

# Theorising Gambling Self-Exclusion Agreements: The Inadequacy of Procedural Autonomy

Bernard Long

University College Cork, Ireland

## Abstract

Gambling self-exclusion agreements enable a person to have themselves prevented from gambling for some future period. In light of evidence of their effectiveness in helping problem gamblers manage their addiction, these agreements enjoy growing popularity. In particular, several jurisdictions now oblige gambling operators to offer self-exclusion to their clientele. If self-exclusion has a unique value that is distinct from paternalistic measures, such as forced exclusion, it is surely because it prizes the gambler's autonomy. In this article, however, I will argue that self-exclusion's theoretical basis cannot, in fact, be found in a procedural theory of autonomy that only regards agents' own values and decisions. Rather, I will contend that if agents may bind their future selves in only some ways—for example, by preventing themselves from gambling but not preventing themselves from self-excluding or selling themselves into slavery—it can only be because of a normative, substantive claim.

---

Keywords: *Autonomy; Gambling addiction; Gambling self-exclusion agreements; Procedural autonomy*

## I. Introduction

Gambling self-exclusion agreements are a form of precommitment whereby a person may request to be prevented from gambling for some future period or indefinitely.<sup>1</sup> They are used primarily, though not exclusively, by problem gamblers<sup>2</sup> to forestall future destructive bouts of indulgence in their addiction.<sup>3</sup>

- 
1. As regards what is meant by prevention, operators in the UK are required to 'refuse service' to self-excluders and, in the case of holders of non-remote licences, remove self-excluders from the premises. See The United Kingdom Gambling Commission, *Licence Conditions and Codes of Practice*, UKGC, 2020, Part 2, ss 3.5.1(1), 3.5.1(6)(d), 3.5.3(1), 3.5.8(1) [LCCP].
  2. Although distinctions between terms such as 'problem gambling', 'gambling addiction', 'Gambling Disorder', and their cognates are significant for clinical purposes, I will use them interchangeably in this article.
  3. Though self-exclusion agreements are primarily used as a tool to address gambling addiction, it is wrong to assume that the presence of such an agreement is *prima facie* evidence of problem gambling. In one study, 26.3% of respondent self-excluders were motivated to self-exclude due to annoyance with a specific operator. These respondents had therefore self-excluded in order to "punish" operators for perceived wrongdoing, rather than to address any problematic gambling patterns of their own. Tobias Hayer & Gerhard Meyer, "Internet Self-Exclusion: Characteristics of Self-Excluded Gamblers and Preliminary Evidence for Its Effectiveness" (2011) 9:3 *International Journal of Mental Health and Addiction* 296 at 301. In an English

Self-exclusion is not, of course, an absolute panacea for gambling addiction. An overwhelming majority of studies in the area have demonstrated that some self-excluders continue to gamble during their period of exclusion, in other jurisdictions, by other forms of gambling, or even in venues from which they are apparently excluded.<sup>4</sup> Despite its limits, however, self-exclusion is generally considered to be at least somewhat effective in minimising risk and assisting problem gamblers in managing their addiction.<sup>5</sup> As such, it continues to enjoy a burgeoning popularity. Though self-exclusion agreements originated as an industry driven initiative, their provision by operators (or, alternatively, operators' participation in a state-run scheme) has increasingly become a legislative and/or regulatory requirement globally.<sup>6</sup>

Perhaps due to its relative nascence, self-exclusion is beset with dissonance and ambiguity. This dissonance can be found even in matters as simple as terminology, with self-exclusion *agreements* in the UK paralleled by self-exclusion *orders* in Queensland, for example.<sup>7</sup> These differences may also be reflected in the place of self-exclusion within overall legislative frameworks. Whereas in the UK self-exclusion is a standalone procedure, in New Zealand it is merely part of a broader exclusion procedure. As such, self-exclusion in New Zealand seemingly differs little from forced exclusion procedures, except that it is initiated by the prospective excluded person rather than the holder of an operating licence.<sup>8</sup> This difference may betray a fundamental disagreement over the

---

case, there was a suggestion, tacitly accepted by the court, that the defendant had sought to strategically use self-exclusion agreements as an “each way bet” in order to insulate himself against losses without foregoing profits. *The Ritz Hotel Casino Ltd v Al Geabury*, [2015] EWHC 2294 (QB) at paras 57, 67.

