geriatrics,<sup>3</sup> adolescents,<sup>4</sup> and adults undergoing psychiatric treatment.<sup>5</sup>

Aside from their apparent inhumane nature, restraints are controversial because they are dangerous.<sup>6</sup> According to rough estimates, anywhere from 50 to 150 deaths per year are directly attributable to the use of mechanical restraints.<sup>7</sup> It is not uncommon for patients to be placed in a su-

51 PSYCHIATRIC SERV. 717, 718 (2000). Not surprisingly, the use of chemical sedatives is inversely related to the use of mechanical restraints. John Chiles et al., *Effects of Clozapine on the Use of Seclusion and Restraint at a State Hospital*, 3 HOSP. & CMTY. PSYCHIATRY 269, 269 (1994).

<sup>2</sup> Elyn R. Saks, *The Use of Mechanical Restraints in Psychiatric Wards*, 95 YALE L.J. 1836, 1837 n.5 (1986).

<sup>3</sup> Lois K. Evans & Neville E. Strumpf, *Tying Down the Elderly: A Review of the Literature on Physical Restraint*, 37 J. AM. GERIATRIC SOC'Y 65, 65–67 (1989); see Elizabeth Betemps et al., *Length of Time Spent in Seclusion and Restraint by Patients at 82 VA Medical Centers*, 43 HOSP. & CMTY. PSYCHIATRY 912, 912–14 (1992). The primary motivation for use of restraints on this sub-population is to prevent self-injury from falling. Daniel Weintraub & Michelle Spurlock, *Change in the Rate of Restraint Use and Falls on a Psychogeriatric Inpatient Unit: Impact on the Health Care Financing Administration's New Restraint and Seclusion Standards for Hospitals*, 25 J. GERIATRIC PSYCHIATRY & NEUROLOGY 91, 91 (2002); See Veronica Morgan, *Hospital Falls: A Persistent Problem*, 75 AM. J. PUB. HEALTH 775, 775–77 (1985); see Richard Lofgreen et al., *Mechanical Restraints on the Medical Wards: Are Protective Devices Safe?*, 79 AM. J. PUB. HEALTH 735, 735 (1989).

<sup>4</sup> Sue Burrell, *Due Process Limitations on the Use of Four-Point Restraints in Juvenile Detention Facilities*, 12 CHILD. LEGAL RTS. J. 12, 12–14 (1991) (arguing that the use of four-point restraints on juveniles is constitutionally and professionally unacceptable); see also Howard Bath, *The Physical Restraint of Children: Is it Therapeutic?*, 64 AM. J. ORTHOPSYCHIATRY 40, 48 (arguing that restraints, as opposed to seclusion, may have some value for children/adolescents because it demonstrates the limits of behavior in an “interactive” and “inclusive” manner); see generally William T. Garrison, *Inpatient Psychiatric Treatment of the Difficult Child: Common Practices and Their Ethical Implications*, 6 CHILD. & YOUTH SERVICES REV. 353, 353–363 (1984).

<sup>5</sup> See Fay Bower et al., *A Synthesis of What We Know About the Use of Physical Restraints and Seclusion with Patients in Psychiatric and Acute Care Settings*, 7 ONLINE J. OF KNOWLEDGE SYNTHESIS FOR NURSING 1, 5–8 (2003) (describing the effects of physical restraints on adult psychiatric patients). The scope of the current article is restricted to adult psychiatric inpatients.

<sup>6</sup> Clarke Ross, *Death by Restraint: Horror Stories Continue, but Best Practices Are Also Being Identified*, BEHAV. HEALTHCARE TOMORROW, Apr. 2009, at 21.

<sup>7</sup> *Id.*; Wanda Mohr & Brian Mohr, *Mechanisms of Injury and Death Proximal to Restraint*

pine position (on their stomach, facedown), as opposed to a prone position (on their back, faceup) while in restraints, even though being placed facedown is more dangerous.<sup>8</sup> Professional literature even encourages the supine position,<sup>9</sup> despite reports of deaths due to positional asphyxiation and neck compression as a result of being placed facedown.<sup>10</sup> Regardless of the position, prolonged immobilization can engender other negative physical consequences, such as apnea and hypotension, and cardiac arrest may occur if patients continue to struggle while in restraints.<sup>11</sup>

The use of restraints can also have negative psychological side effects. Patients universally report feelings of fear, depression, anger, and helplessness while in restraints.<sup>12</sup> These feelings can persist for quite some time after the administration of restraints.<sup>13</sup> Negative emotional reactions can be especially serious for patients who have previously experienced trauma, which is not uncommon for psychiatric inpatients. For example, patients who suffer from post-traumatic stress disorder (PTSD) tend to ex-

---

Use, 6 ARCHIVES PSYCHIATRIC NURSING 285, 288–91 (2000).

<sup>8</sup> David Reily et al., *Patient Restraints Positions in a Psychiatric Inpatient Service*, NURSING TIMES, Jan. 2006, at 42 (noting that one-third of all patients were placed in restraints facedown on the floor); J. Southcott & A. Howard, *Effectiveness and Safety of Restraint and Breakaway Techniques in a Psychiatric Intensive Care Unit*, NURSING STAND. 35, 37 (May 2007).

<sup>9</sup> Tardiff & Lion, *supra* note 1, at 347 (noting that the supine form is considered safer for staff because the patient is less able to bite).

<sup>10</sup> Reily et al., *supra* note 8, at 42; see Morh & Mohr, *supra* note 7, at 288–291 (documenting various ways in which restraints can cause death). One study compared the efficacy of the different restraint positions (i.e., supine vs. prone) in terms of the time required to bring the behavior under control. No statistical differences were found between the two positions, although the researcher suggested there are “clinical” differences and recommended the prone position based on patient feedback. Sharon Aschen, *Restraints: Does Position Make a Difference?*, 16 ISSUES MENTAL HEALTH NURSING 87, 89–91 (1995).

<sup>11</sup> George S. Fidone, *Risks in Physical Restraint*, 39 HOSP. & CMTY. PSYCHIATRY 203, 203 (1988).

<sup>12</sup> Rolf Wynn, *Psychiatric Inpatients’ Experiences with Restraint*, 15 J. FORENSIC PSYCHIATRY & PSYCHOL. 124, 131–33 (2004); Julia Brown & Sharon Tooke, *On the Seclusion of Psychiatric Patients*, 5 SOC. SCI. & MED. 711, 717 (1992); William A. Fisher, *Restraint and Seclusion: A Review of the Literature*, 151 AM. J. PSYCHIATRY 1584, 1587; Kevin Huckshorn, *Re-Designing State Mental Health Policy to Prevent the Use of Seclusion and Restraint*, 33 ADMIN. & POL’Y MENTAL HEALTH & MENTAL HEALTH SERVICES RES. 482, 485 (2006); Cynthia Robbins et al., *Consumers’ Perceptions of Negative Experiences and “Sanctuary Harm” in Psychiatric Settings*, 56 PSYCHIATRIC SERVICES 1134, 1137–38 (2005).

<sup>13</sup> Brown & Tooke, *supra* note 12, at 718; Henry Wadeson & William T. Carpenter, *Impact of the Seclusion Room Experience*, 163 J. NERVOUS & MENTAL DISEASE 318, 320–24 (1976).

hibit significantly greater negative emotional reactivity to being placed in restraints.<sup>14</sup> Similarly, rape survivors report experiencing profound negative emotional reactions to being physically controlled.<sup>15</sup> Anecdotal case studies report that restraints can dredge up repressed memories of physical and sexual abuse.<sup>16</sup> This list is hardly comprehensive, as the psychological implications of restraints have been discussed for well over two centuries.<sup>17</sup>

Beyond the potentially serious consequences of mechanical restraints, a second cause of controversy is the lack of uniformity in their application and use. The frequency with which restraints are used is highly variable. One study found that 51% of all patients were restrained at least once during their hospitalization,<sup>18</sup> whereas others report lower rates,<sup>19</sup> even as low

<sup>14</sup> Christopher Freuch et al., *Patients' Reports of Traumatic or Harmful Experiences Within the Psychiatric Setting*, 56 PSYCHIATRIC SERVICES 1123, 1127 (2005) (reporting a study in which 19% of randomly sampled psychiatric inpatients met the Diagnostic and Statistical Manual of Mental Disorder (DSM) criteria for PTSD). Individuals with PTSD scored higher than non PTSD individuals on self-report measures of trauma, indicating they felt less safe, more fearful, more helpless, and more distressed while in restraints. *Id.* at 1127.

<sup>15</sup> See Freunch et al., *supra* note 14, at 1127 (patients with a history of sexual assault were more likely to report medication being used as a threat or coercive device, inadequate privacy, unwanted sexual advances, and sexual assault by staff members).

<sup>16</sup> See generally Sharyl Smith, *Restraints: Retraumatization for Rape Victims*, 23 J. PSYCHOL. NURSING 23, 23 (1995) (finding, on the basis of several case studies, that rape survivors described the experience of restraints as evoking feelings of helplessness similar to those feelings experienced during sexual assault).

<sup>17</sup> Phillippe Pinel, considered the godfather of psychiatry, discussed the use of restraints in 1794: "[Servants] carry [the madman] to his cell while thwarting his efforts and chain him if he is very dangerous or merely lock him up." Fisher, *supra* note 12, at 1584. Even in 1880, the deleterious effects of restraints were known and discussed. Walter Channing, *The Use of Mechanical Restraint in Insane Hospitals*, 103 BOS. MED. & SURGICAL J. 173, 173-77 (1880).

<sup>18</sup> Paul Phillips & Suhayl Nasr, *Seclusion and Restraint and Prediction of Violence*, 140 AM. J. PSYCHIATRY 229, 230 (1983) (51% of randomly sampled patients experienced restraints); Wadeson & Carpenter, *supra* note 13, at 320 (reporting that 66% of patients experienced seclusion during hospitalization).

<sup>19</sup> Judith Lever et al., *Use of Physical Restraints and Their Relationship to Medication Use in Patients in Four Different Settings*, 10 HUMANE MED. HEALTH CARE 1, 1 (2010) (reporting a rate of 21%); Nancy Ray & Mark Rappaport, *Use of Restraint and Seclusion in Psychiatric Settings in New York State*, 46 PSYCHIATRIC SERVICES 1032, 1034 (reporting an average rate of 33%); Susanne Tilentelo et al., *A Study of the Use of Restraint in a Psychiatric Emergency Room*, 34 HOSP. & CMTY. PSYCHIATRY 164, 164 (1983) (finding 24% of patients had been restrained); Frank Lovoie, *Consent, Involuntary Treatment and the Use of Force in an Urban Emergency Department*, 21 ANNALS EMERGENCY MED. 25, 25 (1992) (finding a rate of 26%); see generally Minnick, *supra* note 1 (finding a rate of 5.8%); D. Macpherson, *Deciding to Re-*

as 1%.<sup>20</sup> The duration of time spent in restraints is also variable, ranging anywhere from minutes to upwards of days.<sup>21</sup> The reasons prompting the use of restraints are similarly variable. For example, the actual occurrence of violence tends not to be the most common motivation for the use of restraints.<sup>22</sup> Instead, threats of violence tend to occasion restraints far more often,<sup>23</sup> suggesting that restraints are not generally used to contain ongoing violence.<sup>24</sup> Less pernicious reasons such as disorientation, agitation, or refusal to follow staff instructions also commonly motivate the use of restraints.<sup>25</sup>

---

*strain Medical Patients*, 38 J. AM. GERIATRICS SOC'Y 516, 518 (1998) (finding a rate just over 50%); Elia Salias & Kristian Whalbeck, *Restraint and Seclusion in Psychiatric Inpatient Wards*, 18 CURRENT OPINION PSYCHIATRY 555, 556 (2005) (noting that the rate is highly variable).