4. See for example Roxana Kotter, Anja Kräplin & Gerhard Bühringer, “Casino Self- and Forced Excluders’ Gambling Behavior Before and After Exclusion” (2018) 34:2 *Journal of Gambling Studies* 597; Robert Ladouceur et al, “Analysis of a Casino’s Self-Exclusion Program” (2000) 16:4 *Journal of Gambling Studies* 453; Robert Ladouceur, Caroline Sylvain & Patrick Gosselin, “Self-Exclusion Program: A Longitudinal Evaluation Study” (2007) 23:1 *Journal of Gambling Studies* 85; Sarah E Nelson et al, “One Decade of Self Exclusion: Missouri Casino Self-Excluders Four to Ten Years after Enrollment” (2010) 26:1 *Journal of Gambling Studies* 129; Responsible Gambling Council, *From Enforcement to Assistance: Evolving Best Practices in Self-Exclusion* (Responsible Gambling Council, 2008) at 38-39; Anders Håkansson & Carolina Widinghoff, “Gambling Despite Nationwide Self-Exclusion: A Survey in Online Gamblers in Sweden” (2020) 11 *Frontiers in Psychiatry*, DOI: <https://doi.org/10.3389/fpsy.2020.599967>.
5. See for example Sally M Gainsbury, “Review of Self-exclusion from Gambling Venues as an Intervention for Problem Gambling” (2014) 30:2 *Journal of Gambling Studies* 229; Roxana Cotter et al, “A Systematic Review of Land-Based Self-Exclusion Programs: Demographics, Gambling Behavior, Gambling Problems, Mental Symptoms, and Mental Health” (2019) 35:2 *Journal of Gambling Studies* 367; Dylan Pickering, Alex Blaszczyński & Sally M Gainsbury, “Multi-Venue Self-Exclusion for Gambling Disorders: A Retrospective Process Investigation” (2018) 38 *Journal of Gambling Issues* 127.
6. For examples in the English-speaking world, see non-exhaustively *LCCP*, *supra* note 1, Part 2, s 3.5; *National Gambling Act, 2004* (S Afr), No 7 of 2004, s 14(1); *Gambling Act 2003* (NZ), 2003/51, s 310; *Gambling Regulation Act 2003* (Vic), 2003/114, s 3.4.12A; *Casino Control Act 1992* (NSW), 1992/15, s 79(3); *Casino Control Act 1982* (Qld), 1982/78, ss 91N-91O; 58 Pa Code § 503a (2021); 68 IAC 6-3-3 (2018); *Casino Control Act 2006* (Sing), s 165A(1)(c).
7. See *Casino Control Act 1982*, *supra* note 6, s 91O.
8. See *Gambling Act 2003*, *supra* note 6, ss 309-310. According to s 309, a venue manager or licence holder must, if they suspect a patron of being a problem gambler, approach the person

conceptual aspects of self-exclusion. If a problem gambler in New Zealand is advised of the option to self-exclude and rejects it, they may nonetheless be excluded for the same period at the behest of the licence holder. This approach seems to drastically reduce the significance of the gambler's consent to exclusion, with the looming threat of forced exclusion making the deontological 'self-' caveat largely irrelevant in pursuing the *telos* of excluding problem gamblers.<sup>9</sup>

In this article, I will consider one vital, though underexamined and certainly unresolved, ambiguity at the heart of self-exclusion agreements. Self-exclusion is ostensibly distinguishable from forced exclusion and other paternalistic measures by the apparent consent of the self-excluder. Any unique merit which it might be thought to have must surely exist because it values the gambler's autonomy, rather than foregoing that autonomy and subjecting the gambler to paternalistic measures against their will. While I will briefly discuss the issues with paternalism shortly, self-exclusion might be prized because it coheres with a general societal favour for autonomy.<sup>10</sup>

A closer inspection, however, reveals that the matter is more complicated than self-exclusion being the autonomy-valuing cousin to paternalistic forced exclusion. Self-exclusion is not *tout court* autonomy-promoting. While the decision to self-exclude is presumably autonomous,<sup>11</sup> it precludes a subsequent (and possibly autonomous) decision to gamble. This preclusion is as limiting to the gambler's later autonomy as forced exclusion. Of course, many will argue that this limitation of autonomy is justified by the earlier decision. This, however, generates what I will call the 'self-exclusion problem': whether it is acceptable for a present decision to limit a gambler's subsequent autonomy and, if so, to what extent and/or in what circumstances. This is an issue which is ostensibly not restricted to gambling self-exclusion agreements. However, self-exclusion differs in crucial ways from other instances of self-binding. Most obviously, it differs from what is perhaps the most prominent example of making decisions for one's future self, Advance Healthcare Directives (AHDs). AHDs, unlike

---

with information on problem gambling, including the self-exclusion procedure outlined in s 310. After offering information, the venue manager or licence holder may issue an involuntary exclusion order.

9. Elster opines that freedom must mean the ability to both choose something or reject it. An Australian, he says, is not free to vote because Australia's mandatory voting laws mean they are not equally free not to vote. See Jon Elster, *Sour Grapes: Studies in the Subversion of Rationality* (Cambridge University Press, 1983) at 128-29. If this is the case, a prospective self-excluder in New Zealand, under the potential threat of forced exclusion if they do not self-exclude, is certainly not as free as their British counterpart.
10. A recent article argues that allowing self-exclusion from purchasing firearms as a suicide prevention measure is "a logical extension of modern society's respect for autonomy." Melvin G McInnis et al, "Suicide Prevention and Mood Disorders: Self-exclusion Agreements for Firearms as a Suicide Prevention Strategy" (2021) 13:3 Asia-Pacific Psychiatry, DOI: <https://doi.org/10.1111/appy.12455> at 3. Self-exclusion from firearms ownership is, however, too different from gambling self-exclusion to be of any useful comparative value. Firearms have non-recreational uses which gambling does not, such as self-defence, and pose an immediate threat to others in a way gambling also does not.
11. This is merely an assumption. As I will discuss later, the decision to self-exclude may not take place in autonomy-conducive circumstances.

gambling self-exclusion, relate to future times when the agent lacks capacity to make a decision.<sup>12</sup> As I will discuss later, the capacity of a relapsed gambling addict is contestable. However, the gambler is certainly far more able to express a contrary preference than a person in a vegetative state or, even, a dementia sufferer.<sup>13</sup> Gambling addiction can also be distinguished from substance-related addictions. A problem gambler's desire to partake in their vice cannot be as easily dismissed as that of an alcoholic or drug addict under the influence of extraneous chemicals or suffering withdrawal therefrom.<sup>14</sup> Similarly, self-exclusion differs from another form of precommitment used to address addiction, disulfiram. Disulfiram, sold as Antabuse, discourages alcohol consumption by inducing negative physiological responses to alcohol. However, once disulfiram has been ingested, its physiological consequences are inevitable. In other words, the