<sup>20</sup> Bruce Way & Steven Banks, *Use of Seclusion and Restraint in Public Psychiatric Hospitals: Patient Characteristics and Facility Effects*, 41 HOSP. & CMTY. PSYCHIATRY 75, 75 (1990) (finding a rate of seclusion in some hospitals as low as 0.4%); Brown & Tooke, *supra* note 12, at 715.

<sup>21</sup> Currier & Allen, *supra* note 1, at 717–18 (providing the most comprehensively based estimate of the average duration spent in restraints, which is 3.3 hours); Robert Okin, *Variation Among State Hospitals in the Use of Seclusion and Restraint*, 36 HOSP. & CMTY. PSYCHIATRY 648, 649 (1985) (the average durations varied between 4.9 and eighteen hours).

<sup>22</sup> R. Kaltiala-Heino et al., *Reasons for Using Seclusion and Restraint in Psychiatric Inpatient Care*, 26 INT'L J. L. & PSYCHIATRY 139, 144 (2003) (finding that the use of restraints was motivated by the occurrence of violence in 11% of instances); Selim El-Bardi & Graham Mell-sop, *A Study of the Use of Seclusion in an Acute Psychiatric Service*, 36 AUSTL. & N.Z. J. PSYCHIATRY 399, 400–02 (2002); *but see* Way & Banks, *supra* note 20, at 386 (finding that assault on staff was the most common motivation for the use of restraints).

<sup>23</sup> El-Bardi & Mell-sop, *supra* note 22, at 400–02; Phillips & Nasr, *supra* note 18, at 230; Telintelo, *supra* note 19, at 164–65 (reporting that behaviors indicative of violence were the most common percipient of restraints).

<sup>24</sup> It is possible that restraints are used equally as often in response to threats of violence as they were to contain ongoing violence, but that threats of violence are more common than actual violence, hence the actual frequency of restraint use in response to threats is greater than actual violence. *See supra* note 23.

<sup>25</sup> Kirk Heilbrun et al., *Physical Control of Patients on an Inpatient Setting: Civil vs. Forensic Populations*, 66 PSYCHIATRIC Q. 133, 136 (1995) (finding that agitation is one of the leading predictors for the use of restraints); *cf.* Leslie Zun & LaVonne Downey, *Level of Agitation of Psychiatric Patients Presenting to an Emergency Department*, 15 PRIMARY CARE COMPANION J. CLINICAL PSYCHIATRY 108, 109–11 (2008); Miriam Sheridan, *Precipitants of Violence in a Psychiatric Inpatient Setting*, 41 HOSP. & CMTY. PSYCHIATRY 776, 779 (1990) (finding that “conflicts with staff” is the best predictor for the use of restraints); Paul Soloff, *Behavioral Precipitants of Restraint in the Modern Milieu*, 19 COMPREHENSIVE PSYCHIATRY 179, 182 (1978) (arguing that “violation of community and administrative limits” is the best predictor for use of restraints); *see* Fisher, *supra* note 12, at 1587 (suggesting disorientation often leads to the use of restraints).

A number of studies have examined the correlates of restraint use.<sup>26</sup> These studies, as a whole, suggest that neither clinical nor demographic variables are consistently related to the use of restraints. Some studies find that age is related to the use of restraints, with younger patients restrained more frequently,<sup>27</sup> whereas other studies do not find any such association.<sup>28</sup> Similarly, certain mental disorders, especially those exhibiting active symptoms, such as psychoses, are related to restraint use in some studies but not in others.<sup>29</sup> These findings are consistent with the results of controlled experiments, which reveal substantial disagreement among physicians about the causes and recommendations for the use of restraints.<sup>30</sup> Among the more consistent findings is a fairly large "facility effect," in which certain facilities are considerably more or less likely to use restraints.<sup>31</sup> This effect suggests that the observed variability can be par-

---

<sup>26</sup> E.g., Barbara Berland et al., *Patient Characteristics Associated with the Use of Mechanical Restraints*, 5 J. GEN. INTERNAL MED. 480, 483, 485 (1990) (finding that disruptive behavior, risk of falling, and cognitive impairment were related to restraint use).

<sup>27</sup> E.g., Way & Banks, *supra* note 20, at 77–78.

<sup>28</sup> See Fisher, *supra* note 12, at 1586 ("[A]ttempts to correlate demographics such as age, gender, and race are so contradictory that clear-cut trends rarely emerge.").

<sup>29</sup> *Id.* at 1586–87.

<sup>30</sup> R. James Holzworth & Celia E. Wills, *A Nurses' Judgments Regarding Seclusion and Restraint of Psychiatric Patients: A Social Judgment Analysis*, 22 RESEARCH NURSING & HEALTH 189, 189–99 (1999) (providing the results of a qualitative study that found the potential for harm and the unsuccessful search for alternatives both influence nurses' decisions to use restraints); *but cf.* Sandy Marangos-Frost & Donna Wells, *Psychiatric Nurses' Thoughts and Feelings about Restraint Use: A Decision Dilemma*, 31 J. ADVANCED NURSING 362, 365–67 (2000) (showing that nurses who apply restraints feel conflicted since they must choose between physically restraining patients and allowing them to potentially harm themselves or others). The results of these studies must be accepted cautiously however, as the methodology is highly susceptible to social desirability effects (i.e., providing the most socially desirable response). Interestingly, one study found that male staff members are more likely to believe restraint is a negative experience (for patients), whereas female staff members are more likely to believe restraints are a positive experience. Valerie Klinge, *Staff Opinions About Seclusion and Restraint at a State Forensic Hospital*, 45 HOSP. & CMTY. PSYCHIATRY 138, 140–41 (1994).

<sup>31</sup> See Way & Banks, *supra* note 20, at 79–80; *cf.* Sandra Forquer et al., *Predictors of the Use of Restraint and Seclusion in Public Psychiatric Hospitals*, 23 ADMIN. & POL'Y MENTAL HEALTH 527, 529–31 (1996) (noting that some hospitals are more likely to use restraints or seclusion); Elizabeth Betemps et al., *Hospital Characteristics, Diagnoses, and Staff Reasons Associated with Use of Seclusion and Restraint*, 44 HOSP. & CMTY. PSYCHIATRY 367, 367–71 (1993) (noting hospitals in certain regions of the United States are more likely to use restraints).

There is a potential selection effect here—what is known in psychology as the "third variable problem." BRETT W. PELHAM & HART BLANTON, CONDUCTING RESEARCH IN PSYCHOLOGY 62–63 (Rob Hugel et al. Eds., 3rd ed. 2007). The fact that certain facilities have a

tially explained by disparate philosophies about restraint use.

While some advocates consider restraints to be a barbaric practice, those in charge of administering restraints deem them to be a necessary evil and unavoidable aspect of their profession. They argue that restraints provide safety for staff and patients, promote milieu on the ward, and calm belligerent patients.<sup>32</sup> Although some would prefer to use seclusion over restraints,<sup>33</sup> overcrowding and a lack of resources (i.e., seclusion rooms) often preclude this option. There are also some instances in which no other alternative can effectively avert violence. It should be noted that the use of restraints as a punitive or deterrent device is explicitly prohibited.<sup>34</sup>

---

higher rate of restraint use may not be surprising since certain facilities, because of referral policies, practices, resource availability, etc., tend to encounter more violent patients which might in turn necessitate a higher rate of restraint use. Cf. Ben Burstein, *Using Mechanical Restraints on Acutely Disturbed Psychiatric Patients*, 26 HOSP. & CMTY. PSYCHIATRY 757, 757 (1975) ("Frequently nonrestraint is achieved at the expense of patient selectivity; more violent patients may be sent to other—often state—hospitals."). To demonstrate a true "facility effect" it would be necessary to randomly assign patients to facilities and then compare the rate of restraint use, or at the very least, statistically control for the danger level presented by patients at each facility.

<sup>32</sup> Leslie Stilling, *The Pros and Cons of Physical Restraint & Behavior Controls*, 30 J. PSYCOSOCIAL NURSING 18, 18–20 (1992) (discussing the justifications of restraints); Nancy Ray et al., *Patient Perspectives on Restraint and Seclusion Experiences: A Survey of Former Patients of New York State Psychiatric Facilities*, 20 J. PSYCHIATRIC REHABILITATION 11, 14 (1996) (noting that patient protection, calming patients, teaching patients "internal controls" and preserving milieu are reasons for the use of restraints); Harold Mazzarins et al., *Utilizing Mechanical Restraints*, 10 CHILD & YOUTH SERVICES 153, 153–54 (1988) (restraints are used to prevent injury to self or others).

<sup>33</sup> Seclusion refers to the "involuntary confinement of a [person] alone in a room, which the [person] is physically prevented from leaving, for any period of time." JOINT COMM' ON ACCREDITATION FOR HEALTHCARE STANDARDS, STANDARDS FOR BEHAVIOR HEALTH CARE 44 (2004).

<sup>34</sup> Most governing statutes explicitly state this. California's Code of Regulation 22 CCR § 77103(c) is representative: "Behavioral restraint and seclusion shall not be used as punishment . . ." CAL. CODE REGS. tit. 22, § 77103(c) (2011). The American Psychiatric Association admonishes that hospital personnel must never use restraints for punitive purposes or staff convenience. Tardiff & Lion, *supra* note 1, at 352. There is some ambiguity, however, as to whether *punishment* is used in a "behavioral therapy" sense or in a punitive sense. See David Wexler, LEGAL ASPECTS OF SECLUSION AND RESTRAINT, in THE PSYCHIATRIC USES OF SECLUSION AND RESTRAINT 122, 122 (Kenneth Tardiff ed., 1985). Even in the former instance, in which restraints are used to modify behavior, they seem generally impermissible without informed consent. However, a central tenet of informed consent is that it is revocable at any time, so the feasibility of this justification seems limited. *Id.* at 118. An additional end run is to claim that behavior modification is necessary to prevent dangerous or violent behavior. Thus, restraints for the purpose of behavior modification may be permissible if they reduce the likelihood of dangerous behavior. *Id.* at 120. But since arguably all forms of treatment are directly or indirectly geared toward reducing danger, this distinction is disingenuous.

The necessary evil argument has not escaped criticism. Following the so-called "British experience,"<sup>35</sup> a movement in the United Kingdom to eschew the use of restraints in psychiatric practice, restraints are almost never used, yet the prevalence of violence is not greater in British psychiatric wards.<sup>36</sup> Generalizability issues notwithstanding,<sup>37</sup> the data seem to indicate that restraints are not necessary to achieve their safety or milieu objectives.<sup>38</sup>

These findings are consistent with a burgeoning literature in the United States on alternatives to restraints. For instance, one study found that a simple training program, in which clinicians experience restraints first hand, substantially reduced the use of restraints without a corresponding increase in the occurrences of violence.<sup>39</sup> Another study found that simply

<sup>35</sup> Saks, *supra* note 2, at 1843; cf. Elias Cohen & Ann Kruschwitz, *Restraint Reduction: Lessons from the Asylum*, 3 J. ETHICS L., & AGING 25, 30 (1997) (calling it the "European experience"). The European orientation towards treatment of the mentally ill immigrated to the United States in the late 1960's, but it was not until seminal court cases in the 1980s that the use of restraints on psychiatric populations came to the public's awareness. See Bower et al., *supra* note 5, at 2.

<sup>36</sup> Saks, *supra* note 2, at 1845-47.

<sup>37</sup> "Generalizability" refers to the degree to which the findings would apply, or *generalize*, to other populations, such as American psychiatric patients. One issue that potentially questions the generalizability of the British findings is that American psychiatric populations tend to have different prevalence rates of certain mental disorders that predispose patients to injurious behavior. For example, concomitant substance abuse disorders are among the mental disorders with the highest correlation to violence. JOHN MONAHAN ET AL., *RETHINKING RISK ASSESSMENT THE MACARTHUR STUDY OF MENTAL DISORDER AND VIOLENCE* 64 (2001); cf. Michael A. Norko, *Dangerousness—A Failed Paradigm for Clinical Practice and Service Delivery*, 28 J. AM. ACAD. PSYCHIATRY L. 282, 287 (2000) (citing a study that proved there is no statistically significant correlation between mental illness and violence in the absence of substance abuse). The prevalence of such disorders is about three times greater amongst American psychiatric patients than their British counterparts. Michael Soyka, *Substance Misuse, Psychiatric Disorder and Violent and Disturbed Behavior*, 176 BRIT. J. PSYCHIATRY 345, 345, 347 (2000). Thus, the observed lack of violence following the eschewal of restraints could be an artifact stemming from the relative lack of violence to begin with. This argument is supported by empirical studies from western European countries. Len Bowers et al., *International Variation in Containment Measures for Disturbed Psychiatric Inpatients: A Comparative Survey*, 44 INT'L J. NURSING STUD. 357, 357-58, 360-64 (2005). It is also possible that the use of restraints could *cause* (or increase) the occurrence of violence, perhaps because they frighten patients, but this is an empirical question for which there is currently no data.

<sup>38</sup> See Harvey Gordon et al., *The Use of Mechanical Restraints in the Management of Psychiatric Patients: Is it Ever Appropriate?*, 10 J. FORENSIC PSYCHIATRY 173, 177, 180 (1999) (discussing the British experience and arguing that there are instances in which restraints may prove "helpful").

<sup>39</sup> Peter L. Forster et al., *Staff Training Decreases Use of Seclusion and Restraint in an*

increasing the number of staff decreased the use of restraints.<sup>40</sup> And another study found that fostering a collaborative relationship between staff and patients decreased the use of restraints without an increase in violence.<sup>41</sup> These findings, in conjunction with the British experience, cast doubt on the necessity of the use of restraints. The fact that restraint use can be decreased without a corresponding increase in violence implies that restraints are unnecessary in all but perhaps the most extreme cases. Indeed, this is the main argument levied by advocates who call for the abolition of restraints in the United States.

From a methodological standpoint, however, most of the available research is equivocal. In general the research tends to be correlational,

---

*Acute Psychiatry Hospital*, 8 ARCHIVES PSYCHIATRY NURSING 269, 270 (1999) ("Each staff member experienced 5-point restraints first-hand, and many cited that experience as pivotal in their decision whether or not to restrain a patient in a state of agitation when queried 1 year after the course."); Robert E. McCue, *Reducing Restraint in a Public Psychiatric Inpatient Service*, J. BEHAV. HEALTH SERV. RESEARCH 217, 221–23 (2004); Maureen Lewis et al., *Crisis Prevention Management: A Program to Reduce the Use of Seclusion and Restraint in an Inpatient Mental Health Setting*, ISSUES MENTAL HEALTH NURSING 159, 162–63 (2009) (finding that a crisis prevention management program decreased the use of seclusion restraint); cf. Dennis Donat, *An Analysis of Successful Efforts to Reduce the Use of Seclusion and Restraint at a Public Psychiatric Hospital*, 54 PSYCHIATRIC SERVICES, 1119, 1119–23 (2003) (demonstrating various other methods that can be used to reduce the use of seclusion and restraints without increasing violence); cf. David J. Hellerstein et al., *Decreasing the Use of Restraint and Seclusion Among Psychiatric Inpatients*, 13 J. PSYCHIATRIC PRAC. 308, 309, 311–15 (2007) (finding that training programs and certain management structures reduced the use of seclusion and restraints without increasing violence); Sandra A. Barton et al., *Achieving Restraint-Free on an Inpatient Behavioral Health Unit* 47, 47 J. PSYCHOSOCIAL NURSING 34, 39 (2009) (finding that training programs practically eliminated the use of mechanical restraints); Jessica A. Jonikas et al., *A Program to Reduce use of Physical Restraint in Psychiatric Inpatient Facilities*, 55 PSYCHIATRIC SERVICES, 818, 818–20 (2004) (demonstrating that advance crisis management counseling and a training program to teach patients how to manage their anger reduced the need for restraints).

<sup>40</sup> Dennis Donat, *Impact of Improved Staffing on Seclusion/Restraint Reliance in a Public Psychiatric Hospital*, 25 PSYCHIATRIC REHABILITATION J. 413, 413–15 (2002). In fact, other studies have found that the number of available staff members is inversely related to the rate of restraint use. Kathryn Brooks et al., *Patient Overcrowding in Psychiatric Hospital Units: Effects on Seclusion and Restraint*, 22 ADMIN. & POL'Y. MENTAL HEALTH 133, 135, 138–43 (1994) (finding a positive correlation between restraint use and overcrowding); Jose de Cangas, *Nursing Staff and Unit Characteristics: Do They Affect the Use of Seclusion?*, 29 PERSP. CARE 15, 21 (1993) (finding that the number of seclusion hours on a unit increased as the number of available staff decreased).

<sup>41</sup> Ann Sullivan et al., *Reducing Restraints: Alternatives to Restraints on an Inpatient Psychiatric Service—Utilizing Safe and Effective Methods to Evaluate and Treat the Violent Patient*, 76 PSYCHIATRY Q. 51, 59–61, 62 (2005).

which limits the ability to make casual claims from the data. Furthermore, the studies likely suffer from “observer/expectancy effects,” which might cause both staff and participants to behave differently given their awareness of the study.<sup>42</sup> These effects are particularly likely to occur if a staff member developed the tested alternative, since she might have a vested interest in its success.<sup>43</sup> There are also reporting issues related to observer effects. Most of the studies compare rates of restraint use and violence occurrence rates, both of which are typically ascertained by official hospital records. But no study has examined whether the reporting habits are consistent following the implementation of the alternative approach. These methodological issues potentially limit the internal validity—that is, the ability to make causal statements about the efficacy of the restraint alternatives. None of the proposed alternatives have been subjected to randomized clinical trials,<sup>44</sup> the gold standard for determining efficacy in the medical field.<sup>45</sup> Hence, the empirical findings should be considered preliminary rather than definitive proof that restraints are superfluous.<sup>46</sup>

---

<sup>42</sup> For an example in which observer effects could have influenced the results, see Hank Visalli et al., *Reducing High-Risk Interventions for Managing Aggression in Psychiatric Settings*, 11 J. NURSING CARE QUALITY 54, 55 (1997), which found a reduction in the use of restraints and violence following the implementation of a campaign to implement less-restrictive measures. It could be the case, however, that staff members habitually overused restraints, but, following the campaign, were simply less apt to use restraints (for reasons unrelated to the specifics of the campaign). Hence, one would expect that restraints to be used less frequently without a corresponding increase in violence because restraints were previously used unnecessarily.

<sup>43</sup> This is not necessarily chicanery. Many cognitive biases arise when looking to confirm a hypothesis. For example, confirmation bias is the tendency to seek evidence that is consistent with a hypothesis and discount evidence that is inconsistent with it. See Raymond S. Nickerson, *Confirmation Bias: A Ubiquitous Phenomenon in Many Guises*, 2 REV. GEN. PSYCHOL. 175, 175–211 (1998). Importantly, the confirmation bias typically occurs outside the realm of consciousness, hence it should not be interpreted as intentional deception. *Id.* at 202.

<sup>44</sup> See generally Eila Sailas & Mark Fenton, *Seclusion and Restraint for People with Serious Mental Illnesses (Review)*, COCHRANE COLLABORATION 1, 7 (2009) (“There is a surprising and shocking lack of published trials assessing the effects of secluding or restraining people with schizophrenia or similar psychotic illnesses.”).

<sup>45</sup> Jan Bergk et al., *Feasibility of Randomized Controlled Trials on Seclusion and Mechanical Restraint*, 5 CLINICAL TRIALS 356, 357–59 (2009) (discussing, with approval, the ethical, legal, and methodological issues associated with testing the efficacy of restraint procedures); see also Tardiff & Lion *supra* note 1, at 341.

<sup>46</sup> One might still argue in favor of the proposed alternatives on the basis that the use of restraints merely contains behavior whereas the alternatives might help to modify behavior. *E.g.*, Elaine Walsh & Brooke Randell, *Seclusion and Restraint: What We Need to Know*, 8 J. CHILD & ADOLESCENT PSYCHIATRIC NURSING 28, 33–34 (1995). However, one should bear in mind that the expressed sole purpose of the use of restraints is to contain rather than modify behavior.

The debate over the propriety of the use of restraints has reached a stalemate. Advocates, relying on empirical evidence, claim that restraints are unnecessary, and professionals, pursuant to standard care, continue to use restraints. This Comment suggests that a satisfactory resolution might lie somewhere in between these poles. Very few advocates would argue that the use of restraints is never permissible. Instead, it seems that advocates argue that the *de facto* use of restraints is arbitrary, excessive, and ultimately abusive. Insofar as this argument is valid, this Comment suggests that the current state of affairs is not solely attributable to the professionals who use restraints in practice. The current situation is at least partially the result of the case law governing the use of restraints, which, as is discussed in the next section, grants considerable deference to professionals to make decisions regarding the use of restraints. The law only requires “danger” in order to use restraints. It does not articulate the requisite level or type of danger that would justify the use of restraints. As a result, there are no means by which to evaluate the decision to use restraints, nor any real accountability for restraint use. Delineating an acceptable threshold level of danger that legitimates the use of restraints would help overcome both these limitations. The purpose of this paper is to explore this threshold.

This paper proceeds in five sections. Part II describes the case law pertaining to the use of restraints, which is both circular and empty, and which motivates the formal analysis contained in this paper. Part III explains how the decision to use restraints is fundamentally a diagnostic decision. This decision, like any diagnostic decision, inherently turns on a tradeoff between two types of errors: false positives (i.e., applying restraints when not required) and false negatives (i.e., failing to apply restraints when required). This section presents the results of an empirical study that elicited judgments from two different groups of stakeholders, practicing clinicians and former psychiatric patients, regarding the appropriate tradeoff. Not surprisingly, the judgments were radically disparate, indicating that different stakeholders have very different beliefs about when the use of restraints is legitimate. Part IV considers the normative implications for the use of restraints. Because the judgments of both stakeholder groups are relevant but conflicting, a normative threshold ought to be an amalgamation of all views. This section describes how each

---

Use of restraints to modify behavior is generally impermissible. *See supra* note 34.

stakeholder's tradeoff can be differentially weighted to yield a single, normative threshold for when the use of restraints is legitimate. Part V provides a substantive discussion of the findings and their practical implications for the law.

## II. CASE LAW ON THE USE OF RESTRAINTS

The U.S. Supreme Court addressed the issue of the use of mechanical restraints in *Youngberg v. Romeo*.<sup>47</sup> The *Youngberg* Court held that the Due Process Clause of the Fourteenth Amendment established liberty interests to reasonably safe conditions of confinement and freedom from unreasonable bodily restraints.<sup>48</sup> The Court noted that there are situations in which restraints are permissible, and stated that the relevant query is "not simply whether a liberty interest has been infringed but whether the extent or nature of the restraint or lack of absolute safety is such as to violate due process."<sup>49</sup> The standard for evaluating the use of restraints requires only that professional judgment be exercised,<sup>50</sup> based on the rationale that the judiciary is not more qualified to make such decisions.<sup>51</sup> Thus, the Court appears to show great deference to "professionals" and their judgment.<sup>52</sup> Notably, the Court defined a professional as "a person competent, whether by education, training or experience, to make the particular decision at issue."<sup>53</sup> The Court explicitly left the responsibility with lower courts to de-

---

<sup>47</sup> 457 U.S. 307 (1982).