- 
12. See for example *Mental Capacity Act 2005* (UK), ss 24-26.
  13. Dementia sufferers were at the centre of one of the most prominent twentieth-century debates around advance directives, between Ronald Dworkin, Daniel Callahan, Rebecca Dresser, and Sanford Kadish. For a summary of this debate, see Michael J Newton, "Precedent Autonomy: Life-Sustaining Intervention and the Demented Patient" (1999) 8:2 *Cambridge Quarterly of Healthcare Ethics* 189.
  14. Comparisons between gambling and substance addictions are complex, being marked by similarities and dissimilarities. For example, research indicates that sufferers of Gambling Disorder and Substance Use Disorders alike suffer cravings and cognitive impairment. See American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders: DSM-5* (American Psychiatric Association, 2013) at 491, 509, 520, 523, 534, 541, 550, 561, 571, 577 [*DSM-5*]; Hermano Tavares et al, "Comparison of Craving Between Pathological Gamblers and Alcoholics" (2005) 29:8 *Alcoholism: Clinical & Experimental Research* 1427; Benjamin J Morasco et al, "Psychological Factors that Promote and Inhibit Pathological Gambling" (2007) 14:2 *Cognitive & Behavioral Practice* 208; Anna E Goudriaan et al, "Neurocognitive Functions in Pathological Gambling: A Comparison with Alcohol Dependence, Tourette Syndrome and Normal Controls" (2006) 101:4 *Addiction* 534; Andrew J Lawrence et al, "Problem Gamblers Share Deficits in Impulsive Decision-making with Alcohol-dependent Individuals" (2009) 104:6 *Addiction* 1006; Antonio Verdejo-Garcia et al, "Neural Substrates of Cognitive Flexibility in Cocaine and Gambling Addictions" (2015) 207:2 *British Journal of Psychiatry* 158. Clinically, Gambling Disorder shares various diagnostic criteria with certain Substance Use Disorders, including tolerance and withdrawal symptoms. See *DSM-5*, *supra* note 14 at 490-91, 509, 541, 550-51, 561, 571, 577-78, 585. Put plainly, the psychological urge to relapse is likely equally strong amongst recovering gambling addicts and their substance-addicted counterparts. However, problem gambling, as a behavioural addiction, is unlikely to deleteriously affect the brain in the same manner as recurrent substance use. See Jon E Grant & Samuel R Chamberlain, "Gambling Disorder and Its Relationship with Substance Use Disorders: Implications for Nosological Revisions and Treatment" (2015) 24:2 *The American Journal on Addictions* 126 at 128. For example, Lawrence et al, *supra* note 14 found that those with alcohol dependence exhibit greater impairment of executive functions, such as working memory, than problem gamblers. Elsewhere, Volkow et al suggest that prolonged substance abuse may alter addicts' brain chemistry, negatively affecting inhibitory control, emotional regulation, and decision-making. See Nora D Volkow et al, "Addiction: Beyond Dopamine Reward Circuitry" (2011) 108:37 *Proceedings of the National Academy of Sciences* 15037 at 15041. Returning to the *DSM-5* diagnostic criteria, while Gambling Disorder and Substance Use Disorders share withdrawal as a symptom, withdrawal from the former includes only restlessness and irritability, whereas withdrawal from some substances includes a range of physical symptoms, such as nausea, tremors, and insomnia. See *DSM-5*, *supra* note 14 at 499, 517-18, 547-48, 557-58, 569, 575. The recovering substance addict is therefore subject to a greater range of autonomy-impugning factors than their gambling-addicted counterpart. Whatever is to be said about these respective addictions' similarities for clinical purposes, I consider this sufficient to justify gambling addiction's unique treatment in this article.

alcoholic cannot choose to relapse without suffering the adverse effects of disulfiram. In the case of self-exclusion, there is no inevitable, physical response to relapse. Rather, its enforcement relies on the cooperation of others, i.e., the gambling operator and its staff. Justifying the necessary role of these parties in sustaining the self-exclusion agreement is what requires theoretical engagement with the self-exclusion problem.

Finally, self-exclusion is distinct from two other forms of self-binding referred to in passing in this article—namely, self-exclusion from firearms purchases<sup>15</sup> and covenant marriage<sup>16</sup>—in that gambling is a primarily self-regarding activity. In the former case, firearms pose a significant and immediate threat to the safety of others, while in the latter, a party to a covenant marriage has made promises to another party, their spouse. On the other hand, a theoretical basis for gambling self-exclusion cannot depend on the interests of others.<sup>17</sup> It therefore poses a distinct, though possibly not unique,<sup>18</sup> theoretical question.

In this article, I will argue that self-exclusion cannot be justified according to a content-neutral, procedural approach to autonomy. Rather, self-exclusion can only be justified based on a normative position that, at least in the case of the self-excluder, abstention from gambling is preferable to participation in gambling.

## II. The appeal and underlying problem of self-exclusion

Amongst Johnny Cash's voluminous catalogue is a song entitled *When Uncle Bill Quit Dope*.<sup>19</sup> Though more obscure than some of his vaunted classics, Cash presents in this song a story of remarkable emotional and, indeed, philosophical depth.<sup>20</sup> He tells the story of a cocaine-addicted man who implores his sister, Louise, to lock him in an upstairs room and refuse to release him until "[he] don't want dope no more."<sup>21</sup> After fourteen arduous days and nights, Louise opens the door of Bill's makeshift prison cell to find him in a renewed state of health and happiness, having overcome his addiction to "that demon cocaine."<sup>22</sup> The covenant Cash describes between Uncle Bill and Louise is, in fact, a Ulysses pact, named for the legendary Greek hero known in Greek as Odysseus. Ulysses pacts derive their name from a particular incident in Homer's *Odyssey*. Ulysses, keen to

---

15. See McInnis et al, *supra* note 10.

16. See text accompanying *infra* note 89.

17. Gambling addiction is not, of course, *entirely* self-regarding, as it will inevitably affect those close to the sufferer, whether they be financially dependent on the sufferer or emotionally invested in them. Nonetheless, gambling addiction does not affect others as directly as the other examples I have provided.