<sup>48</sup> *Id.* at 323-25.

<sup>49</sup> *Id.* at 320.

<sup>50</sup> See *id.* at 323 ("[setting forth a test that establishes liability when there is] such a substantial departure from accepted professional judgment, practice or standards as to demonstrate that the person responsible actually did not base the decision on such a judgment.").

<sup>51</sup> *Id.* at 321.

<sup>52</sup> This deference substantially forecloses any liability claims that could be brought by patients for improper use of restraints. As the Court went on to state, "[T]he decision, if made by a professional, is presumptively valid; liability may be imposed only when the decision by the professional is such a substantial departure from accepted professional judgment, practice, or standards as to demonstrate that the person responsible actually did not base the decision on such a judgment." *Id.* at 323. This reasoning is either tautological in that the decision is correct by virtue of the fact that a professional made it, or incoherent since the judgment made by a professional is necessarily a professional judgment. Following this holding, lower courts have been extremely reluctant to impose liability, even in the face of negligence. See *Feagley v. Waddill*, 868 F.2d 1437, 1441-43 (5th Cir. 1989); *Gilbert v. Tex. Mental Health & Mental Retardation*, 919 F. Supp. 1031, 1038, 1042 (E.D. Tex. 1996).

<sup>53</sup> *Youngberg*, 457 U.S. at 323 n.30. This liberal definition of a "professional" is at odds

termine whether adequate professional judgment was exercised.<sup>54</sup>

Statutes governing the use of restraints are quite similar among jurisdictions. Almost all jurisdictions approve the use of restraints, provided they are the “least restrictive alternative” that will adequately contain the danger.<sup>55</sup> Some statutes are more stringent in that restraints may be applied only after less restrictive alternatives have failed,<sup>56</sup> but the statutes do not specify which modalities are the “least” restrictive, and there is not consensus on this issue.<sup>57</sup>

The statutes are also similar in the functional language they use to describe the necessary conditions for restraint use. Connecticut’s statute is fairly representative, “No patient may be placed involuntarily in seclusion

with the definition proposed by certain regulatory committees. See Coffin *infra* note 54, at 598–99.

<sup>54</sup> Christopher Coffin, *Case Law and Clinical Considerations Involving Physical Restraint and Seclusion for Institutionalized Persons with Mental Disabilities*, 23 MENTAL & PHYSICAL DISABILITY L. REP. 597, 598–99 (1999) (the “professional judgment standard” opens many more questions than it answers). For an illuminating discussion, see Susan Stefan, *What Constitutes Departure from Professional Judgment?*, 17 MENTAL & PHYSICAL DISABILITY L. REP. 207 (1993), which discusses the conditions under which professional judgment was not adequately exercised, and noting the substantive questions that follow from such departures.

<sup>55</sup> ALASKA STAT. § 47.30.825(d) (2004) (Alaska); ARIZ. ADMIN. CODE § R9-21-204 (2003) (Arizona); CAL. HEALTH & SAFETY CODE § 1180.4(j) (2003) (California); GA. COMP. R. & REGS. 290-4-6-.02(1)(c)(1) (1991) (Georgia); COLO. REV. STAT. § 26-20-103(1)(b)(I) (2010) (Colorado); 440 IND. ADMIN. CODE 1.5-3-13(4) (2008) (Indiana); KAN. STAT. ANN. § 59-2977(a) (1996) (Kansas); KY. REV. STAT. ANN. § 202A.241 (West 2005) (Kentucky); LA. REV. STAT. ANN. § 28:171(D)(1) (2006) (Louisiana); MICH. COMP. LAWS ANN. § 330.1740(2) (1996) (Michigan); MO. REV. STAT. § 630.175 (2008) (Missouri); MONT. CODE ANN. § 53-21-146 (2009) (Montana); 175 NEB. ADMIN. R. & REGS. Ch. 19 § 006 (2011) (Nebraska); N.J. Stat. Ann. § 30:4-27.11e(a)(3) (West 1991) (New Jersey); N.Y. MENTAL HYG. LAW § 33.04 (McKinney 1980) (New York); R.I. Gen. Laws § 42-72.9-3(b) (2000) (Rhode Island); S.D. CODIFIED LAWS § 27A-12-6.2 (1991) (South Dakota); TENN. COMP. R. & REGS. 0940-03-09-.07(1) (2008) (Tennessee); VA. CODE ANN. § 37.2-400(A)(6) (2009) (Virginia); WASH. ADMIN. CODE § 246-337-110(1) (2005) (Washington); W. VA. CODE. R. § 64-59-10.13 (2011) (West Virginia); WYO. R. & REGS. HEALTH STATE-PATIENT BILL OF RIGHTS Ch. 1 § 4 (1990) (Wyoming).

<sup>56</sup> E.g., COLO. REV. STAT. § 26-20-103(1)(b)(I) (2010) (“[S]ubject to the provision of this article, an agency may only use restraint after the failure of less restrictive alternatives.”).

<sup>57</sup> Yvette Sheline & Teresa Nelson, *Patient Choice: Deciding Between Psychotropic Medication and Physical Restraints in an Emergency*, 21 BULL. AM. ACAD. PSYCHIATRY & L. 321, 321–322, 328 (1993). However, one study has found some agreement on the relative degree of restrictiveness amongst the various forms of restraints. Grant Harris et al., *Staff and Patient Perceptions of the Least Restrictive Alternatives for the Short-term Control of Disturbed Behavior*, 17 J. PSYCHIATRY & L. 239, 259 (1989) (“[M]echanical restraint combined with constant observation was clearly the most intrusive component, followed in descending order of intrusiveness by seclusion, sedative injection, loss of clothing, a sedative pill and manual restraint.”).

or a mechanical restraint unless necessary because there is imminent physical danger to the patient or others and a physician so orders.”<sup>58</sup> There are, however, some slight variations in the statutes with regard to the temporal adjective describing the danger. Connecticut requires “imminent danger,”<sup>59</sup> Illinois requires “immediate danger,”<sup>60</sup> and some statutes do not specify a temporal property at all, simply requiring “danger.”<sup>61</sup>

Nearly all statutes are remarkably vague with respect to defining “danger,” or the types of acts that would satisfy the statutory definition of danger. Connecticut, for example, only states that the danger must be a “physical” danger.<sup>62</sup> Other, more elaborate definitions typically either beg the question or are largely circular. For instance, Louisiana statutorily defines danger as “clinically significant threats or attempts to commit suicide or to inflict serious harm to self, or a substantial risk or serious physical assault on another person, as evidenced by dangerous actions or clinically significant threats.”<sup>63</sup>

In sum, the *Youngberg* Court recognized that patients have a fundamental right to be free from restraint except under specific conditions, and, at the same time, the Court vested professionals with the authority to determine these conditions. The *Youngberg* logic is self-defeating for at least two reasons. The first stems from relegating this determination to professionals, who are clearly not disinterested. A restrain-first, ask-questions-later, better-safe-than-sorry philosophy is reasonable from the clinician’s perspective because it can effectively avert violence. However, this policy will likely result in numerous individuals being gratuitously restrained, which will do little to ensure that the promulgated liberty interests are protected. The second problem is the Court’s deferential posture toward the professionals. As currently formulated, professional judgment cannot be

<sup>58</sup> CONN. GEN. STAT. § 17a-544(a) (1991).

<sup>59</sup> *Id.*

<sup>60</sup> 405 ILL. COMP. STAT. 5/2-108(f) (2002).

<sup>61</sup> Indiana’s statute, for example, approves the use of restraints to “prevent danger of abuse or injury to the patient or other patients.” IND. CODE § 12-27-4-1(1) (2007). Notice there is no specified time frame for when the abuse or injury is forecasted to occur. An indefinite time frame is incoherent since virtually everyone will be dangerous at some point in life.

<sup>62</sup> CONN. GEN. STAT. § 17a-544(a) (1991) (“[N]o patient may be placed involuntarily in seclusion or a mechanical restraint unless necessary because there is imminent physical danger to the patient or others and a physician so orders.”); *cf.* 405 ILL. COMP. STAT. 5/2-108(f) (2002) (“unless there is an immediate danger that the recipient will physically harm himself or others.”).

<sup>63</sup> LA. Rev. Stat. Ann. § 28:171 (D)(2) (West 2011).

second-guessed, provided the judgment is within the realm of “standard professional care.”<sup>64</sup> Aside from the fact that the empirical research paints a confusing picture about what level of care is considered “standard,”<sup>65</sup> the *Youngberg* reasoning is specious. The reasoning is akin to accepting that  $2 + 2 = 50$  simply because a mathematics professor made the assertion. While a math professor might be qualified to make such statements, it is fallacious to assume that qualifications are a proxy for validity.<sup>66</sup> Compounding the issue are the nebulous statutes that govern the use of restraints. The statutes generally authorize the use of restraints to “prevent danger,” but because any non-neonate presents some potential for danger, this requirement is empty.

One way to potentially improve the situation is to reify the concept of “danger” for the purposes of using restraints. As with civil commitment, the fundamental issue is determining the level of danger at which intervention is justified.<sup>67</sup> This level, or *threshold*, should incorporate not only the professional’s desire to prevent violence, but also the patient’s liberty interests. A threshold that delineates the acceptable range of risk that justifies the use of restraints could remediate both shortcomings of *Youngberg*. First, a range of risk deemed to be “acceptable” would inherently incorporate patients’ liberty interests. Second, a delineated range provides a normative standard to evaluate and enforce de facto restraint decisions. Thus, while the use of restraints will continue to be controversial, their application might be less capricious and presumably less deplorable if there were a standardized level of danger when the use of restraints is permissible.

---

<sup>64</sup> *Youngberg v. Romeo*, 457 U.S. 307, 323 (1982).

<sup>65</sup> See *supra* notes 52–57, 62–63 and accompanying text.

<sup>66</sup> This form of reasoning is known as “the fallacy of arguing from authority.” RICHARD POPKIN & AVRUM STROLL, *PHILOSOPHY MADE SIMPLE*, 263 (1993) (“It is not the prestige of an authority that makes a statement true or false but, rather, the citing of evidence to confirm or confute the statement. The fact that an authority has made a statement cannot be *itself* regarded as evidence; what constitutes evidence are the *facts* that the authority produces—and these are quite different from a mere verbal pronouncement.”).

<sup>67</sup> Robert Schopp & Michael Quattrocchi, *Predicting the Present: Expert Testimony and Civil Commitment*, 13 BEHAV. SCI. & L. 159, 164 (1995) (“Every subject, like the attorneys, the witnesses, the judge, and virtually anyone who is not currently in a permanent vegetative state, presents some risk of harm. To avoid rendering the [dangerousness] requirement trivial, it must be understood as referring to some threshold of risk.”). See generally Nicholas Scurich & Richard S. John, *The Normative Threshold for Psychiatric Civil Commitment*, 50 JURIMETRICS J. 425 (2010) (utilizing the same analytical framework to evaluate dangerousness in the context of involuntary civil commitment).