18. I do not wish to preclude the applicability of the self-exclusion problem to other cases. For example, internet addiction and pornography addiction may be sufficiently analogous to gambling addiction for the theoretical question posed here to be relevant.

19. "Uncle Bill Quit Dope" (music) Deena Kaye Rose, USA IPI00762805625 (1970) registered.

20. This is far from the only instance of philosophical depth in Cash's musical catalogue. See John Huss & David Werther, eds, *Johnny Cash and Philosophy: The Burning Ring of Truth* (Open Court, 2008).

21. "Uncle Bill Quit Dope", *supra* note 19.

22. *Ibid.*

hear the Sirens' song but aware of the ruination that might accompany it, fills the ears of the sailors under his command with wax, temporarily deafening them, and instructs them to tie him to the ship's mast and not release him even if he later orders them to. Upon hearing the Sirens' enticements, Ulysses is overcome with temptation and protests to his men to release him but, relying on his prior instruction, they do not.<sup>23</sup> The comparison with self-exclusion agreements is obvious. In the same manner as Uncle Bill and Ulysses, the gambler acknowledges that they will face, at some future time, an urge to engage in behaviour they presently think imprudent. In order to forestall this future behaviour, they decide in advance to enlist others to prevent their subsequent indulgence. In other words, the gambler decides at one time ( $t^1$ ) that they should be prevented from gambling at a later time ( $t^2$ ) and requests (or, in jurisdictions where providing self-exclusion is mandatory, demands) that others (gambling operators and their staff) do so.

Johnny Cash is not alone in recognising the applicability of Ulysses pacts to addiction.<sup>24</sup> In *Ulysses and the Sirens*, for example, Jon Elster argues that humans exhibit an imperfect rationality.<sup>25</sup> Ulysses rationally knows that he should not succumb to the Sirens' enticements and would, if he were a perfectly rational actor, need no more than that knowledge to resist them. His capacity to act on this knowledge, however, is impaired by non-rational factors in his human constitution, such as the weakness of his will and overwhelming emotion. Elster elsewhere opines that addictive behaviour is "the most striking instance ... of ambivalence or weakness of will."<sup>26</sup> Because the addict often experiences the pattern of attempted recovery and relapse multiple times, they realise that they must "confront not only their addiction, but their inability to quit by simply deciding to do so."<sup>27</sup> Precommitment strategies, including Ulysses pacts, are a logical solution to this predicament. Inasmuch as addiction is caused by rationality-disrupting emotional factors, it is part of a strand of thought almost as old as Ulysses' own pact with his sailors—*akrasia*. In *Nicomachean Ethics*, Aristotle observes that an agent may act irrationally because they "have the rational principle but do not abide by it, since [they] are defeated by a weaker passion, and do not act without previous deliberation"<sup>28</sup> or "by reason of their quickness [or] the violence of their passions do not await the argument, because they are apt to follow their imagination."<sup>29</sup> This notion of an akratic

23. See Homer, *The Odyssey*, 2d ed by and translated by Albert Cook (WW Norton & Co, 1993) at 134-35.

24. For a detailed outline of scholarly accounts of the relationship between Ulysses pacts and addiction, see Kirsten Bell, "Thwarting the Diseased Will: Ulysses Contracts, the Self and Addiction" (2015) 39:3 *Culture, Medicine & Psychiatry* 380.

25. See Jon Elster, *Ulysses and the Sirens: Studies in Rationality and Irrationality* (Cambridge University Press, 1979) at 36-37 [Elster, *Ulysses and the Sirens*].

26. Jon Elster, *Ulysses Unbound: Studies in Rationality, Precommitment and Constraints* (Cambridge University Press, 2000) at 63.

27. *Ibid.* Indeed, one of the current diagnostic criteria for Gambling Disorder is having repeated, failed attempts to quit. See *DSM-5*, *supra* note 14 at 585.