### III. DECISION THEORY

This section explicates the logic underlying decision theory. Decision theory is a method to formally balance costs and benefits in a decision dilemma. One relevant benefit of formalizing the decision dilemma is the transparency of the cost-benefit policy. Transparency permits an evaluation of whether the underlying policy is appropriate, which in turn indicates when the use of restraints is justified.

#### A. THE LOGIC

Decision theory was originally developed by applied mathematicians,<sup>68</sup> but it has become a mainstay within contemporary economics<sup>69</sup> and psychology.<sup>70</sup> The theory relies on a set of axioms, known as coherence axioms,<sup>71</sup> to derive an optimal decision,<sup>72</sup> given a set of preferences.

<sup>68</sup> The axioms underpinning decision theory were first discovered by Frank Ramsey, a philosopher and mathematician, in 1931. FRANK P. RAMSEY, *THE FOUNDATIONS OF MATHEMATICS AND OTHER LOGICAL ESSAYS* 158–98 (R.B. Braithwaite ed., 1931). The theory was elaborated upon by Bruno De Finetti, John von Neumann, Oskar Morgenstern, and Leonard Savage. See generally JOHN VON NEUMANN & OSKAR MORGENSTERN, *THEORY OF GAMES AND ECONOMIC BEHAVIOR* (Sixtieth-Anniversary ed. 2004) (discussing economic behavior and “the application of ‘games of strategy’”); LEONARD SAVAGE, *THE FOUNDATIONS OF STATISTICS* (Walter Shewart ed., 1954) (illustrating the statistical foundations and requirements of subjective probability); BRUNO D. FINETTI, *FORESIGHT: ITS LOGICAL LAWS, ITS SUBJECTIVE SOURCES*, in *STUDIES IN SUBJECTIVE PROBABILITY* 93 (Henry E. Kyburg, Jr. & Howard E. Smokler eds., Henry E. Kyburg, Jr. trans., 1964) (providing an introduction in probability theory).

<sup>69</sup> See Colin Camerer, *Wanting, Liking, and Learning: Neuroscience and Paternalism*, 73 U. CHI. L. REV. 87, 88–93 (2006) (discussing the evolution of decision theory in economics).

<sup>70</sup> Psychology has become infused in contemporary economic thought, in a paradigm that is now called “behavioral economics.” Psychologists who study judgment and decision-making have long used decision theory to evaluate the efficiency of decisions made under uncertainty. The seminal work was conducted by Ward Edwards, a psychologist who pioneered the judgment and decision making (behavioral decision theory) field. Ward Edwards, *The Theory of Decision Making*, 51 PSYCHOL. BULL. 380, 380, 410–11 (1954).

<sup>71</sup> The coherence axioms imply a type of internal consistency in choice. There are three axioms that follow from two general principles known as “weak ordering” and the “sure-thing principle.” Weak ordering has two parts. The first is that choices must be “connected,” such that X can be preferred to Y, or vice versa, or one can be indifferent between X and Y. If X cannot be compared to Y, then the connected axiom is violated. The second axiom of weak ordering is known as “transitivity.” Transitivity holds that if X is preferred to Y and Y is preferred to Z, then X is preferred to Z. The sure-thing principle holds that if the outcome is the same no matter the choice, then the choice should not depend on the outcome. JON BARON, *THINKING AND DECIDING* 233–36 (3d ed. 2000).

These preferences refer to the desirability of the costs and benefits associated with the possible outcomes of the decision. Importantly, the theory does not dictate what the preferences ought to be, which are determined by the decision-maker. Rather the theory states what the optimal choice should be given a specified set of preferences.

The theory itself is not mathematically convoluted, but Figure 1 below can pictorially convey the crux of the logic. Accordingly, all equations will be relegated to footnotes.

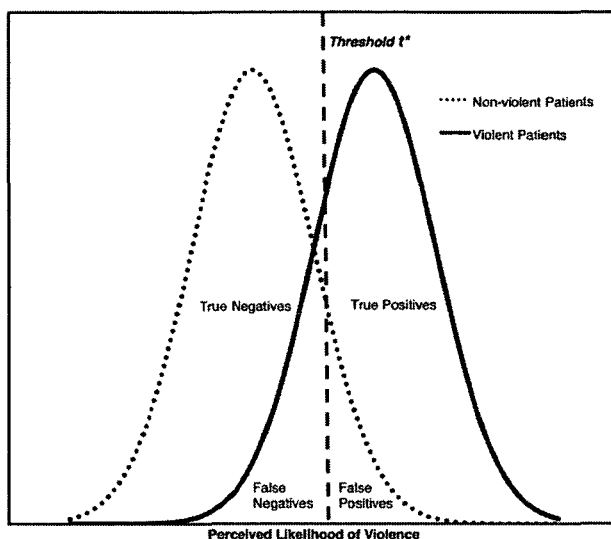


Figure 1 depicts two different groups of patients: the group on the right represents violent patients; and the group on the left represents non-violent patients.<sup>72</sup> The objective is to determine which patients are violent

<sup>72</sup> The term “optimal” is for our purposes synonymous with the term “rational.” Rationality concerns methods of thinking, not necessarily the conclusions that are generated or reached. *Id.* at 55. Hence, calling something “irrational” suggests a disagreement with the thinking process. *Id.* Irrationality, taken to be the antithesis of rationality, in economics and psychology is generally defined as deviating from the axioms of expected utility (EU). *Id.* For example, one might violate transitivity yet still reach the desired outcome. In psychology and economics this would be considered irrational behavior.

<sup>73</sup> The term “violent” will be used in place of the common term “dangerous.” “Dangerous” is an evaluative judgment about the degree of risk posed, which in most contexts depends on normative and legal considerations. See John Monahan & David Wexler, *A Definite Maybe*:

and implement restraints on those patients only. However, since clinicians are neither omniscient nor possess perfect detection capabilities, there is overlap in the two groups, implying that some non-violent patients appear violent and some violent patients appear non-violent. Appearance refers to the perceived likelihood that a given patient will be violent.<sup>74</sup> That is, a personal belief about the likelihood that a given patient will engage in violence.

The horizontal axis of Figure 1 reflects the perceived likelihood that a patient will engage in violence.<sup>75</sup> For the most part, violent patients are perceived to be more likely to engage in violence than non-violent patients. But some violent patients actually appear to have a lower likelihood of violence than some non-violent patients. In short, the perceived likelihood that a patient will engage in violence is an imperfect indicator of whether a patient actually will engage in violence. It is, however, the best indicator that is available, so the decision inevitably boils down to avoiding certain errors.

---

*Proof and Probability in Civil Commitment*, 2 L. & HUM. BEHAV. 37, 37-41 (1978). "Violence," as currently used will be left in the abstract, while fully acknowledging that a formal operational definition is required to give the term concrete meaning. Also, referring to patients as "violent" or "not violent" is not to suggest a sort of determinism about the autonomy of psychiatric patients. Violence is often highly situational or conditional. See Edward Mulvey & Charles Lidz, *Conditional Prediction: A Model for Research on Dangerousness to Others in a New Era*, 18 INT'L J. L. & PSYCHIATRY 129, 131-35 (1995). However, it is much easier to refer to patients as either violent or not (given certain conditions and situations, etc.), and so this language will be used.

<sup>74</sup> This likelihood refers to a personal perception held by the clinician. This conceptualization follows from the subjective interpretation of probability, which regards probability as a degree of belief in a proposition. SAVAGE, *supra* note 68, at 3. This form of probability is different than the frequentist conception of probability, according to which probability is based on observed frequencies in the long run. For example, the probability of a heads in a coin flip is 0.5 because in the long run one would expect to observe a heads half the time. Some events are unique, and therefore do not lend themselves to the frequentist conception of probability. The subjective interpretation is appropriate for unique events. See generally JOHN G. KEMENY, *A PHILOSOPHER LOOKS AT SCIENCE* 65-81 (1959) (discussing the various conceptions of probability). A subjective probability can be based on innumerable factors, including past experience, professional training, etc., and it may or may not be well calibrated, but that is not important for the application of decision theory. Subjective probabilities can span the unit interval of [0-1.0].

<sup>75</sup> It is important to not conflate the term "likelihood" with the term "risk." Risk typically is considered the probability of an event weighted by the magnitude of its outcomes. "Likelihood" is meant to refer strictly to the likelihood of the event occurring (i.e., probability), without regard to the severity of its occurrence (i.e., magnitude).

Signal detection theory (SDT),<sup>76</sup> which is subsumed under decision theory, is used to formally implement the preferred tradeoff between these errors. The two essential components of SDT are conveyed in Figure 1. The first component is indicated by the degree of overlap between the two groups, which directly depends on the ability to discern violent patients from non-violent patients, a property referred to as “diagnosticity.”<sup>77</sup> All else being equal, greater diagnosticity results in less error because violent patients are more discernable from nonviolent patients.

The second component is the decision threshold “ $t^*$ ,” which is used to make binary decisions from continuous information. In this context, the threshold indicates the point at which the perceived likelihood of violence is sufficient to warrant the use of restraints. Patients with a perceived likelihood above this threshold are restrained and patients below this threshold are not restrained. Again, since some non-violent patients appear violent and some violent patients appear non-violent, the binary decision to use restraints fundamentally turns on a tradeoff between restraining non-violent patients and failing to restrain violent patients, or false positives

---

<sup>76</sup> SDT was originally developed in the context of psychophysics, but it has been successfully applied in a variety of contexts such as weather forecasting, medical decision-making, oil exploration, quality control, aviation failure testing, and the law. CLYDE COOMBS ET AL., MATHEMATICAL PSYCHOLOGY: AN ELEMENTARY INTRODUCTION 165–99 (1970); Michael L. DeKay, *The Difference Between Blackstone-Like Error Ratios and Probabilistic Standards of Proof*, 21 L. & SOC. INQUIRY 95, 100 (1996) (using SDT to model legal discussions); John A. Swets et al., *Psychological Science Can Improve Diagnostic Decisions*, 1 PSYCHOL. SCI. IN. PUB. INT. 1, 10–20 (2000) (discussing the various contexts in which SDT has been applied). Dr. Douglas Mossman has repeatedly used SDT to analyze clinical decisions about dangerousness. See, e.g., Douglas Mossman, *Critique of Pure Risk Assessment or, Kant Meets Tarasoff*, 75 CIN. L. REV. 523, 556–63 (2006) (This article won the Manfred S. Guttmacher Award in 2008.); Douglas Mossman, *Symposium: Dangerous Decisions: An Essay on the Mathematics of Clinical Violence Prediction and Involuntary Hospitalization*, 2 U. CHI. L. SCH. ROUNDTABLE 95, 105–18 (1995) [hereinafter *Dangerous Decisions*]. See generally Douglas Mossman & Eugene Sommoza, *Balancing Risks and Benefits: Another Approach to Optimizing Diagnostic Tests*, 4 J. NEUROPSYCHIATRY 331 (1992).

<sup>77</sup> A common numeric metric to describe the degree of overlap is  $d'$ , the standardized difference between the means of the respective groups. John A. Swets, *Indices of Discrimination on Diagnostic Accuracy: Their ROCs and Implied Models*, 99 PSYCHOL. BULL. 100, 102 (1986). Formally, this measure indicates the sensitivity of the discriminate ability. Sensitivity is the true positive rate (i.e., the probability of correctly detecting violence, given the person is violent), which can be used to construct a receiver operation characteristic curve (ROC). An ROC is a superior metric of accuracy because it is independent of the base rate and the decision criterion. For reviews, see John A. Swets, *Measuring the Accuracy of Diagnostic Systems*, 240 SCI. 1285 (1988); JOHN SWETS & RONALD PICKETT, *EVALUATION OF DIAGNOSTIC SYSTEMS: METHODS FROM SIGNAL DETECTION THEORY* 29 (1982). See *infra* pp. 94–96 and accompanying notes.

and false negatives, respectively.<sup>78</sup> The acceptable tradeoff is determined by the consequences, or in decision parlance the “utility”<sup>79</sup> associated with each outcome. If false positives are more costly than false negatives, the threshold should be adjusted right to decrease the occurrence of this error type. Conversely, if false negatives are more costly, the threshold should be adjusted left to decrease the occurrence. Decision theory specifies the threshold value that achieves the decision-maker’s preferred balance of error.<sup>80</sup>

There is an additional component that is not readily apparent in Figure 1. The prevalence, or “base rate,”<sup>81</sup> of violence in the patient population also influences the frequency of errors.<sup>82</sup> When an event is rare, over-

<sup>78</sup> Positive or negative refers to whether or not restraints are applied, respectively. True or false refers to whether this is accurate with respect to the true state of affairs.

<sup>79</sup> Utility is a concept used to compare abstract and seemingly incommensurable things. For example, comparing X dollars to Y dollars is simple, but comparing X dollars to the value of love is not straightforward. Utility captures the holistic value of intangible things such as love, virtue, and respect and allows the relative importance of these attributes to be compared to other things. See BARON, *supra* note 71, at 223–24.

<sup>80</sup> It is important to acknowledge that the values of the correct decisions (i.e., true positives and true negatives) are relevant to the decision task. Many juridical decision-analyses assume that the desirability of the true outcomes is equally desirable and exclusively focus on the tradeoff of the errors. *E.g.*, John Kaplan, *Decision Theory and the Fact-Finding Process*, 20 STAN. L. REV. 1065, 1071 (1968); Richard Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021, 1036 (1977). This is partially motivated by Blackstone’s famous maxim that convicting an innocent (false positive) is equal in cost to acquitting ten guilty defendants (false negatives). DeKay, *supra* note 76, at 97. Similar types of tradeoffs have been expressed within the context of civil commitment. John Monahan, *Strategies for an Empirical Analysis of the Prediction of Violence in Emergency Civil Commitment*, 1 L. & HUM. BEHAV. 363, 370 (1977) (“Paraphrasing Blackstone, it may be better that ten ‘false positives’ suffer commitment for three days than that one ‘false negative’ go free to kill someone during that period.”). Mossman also assumes that the value of the correct decisions is equal within the context of civil commitment. See *Dangerous Decisions*, *supra* note 73, at 109.

The assumption that the correct decisions are equal in their desirability, however, is not undisputed within the judicial context. Erik Lillquist, *Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability*, 36 U.C. DAVIS L. REV. 85, 107–11 (2002); cf. Larry Laudan et al., *Re-Thinking the Criminal Standard of Proof: Seeking Consensus About the Utilities of Trial Outcomes*, 7 INT’S COMMENT ON EVIDENCE 2, 12, 14 (2009) (calling it a “Blackstonian fantas[y]” and an “egregious miscalculation” to consider only the costs associated with the errors and not the benefits of the correct outcomes).

<sup>81</sup> Jonathan J. Koehler, *When do Courts Think Base Rate Statistics are Relevant?*, 42 JURIMETRICS J. 373, 374 (2002) (“The social science literature defines a base rate as a proportion—the relative frequency with which an event occurs or an attribute is present in some reference population.”).

<sup>82</sup> This is the signal detection equation:

prediction (false positives) is more likely; when an event is common, under prediction (false negatives) is relatively more likely to occur.<sup>83</sup> This is illustrated in Figure 2 below, which holds the diagnosticity and threshold constant.<sup>84</sup>

$$\frac{P(E/V)}{P(E/NV)} \geq \frac{P(NV)}{P(V)} \frac{U_m - U_{fp}}{U_{tp} - U_{fn}}$$

COOMBS ET AL., *supra* note 73, at 171. The components of the equation are as follows. The left-most term is the likelihood ratio, which corresponds to diagnosticity. It is the ratio of the conditional probability of observing the evidence when ("given") the patient is violent compared to the conditional probability of observing the evidence when ("given") the patient is not violent. The center term is the inverse of the prior odds, which, for our purposes, is just the base rate in odds form. And the right most term is the utility ratio. The product of the prior odds and the utility ratio is  $t^*$ . Hence, it logically follows that as the prior odds change, so would  $t^*$ , all else being equal. See also Anne M. Bisantz, *Signal Detection Theory Demonstration*, U. VIRGINIA (1999), <http://cog.sys.virginia.edu/csees/SDT/index.html> (displaying an interactive and pictorial program that demonstrates this equation).

<sup>83</sup> To demonstrate the impact of base rates, consider the following example. Suppose a test can detect cancer with 99% accuracy. Suppose this test is administered to a population in which 1 person in 1,000 has cancer. Given a person tests positive for cancer, what is the probability that the person actually has cancer? The answer is not 99%. Since the test errors 1% of the time, 1% of the population, or 10 people, who do not have cancer will test positive. Assuming the one person with cancer was actually detected, the probability of actually having cancer, given a positive test result, is 1 in 11 or about 9%.

Now suppose the same test is applied to a population in which 500 out of 1,000 people actually have cancer. The test still errors 1% of the time, identifying 5 cancer-free people as having cancer, and, since it is 99% accurate, it identifies 495 of the 500 who do have cancer. Given a positive test result, then, the probability the person actually has cancer is 495/500, or greater than 99%. See Eric S. Janus & Paul E. Meehl, *Assessing the Legal Standard for Predictions of Dangerousness in Sex*, 3 PSYCHOL. PUB. POL'Y & L. 33, 48 (1997).

<sup>84</sup> See *supra* note 74 and accompanying text.

Figure 2a. Effects of a Low Base Rate of Violence on the Relative Frequency of Errors

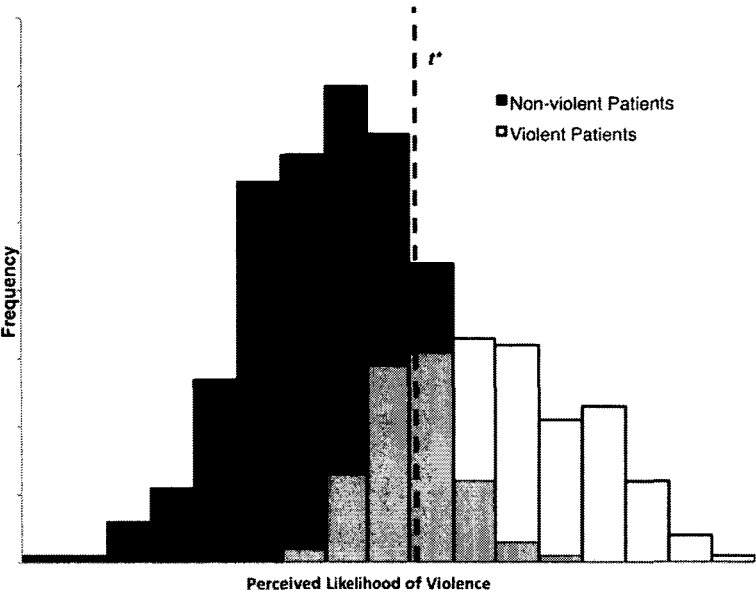
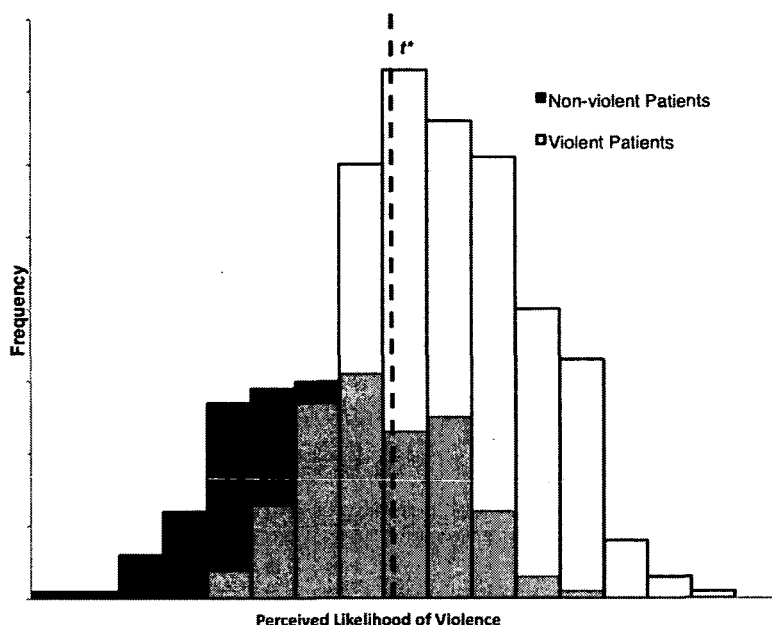


Figure 2b. Effects of a High Base Rate of Violence on the Relative Frequency of Errors



In Figure 2a., violence is relatively rare; the distribution of non-violent patients is twice the size of the distribution of violent patients. The opposite is true for Figure 2b., in which the distribution of violent patients is twice the size of the distribution of non-violent patients. Notice that the proportion of false positives (the proportion of non-violent patients to the right of  $t^*$ ) is considerably greater in 2a. than in 2b., and that the proportion of false negatives (the proportion of violent patients to the left of  $t^*$ ) is greater in 2b. than in 2a. This is because the base rate influences the relative proportion of a given error. The threshold must be adjusted to compensate for the base rate in order to produce the desired frequency of errors (assuming diagnosticity is fixed).

## B. THE INGREDIENTS

Decision theory holds that the decision problem is comprised of three

components, but the theory does not specify what the components are. Probabilities and utilities are not self-defining. This section is concerned with specifying the components. The first subsection addresses the utility issue. We interviewed different stakeholder groups about the use of mechanical restraints, and solicited value judgments about the relative preferability of the various outcomes. The data bearing on the tradeoffs are described. The values pertaining to the other two components—diagnosticity and base rates—are approximated from extant empirical research. Finally, we use these approximations in conjunction with the elicited utility functions to converge on the threshold for restraint use.

### 1. Methods

A total of thirty-two participants<sup>85</sup> were solicited from two separate stakeholder groups: (1) former mental patients; and (2) psychiatrists (hereinafter “doctors”).<sup>86</sup> Of these, twenty-two participants responded to a questionnaire about their experiences with restraints,<sup>87</sup> which probed their personal thoughts and feelings, emphasizing that there is no right or wrong answer. The questionnaire elicited a “preferred marginal tradeoff,” defined as “the number of treatment errors that are acceptable in order to treat correctly one additional person with the disease.”<sup>88</sup> This is akin to eliciting a

---

<sup>85</sup> This was a convenience sample procured through associates of the Saks Institute for Mental Health Law, Policy, and Ethics at the USC Gould School of Law. Approval from the USC Institutional Review Board (IRB) to conduct this research was granted to Nicholas Scurich, *An Empirical Study of The Use of Mechanical Restraints in Psychiatric Practice* (Spring, 2011) (questionnaire and data on file with author).

<sup>86</sup> A few of the doctors pointed out that psychiatric nurses, not doctors, are typically responsible for ordering the use of restraints. Legally, a person from either of these groups is able to order restraints. *Youngberg v. Romeo*, 457 U.S. 307, 323 n.30 (1982). Pursuant to some guidelines, the doctor must formally approve the order within a certain time frame. 42 C.F.R. § 482.13(e)(7) (2011). The reported findings cannot be generalized beyond doctors (psychiatrists) because data were not collected from forensic nurses. There is, however, no reason to believe that psychiatric nurses would have a systematically different perspective than doctors with respect to the use of mechanical restraints.