28. Aristotle, *Nicomachean Ethics*, translated by WD Ross (Batoche Books, 1999) at 118. Aristotle refers to this form of *akrasia* as *astheneia*, which essentially describes weakness of the will.

29. *Ibid.* Aristotle calls this *propeteia*, which may be referred to as impetuosity. *Akrasia* is used twice in the Koine Greek New Testament, in Matthew 23:25 and 1 Corinthians 7:5. In the

(i.e. weak-willed or emotional and, ultimately, irrational) agent is useful in explaining addicts' relapse and understanding the value of Ulysses pacts in addressing it. Heather and Segal, echoing Elster, consider addiction to be a "paradigm case for the study of weakness of will."<sup>30</sup> They argue that the notion of an akratic addict refutes what they call the strong sense of the causal attribution hypothesis of addiction.<sup>31</sup> Advanced *inter alia* by Davies, this hypothesis contends that addiction is a myth and that the behaviours of putative addicts are as voluntary as any other personal choice.<sup>32</sup> Drawing on the work of Donald Davidson,<sup>33</sup> Heather and Segal explain the possible irrationality of an addict's relapse with the example of a smoker who is trying to quit. The smoker has, as Davies contends, a reason (*r*) to relapse, namely that it will bring immediate enjoyment. Knowing that immediate enjoyment is not the only relevant interest, however, the smoker has a more comprehensive reason (*r'*) not to relapse. *r'*, being a reason based on all the information presently available to the smoker, includes *r*. If the smoker relapses, this behaviour is akratic and irrational because it is motivated by an "insufficient" reason, *r*, in comparison to the overall better reason, *r'*, which subsumes *r*.<sup>34</sup> This outline of *akrasia* is easily reconcilable with the behaviours of gambling addicts. Repeated failed attempts to abstain from gambling are a diagnostic criterion for Gambling Disorder (GD).<sup>35</sup> More obviously, gambling addicts suffer from a range of rationality-diminishing factors, such as cravings, trance-like states, and even memory

---

English Standard Version (ESV), it is translated as 'self-indulgence' and 'lack of self-control', respectively. Elsewhere, in Romans 7:15, Paul the Apostle does not use the word *akrasia* but essentially describes the concept in an explanation of sin, writing, "For I do not understand my own actions. For I do not do what I want, but I do the very thing I hate" (ESV).

30. Nick Heather & Gabriel Segal, "Understanding Addiction: Donald Davidson and the Problem of *Akrasia*" (2013) 21:6 *Addiction Research and Theory* 445 at 448 [Heather & Segal, "Understanding Addiction"]. See also Nick Heather, "Addiction as a Form of *Akrasia*" in Nick Heather & Gabriel Segal, eds, *Addiction and Choice: Rethinking the Relationship* (Oxford University Press, 2017) 133; Nick Heather, "The Concept of *Akrasia* as the Foundation for a Dual Systems Theory of Addiction" (2020) 390 *Behavioural Brain Research*, DOI: <https://doi.org/10.1016/j.bbr.2020.112666>.
31. See Heather & Segal, "Understanding Addiction", *supra* note 30 at 446-51.
32. See generally John Booth Davies, *The Myth of Addiction*, 2d ed (Harwood Academic, 1997). Heather & Segal cite the following passage from the prologue of *The Myth of Addiction*: "[Most] people who use drugs do so for their own reasons, on purpose, because they like it, and because they find no adequate reason for not doing so; rather than because they fall prey to some addictive illness which removes their capacity for voluntary behaviour." *Ibid* at x.
33. See Donald Davidson, "How is Weakness of the Will Possible?" in Donald Davidson, ed, *Essays on Actions and Events* (Oxford University Press, 1980) 21.
34. Heather & Segal, "Understanding Addiction", *supra* note 30 at 449. Heather and Segal acknowledge one unsettled aspect of Davidson's theory of *akrasia*, *ibid* at 450. Specifically, Davidson appears to believe that the akratic agent maintains knowledge that the better reason for acting (in Heather and Segal's smoker example, *r'*) is in fact better, even while committing the irrational act (the act motivated by *r*, in Heather and Segal's example). This, they note, jars with other accounts which identify a preference reversal in the akratic agent, i.e., that the agent considers *r* to be better than *r'* while acting on *r*. This distinction, which somewhat reflects that between *astheneia* and *propeteia*, is not relevant to the present discussion.
35. See *DSM-5*, *supra* note 14 at 585.

blackouts.<sup>36</sup> Gambling Disorder is also strongly related to a high degree of impulsivity.<sup>37</sup> All this suggests that gambling addicts are likely to choose *r*, the inferior reason to gamble, over *r'*, the superior, all-things-considered reason to abstain.