<sup>87</sup> Three of the returned surveys were unable to be used because they were either incomplete or completed incorrectly. This yielded twenty-two usable surveys from the thirty-two that were disseminated, a response rate of 68.8%.

<sup>88</sup> Peter DeNeef & Daniel L. Kent, *Using Treatment-Tradeoff Preferences to Select Diagnostic Strategies: Linking the ROC Curve to Threshold Analysis*, 13 MED. DECISION MAKING 126, 126 (1993).

Blackstone-type ratio regarding the appropriate tradeoff of errors.<sup>89</sup> However, whereas Blackstone held that a false conviction (false positive) is considerably less desirable than a false acquittal (false negative), the current study did not make such a presupposition. Instead, participants were first asked which of the two errors they considered to be worse, and then, depending on their response, were asked how many errors they would tolerate at the expense of just one of the less desirable errors.<sup>90</sup> This approach isolates the utility aspect from other considerations, such as the base rate or diagnosticity.<sup>91</sup>

## 2. Preferred Marginal Tradeoffs

Former mental health patients responded unanimously that false positives are worse than false negatives. The median patient tradeoff indicated that 141 false negatives were equivalent to one false positive (geometric mean: 86 false negatives equivalent to 1 false positive).<sup>92</sup> In contrast, the majority of doctors indicated that false negatives were worse than false positives. The median response from doctors indicated that seven false positives were equivalent to one false negative (geometric mean: 4.5 false positives equivalent to 1 false negative.)

---

<sup>89</sup> *Dangerous Decisions*, *supra* note 76, at 110.

<sup>90</sup> For example, if a participant indicated that false negatives are the least desirable of the two errors, she was then asked, "How many false positives would you tolerate in order to prevent just one act of harm (false negative)?"

If a participant indicated that false positives are the worst error, she was then asked, "How many false negatives would you tolerate in order to prevent just one patient from being unnecessarily restrained (false positive)?" See also Douglas Mossman & Eugene Somoza, *Balancing Risks and Benefits: Another Approach to Optimizing Diagnostic Tests*, 4 J. NEUROPSYCHI. 331, 334 (1992) (describing the same approach).

<sup>91</sup> DeNeef & Kent, *supra* note 88, at 126.

<sup>92</sup> *Id.* The geometric mean, defined as the  $n$ th root of the product of all observations, is useful when the quantity is a ratio. Ratio data are "centered" around 1.0, with ratios less than 1 compressed into the interval (0,1) and ratios greater than 1 spread over the range (1,∞). This creates a problem for arithmetic means, which give more weight to ratios greater than 1 than those less than 1. For example, 2 observations that are inverses, 10.0 and 0.10, would yield an arithmetic mean of 5.05, but a geometric mean of 1.0  $[(10.0)(0.10)]^{1/2}$ . While arithmetic means give greater weight to ratios greater than 1.0, geometric means weigh  $k$  and  $1/k$  equally. Note that the geometric mean is equivalent to the exponential of the arithmetic average of the natural logs of the ratios.

### 3. Diagnosticity

Ascertaining diagnosticity would require controlled experimentation, an endeavor that is potentially quite costly and practically difficult to undertake. Fortunately, studies on the clinical ability to predict violence are voluminous.<sup>93</sup> Although these studies were not concerned with the use of restraints per se, they do provide a rough estimate of the clinical ability to forecast violence.<sup>94</sup> The best meta-analysis<sup>95</sup> of the clinical ability to forecast violence indicates a range of  $d'$  from 0.62 to 1.04,<sup>96</sup> where  $d' = 0$  indicates no diagnostic ability and  $d' = 4$  is considered a practical upper limit.<sup>97</sup> To get an idea of the magnitude of this diagnosticity, consider the following values from other domains: police officer's ability to detect truth-tellers from liars:  $d' = 0$ ; eyewitness's ability to correctly recall a face:  $d' = .08$ ; ability of a mammogram to detect breast cancer:  $d' = 1.3$ ; diagnostic success rate of computer tomography scans (CT scans):  $d' = 2.65$ ; ability of polygraph machine to detect liars:  $d' = 2.70$ .<sup>98</sup> Plainly stated, "clinicians are able to distinguish violent from non-violent patients with a modest, better-than-chance level of accuracy."<sup>99</sup>

<sup>93</sup> See generally JOHN MONAHAN, STRUCTURED RISK ASSESSMENT OF VIOLENCE, in TEXTBOOK OF VIOLENCE ASSESSMENT AND MANAGEMENT 17, 17 (Robert Simon & Kenneth Tardiff eds., 2008) (describing the breadth of research on clinical prediction of violence).

<sup>94</sup> We are unaware of any empirical evidence that examines the accuracy with which professionals are able to discern violent patients for the sole purpose of using mechanical restraints. The extant literature on violence prediction examines the ability of clinicians to predict which patients will be violent when released into the community, generally without supervision. Admittedly, these are two different tasks, but there is no reason to believe that the ability level would differ dramatically.

<sup>95</sup> A meta-analysis is a quantitative summary of previous research.

<sup>96</sup> Douglas Mossman, *Assessing Predictions of Violence: Being Accurate About Accuracy*, 62 J. CONSULTING & CLINICAL PSYCHOL. 783, 788–89 (1994) (analyzing 58 datasets from 44 studies on violence prediction) The metric used by Mossman for describing accuracy is the area under the ROC curve (AUC), which spans a range of 0.5, indicating an ability to discriminate no better than chance, up to 1.0, indicating perfect ability to discriminate. *Id.* at 788. For all datasets combined, the median AUC is 0.73 and mean 0.77 (+/- 0.0048). *Id.* The AUC for the studies that only examined clinical judgment is 0.67. *Id.* at 789. To be consistent throughout this paper, the AUC was converted to  $d'$ : AUC is equal to  $A_z$ ,  $d'$  is equal to  $Z_a X^{1/2}$ . SWETS & PICKETT, *supra* note 77, at 32. Converting the reported values yields: AUC = 0.73,  $d' = .86$ ; AUC = 0.77,  $d' = 1.04$ ; AUC = 0.67,  $d' = .62$ .

<sup>97</sup> *Id.*

<sup>98</sup> Hal Arkes & Barbara Mullers, *Do Juries Meet Our Expectations?*, 26 L. & HUM. BEHAV. 625, 632–33 (2002).

<sup>99</sup> *Dangerous Decisions*, *supra* note 76, at 127.

#### 4. Base Rates

Research finds that the base rate of violence on psychiatric wards varies dramatically. Much of this variability is attributable to inter-study differences in the operational definitions of violence, the length of the study, and the method used to determine whether violence occurred. A lack of standardization in restraint use may also contribute to inter-study variability. For example, highly risk adverse staff would be more likely to use preventative interventions, which presumably lower the base rate by thwarting violence. Other staff may be more concerned with patients' liberty interests and thus more reluctant to engage in preventative intervention, which could result in an increase in the base rate of violence.

Regardless of definitional issues and inter-study variability, the base rate of violence on psychiatric wards appears to be low. In general, a small minority of psychiatric inpatients will engage in violence.<sup>100</sup> Certain patients tend to be repeat offenders in that a small number of individuals tend to commit the majority of violence.<sup>101</sup> In order to meaningfully com-

---

<sup>100</sup> See Antonio Convict et al., *Characteristics of Repeatedly Assaultive Psychiatric Inpatients*, 41 HOSP. & CMTY. PSYCHIATRY 1112, 1113 (1990) (20% of patients engaged in violence in a six-month period); Dale McNeil & Renee Binder, *Clinical Assessment of the Risk of Violence Among Psychiatric Inpatients*, 148 AM. J. PSYCHIATRY 1317, 1320 (1991) (17.4% of patients engaged in violence in an eleven-month period); Bradley Ng et al., *Ward Crowding and Incidents of Violence on an Acute Psychiatric Inpatient Unit*, 54 PSYCHIATRIC SERVICES 521, 523 (April 2011) (finding 12.3% of patients engaged in physical violence during a one-year period); Kenneth Tardiff & Attia Sweillam, *Assault, Suicide and Mental Illness*, 37 ARCHIVES GEN. PSYCHIATRY 164, 168 (1980) (10% of 9,365 patients engaged in violence within one year of admission to a psychiatric hospital); Kenneth Tardiff & Attia Sweillam, *Assaultive Behavior Among Chronic Inpatients*, 139 AM. J. PSYCHIATRY 212, 213 (1982) (7.4% of 5,164 patients engaged in violence in a one-month period); Kenneth Tardiff, *Characteristics of Assaultive Patients in Private Hospitals*, 141 AM. J. PSYCHIATRY 1232, 1234 (1984) (9.3% of 1,768 patients engaged in violence in a two-year period); Jerome Yesavage et al., *Inpatient Evaluation of Aggression in Psychiatric Patients*, 169 J. NERVOUS & MENTAL DISEASE 299, 300 (1981) (23.1% of patients engaged in violence during the first eight days of admission to the Intensive Care Unit).

<sup>101</sup> See Cathy Owen et al., *Repetitively Violent Patients in Psychiatric Units*, 49 PSYCHIATRIC SERVICES 1458, 1458 (1998) (12% of patients committed 68% of all violence); see also Peter Noble & Sheila Rodger, *Violence by Psychiatric Inpatients*, 155 BRIT. J. PSYCHIATRY 384, 385 (1989) (reporting a study in which 52.5% of violent patients committed more than one assault); Cathy Owen et al., *Violence and Aggression in Psychiatric Units*, 49 PSYCHIATRIC SERVICES 1452, 1454 (1998) [hereinafter, *Violence and Aggression*] ("A small cohort of 20 [2.3%] patients caused 67 percent of the 1,289 violent or aggressive incidents during the course of the study."); Edwin Tam et al., *Patterns of Violent Incidents by Patients in a General Hospital Psychiatric Facility*, 47 PSYCHIATRIC SERVICES 86, 86 (1996) (32% of pa-

pare the base rates across studies it is necessary to hold constant the rate of violence per time interval. The number of violent acts per day<sup>102</sup> yields a range from zero to about six acts of violence per day, on average, with an arithmetic mean of roughly 1.9.<sup>103</sup>

### C. COMBINING THE LOGIC AND THE INGREDIENTS

The tradeoffs elicited from the various stakeholders are now translated into a threshold value. According to doctors, one false negative is equivalent to about seven false positives (4.5 using geometric mean), suggesting that restraints should be used whenever the likelihood of violence is above about 0.15 (0.18 using geometric mean). On the other hand, former mental patients believe that about 141 false negatives are equivalent to 1 false positive (86 using geometric mean), implying that the likelihood of violence needs to be in excess of 0.99 (0.98 using geometric mean) in order to use restraints.

To illustrate the magnitude of these risk levels, consider that the chance of a heads in a (fair) coin toss is 0.50. Hence doctors require a probability of violence considerably lower than landing heads in a coin toss in order to implement restraints, whereas patients would require that the chance of violence be much greater. The coin flip analogy is not entirely apt here, because the likelihood of violence estimates refer to subjectively perceived probabilities, not frequency estimates (i.e., a heads occurs 50% of the time in the long run). Still, the likelihood of violence estimates can be likened to the certainty required before one would take a bet that the proposition is true.<sup>104</sup> In other words, a doctor would be indifferent be-

---

tients accounted for 75% of all violence.).

<sup>102</sup> This can be calculated by converting the total time period of the study into the number of days and dividing the number of violent acts by this value. For example, if ten acts of violence were observed in a thirty-day period, the rate of violence per day is 10/30, or about 0.33 acts of violence per day. This approach assumes that a fixed proportion of previously non-violent patients engage in violence each day. This assumption may not be tenable given the fact that a minority of patients are responsible for the majority of violence. See Noble & Rodger, *supra* note 101, at 385.

<sup>103</sup> Ng et al., *supra* note 100, at 523 (0.09 acts of violence per day); Noble & Rodger, *supra* note 101, at 385 (1.29 acts of violence per day); *Violence and Aggression*, *supra* note 101, at 1452 (6.13 violent acts per day); Tam et al., *supra* note 101, at 86 (0.36 acts of violence per day); Convict et al., *supra* note 100, at 1113 (2.76 acts of violence per day).

<sup>104</sup> BARON, *supra* note 71, at 102 (noting that probability can be understood as a personal willingness to bet that a proposition is true).

tween the betting that the use of restraints is appropriate and taking a gamble where the chances of success are 0.15, whereas the patient would not take the gamble unless the chances of success are greater than 0.99. Of course 0.99 is close to near certainty.

#### IV. NORMATIVE IMPLICATIONS

The value judgments about the error tradeoffs were radically disparate. In fact, the doctors and patients did not even agree about which error was less desirable. This underpins the dramatic difference in opinion between the doctors and patients about when the use of restraints is justified, and naturally leads to the question: whose tradeoff judgment *ought* to be used? The *Youngberg* decision provides a beginning point to address this question. Recall *Youngberg* held that the standard for evaluating the use of restraints is whether adequate professional judgment was exercised. But the Court also held that patients have a liberty interest to be free from restraints except under certain conditions. Although professionals are in charge of determining these conditions, the Court clearly held that patients have a legitimate interest in the decision to use mechanical restraints. Thus, at a minimum, *Youngberg* can be taken to hold that no one stakeholder's policy dominates, and that the interests of both are relevant to the decision-making calculus.

Assuming both stakeholders ought to be included in a normative policy, the issue becomes determining the amount of weight to ascribe to each perspective. It seems unlikely that the weights are equipoise for both stakeholder groups. Table 1 below indicates some possible weighting schemes for the tradeoffs of the two groups. The two left-most columns indicate the weights for the relevant stakeholder; the center two columns indicate how this weighting scheme would affect the tradeoff ratio (false negatives relative to false positives); and the right-most column indicates what effect this would have on the threshold, expressed in probabilistic terms.

Table 1. Weighted Stakeholder, Tradeoffs Ratios (TR), and Threshold Probabilities

Stakeholder Group Weights		Weighted Tradeoffs, FN:FP <sup>1</sup>		Threshold Probabilities for Violence <sup>2</sup>	
		Mean <sup>3</sup>	Median <sup>4</sup>	Mean <sup>3</sup>	Median <sup>4</sup>
Doctors (W <sub>1</sub> )	Patients (W <sub>2</sub> )				
0.9	0.1	0.40	0.28	0.29	0.22
0.7	0.3	1.3	1.1	0.57	0.53
0.5	0.5	4.4	4.5	0.81	0.82
0.3	0.7	14.4	17.8	0.94	0.95
0.1	0.9	47.4	70.9	0.98	0.99

<sup>1</sup>Weighted Tradeoff =  $e^{[(W_1/n)(TR_1) - (W_2/n)(TR_2)]}$

<sup>2</sup>Threshold violence probability =  $1 - \left(\frac{1}{1 + TR}\right)$

<sup>3</sup>Uses stakeholder group geometric means of Tradeoffs

<sup>4</sup>Uses stakeholder group medians of Tradeoffs

At one extreme, the doctor's perspective is weighted quite heavily (9:1), and tradeoff ratio is less than 1, with a corresponding threshold probability of less than 0.30. When doctors are weighted just a little over twice patients (7:3), the tradeoff ratio is only slightly greater than one, with a corresponding threshold probability of slightly over 0.50. With equal weights for the two stakeholder groups, the tradeoff is over 4, with corresponding threshold probabilities greater than 0.80. When patients are weighted higher than doctors, the tradeoff ratios climb higher, and the threshold probabilities exceed 0.95.

## V. DISCUSSION

The Court's posture in *Youngberg* is clear: the Court places more faith in the doctor's decision to use restraints than in judicial second-guessing. This holding, in conjunction with the nebulous requirement of "danger," has made it nearly impossible to evaluate the appropriateness of the decision to use mechanical restraints. The analytical framework utilized in this paper, however, provides a means for evaluation by enabling one to discern when the use of restraint is legitimate or illegitimate by

considering whether the requisite level of danger was apparent. Thus, aside from the empirical findings, this paper provides a framework to reify the unhelpful case law and empty statutory language pertaining to the use of mechanical restraints.

The disparate tradeoffs are perhaps the most interesting aspect of this paper. One should bear in mind that the reported tradeoffs are based on aggregate data. Some of the responses were even more extreme than the summary statistics (i.e., median; geometric mean) indicate. For example, one doctor indicated that she would tolerate an infinite number of false positives to avoid just one false negative and stated, "There is [sic] never too many false positives."<sup>105</sup> Responses from some of the patients could not be more diametric, as one patient lamented, "The effect of being restrained has never left me. The trauma has led to a PTSD diagnosis. . . . There has to be another way [to do things]. There has to be."<sup>106</sup> This patient indicated that 1,000 false negatives are equivalent in cost to just one false positive.

Several explanations for these disparities are possible. One is a type of self-serving bias, since patients bear many or all of the costs associated with restraints while doctors are mostly the beneficiary. There could also be divergent perceptions about restraints and their consequences. For example, suffering and abuse might quickly come to mind when patients think about restraints, but doctors may think of instances when restraints were successful or when they believe restraints saved them from harm or injury.<sup>107</sup> The disparity could also result from an empathy gap on the doctor's behalf. None of the doctors in this study had experienced restraints first-hand. Given that staff members who experience restraints firsthand are much less likely to apply restraints to others,<sup>108</sup> it is plausible that doctors would have a different tradeoff after being placed in mechanical restraints themselves. It would be interesting to examine whether a doctor's tradeoff changes after experiencing restraints firsthand.

The tradeoffs reported in this paper and the corresponding thresholds

---

<sup>105</sup> Scurich, *supra* note 85.

<sup>106</sup> *Id.*

<sup>107</sup> *Id.* (statement of surveyed doctor) ("I've seen and heard of too many people being attacked to have a more liberal approach. I believe in restraints for the safety of the staff, other patients, and the patients acting out themselves.").

<sup>108</sup> See Forster et al., *supra* note 39, at 270.

depend on the specific type of restraint used. In the survey, restraints were described as full-body restraints. More or less intrusive forms of restraints would invoke a different set of consequences. It is axiomatic in decision theory that decisions are a function of their potential consequences, so it would be rational for different forms of restraints to have different thresholds that justify their use. As a general heuristic, if the potential consequences are less serious, the necessary level of risk to justify the use of restraints would decrease. As the potential consequences become increasingly dire, the level of potential risk required to use restraints would increase.

The endeavor to create and apply a universal threshold is seemingly inconsistent with certain jurisprudential ideals, such as the notion of "individualized justice"; that is, the treatment of people as individuals, rather than as members of a statistical group, who are either above or below an apriori threshold level of risk.<sup>109</sup> Regardless of whether this is true, making case-by-case determinations is highly unlikely to yield a tradeoff that is consistent normative policy. Indeed, case-by-case determinations are largely capricious and idiosyncratic.<sup>110</sup> The decision to use restraints should include pertinent information about the individual, but it should also be congruent with normative policy. Under the current approach, information about the specific individual is relevant to the perceived likelihood of violence; whether or not this likelihood is sufficient to justify the use of restraints is an issue that depends on the normative value tradeoff. Thus, individual considerations are still relevant to and included in the decision.

It might be argued that the metric utilized in this paper is cryptic and

---

<sup>109</sup> The moral, logical, and mathematical issues involved with using group data to make decisions about individuals are discussed in Nicholas Scurich & Richard S. John, *A Bayesian Approach to the Group Versus Individual Prediction Controversy in Actuarial Risk Assessment*, L. & HUM. BEHAV. (in press). Suffice it to say that any putatively "uniquely individualized" judgment must be a distillation of group tendencies otherwise it is dogma by definition.

<sup>110</sup> William M. Grove & Paul E. Meehl, *Comparative Efficiency of Information (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical-Statistical Controversy*, 2 PSYCHOL. PUB. POL'Y, & L. 293, 307-08 (1996). There is a phenomena known as "broken leg countervailings" which is typically invoked to override empirically established decision rules. Roughly, the idea is that a formula might do quite well at making predictions, but there could be an anomaly not accounted for in the formula that would render the formulaic prediction spurious. Although broken leg countervailings are theoretically possible, there are no concrete examples of their existence, and, moreover, people would not be proficient at detecting them.

not very practical. It is true that doctors would probably not attempt to quantify their estimate of the likelihood that the patient will engage in violence. Practicing physicians probably think in crude, non-numeric terms, describing the likelihood of violence as “likely” or “very likely.” Moreover, the judiciary would have a difficult time second-guessing the doctor’s assessment, especially since the judiciary is not present during the situation. These are valid criticisms, but they are not devastating to the paper’s enterprise. The analysis contained in this paper is normative—indicating how these decisions ought to be made. The fact that decision-makers do not live up to these ideals does not mean the ideals are invalid. Instead, research could be undertaken to help decision-makers come closer to meeting the mark. One potential avenue for improvement is by using actuarial (statistical) or “structured professional judgment” to assess the likelihood of violence.<sup>111</sup> Such approaches are not intensely onerous, and greatly improve the reliability of diagnostic decisions.<sup>112</sup> It is plausible that some form of structured professional judgment could be adapted to be used for the purposes of administering restraints. Perhaps doctors could receive training to help calibrate their risk assessments and transcend the use of crude, non-numeric terms, which have been shown to have no consensus within the psychiatric field.<sup>113</sup>

---

<sup>111</sup> See Monahan, *supra* note 93, at 20–27 (discussing structured approaches to assess violence risk). Structured Professional Judgment does not follow a rigid algorithm, but instead identifies the relevant risk variables and permits the clinician to intuitively amalgamate them into a risk estimate. Jennifer L. Skeem & John Monahan, *Current Directions in Risk Assessment*, 20 CURRENT DIRECTIONS PSYCHOL. SCI. 38, 39 (2011). See Kevin S. Douglas, *Expert Opinion: Assessing Risk for Violence Using Structured Professional Judgment*, 29 AM. PSYCHOL.-L. SOC’Y NEWS L. 12, 12–14 (2009) (providing an overview of Structured Professional Judgment).

<sup>112</sup> Robyn M. Dawes, David Faust & Paul E. Meehl, *Clinical Versus Actuarial Judgment*, 243 SCI. 1668, 1668–1673 (1989) (A seminal paper on the superiority of actuarial approaches to prediction.). Actuarial risk assessment instruments designed specifically to assess violence risk have now been incorporated into computer software. Basically, a clinician inputs the risk variables that are present in the specific case and a computer program calculates the relative risk posed. This process is not unlike that used to determine driving risk in order to calculate car insurance. See generally John Monahan et al., *The Classification of Violence Risk*, 24 BEHAV. SCI. & L. 721 (2006) (“[Classification of Violence Risk is an] interactive software program designed to estimate the risk that a person hospitalized for mental disorder will be violent to others.”).

<sup>113</sup> Zoe N. Hilton et al., *Does Using Non-Numeric Terms to Describe Risk Aid Violence Risk Communication?*, 23 J. INTERPERSONAL VIOLENCE 171, 182–85 (2008).