In view of what I have outlined thus far, self-exclusion has obvious value as a mechanism to aid gambling addicts. However, many unresolved issues remain, and it is these issues which will form the remainder of this portion of the present article. Gambling self-exclusion differs in key ways from the Ulysses pacts successfully deployed by Uncle Bill and by Ulysses himself. The *ad hoc*, informal arrangement between Uncle Bill and Louise is distinguishable from the formal, legal arrangement self-exclusion represents. Indeed, Louise, despite her good intentions, is probably guilty of false imprisonment.<sup>38</sup> Gambling self-exclusion differs in an even starker and, for the purposes of this article, more important way from Ulysses' agreement with his sailors. Because of Ulysses' superior rank, paternalism is not a possible alternative to his Ulysses pact. Any sailor or group of sailors who tied Ulysses to the mast without his consent, whether present or prior, would be guilty of insubordination. This is not the case with a problem gambler who seeks to self-exclude. Paternalism, in the form of forced exclusion, is a possible, as distinct from being a desirable, option for policy-makers. If self-exclusion is to have a unique merit, it is, as I earlier posited, found in its basis in the self-excluder's autonomy.<sup>39</sup> While I will discuss autonomy in greater depth later, grounding self-exclusion, or any measure, in autonomy allows it to avoid the issues inherent in paternalism, in particular its condescension and the affirmative value judgement upon which it relies.

Criticisms that paternalism is condescending are well-founded and many, though they do not warrant full attention here. For Anderson, some forms of paternalism are tantamount to "telling citizens that they are too stupid to run their lives."<sup>40</sup> For Feinberg, the "arrogant and demeaning" practice of paternalism, whereby adults are treated as children by a state standing *in loco parentis*, threatens to deprive citizens of their decision-making power and infantilise

- 
36. See Tavares et al, *supra* note 14; EH Limbrick-Oldfield et al, "Neural Substrates of Cue Reactivity and Craving in Gambling Disorder" (2017) 7 *Translational Psychiatry*, DOI: <https://doi.org/10.1038/tp.2016.256>; Hedy Kober et al, "Brain Activity During Cocaine Craving and Gambling Urges: An fMRI Study" (2015) 41:2 *Neuropsychopharmacology* 628; Mikal Aasved, *The Biology of Gambling* (Charles C Thomas, 2003) at 138-39; Magdalen G Schluter & David C Hodgins, "Dissociative Experiences in Gambling Disorder" (2019) 6 *Current Addiction Reports* 34.
  37. For a recent meta-analysis of this relationship, see Konstantinos Ioannidis et al, "Impulsivity in Gambling Disorder and Problem Gambling: A Meta-analysis" (2019) 44:8 *Neuropsychopharmacology* 1354. Indeed, prior to its recategorisation in *DSM-5*, Gambling Disorder (then called Pathological Gambling) was categorised as an Impulse-Control Disorder, alongside disorders such as Trichotillomania and Pyromania.
  38. In some jurisdictions, operators and their staff are explicitly indemnified against any liability which might otherwise arise from their enforcement of a self-exclusion agreement. See for example *Casino Control Act 1992*, *supra* note 6, s 85(4).
  39. See Florian Wagner-von Papp, "Self-Exclusion Agreements: Should We Be Free Not to Be Free to Ruin Ourselves? Gambling, Self-Exclusion Agreements and the Brain" in Michael Freeman & Oliver R Goodenough, eds, *Law, Mind and Brain* (Ashgate, 2009) 81. Wagner-von Papp describes self-exclusion as "self-paternalistic," *ibid* at 81-82.
  40. Elizabeth S Anderson, "What Is the Point of Equality?" (1999) 109:2 *Ethics* 287 at 301, 330.

them.<sup>41</sup> Self-exclusion purports to avoid this issue. To be sure, a self-excluded gambler being refused service appears outwardly indistinguishable to one being paternalistically refused service. The former, however, avoids the condescension of the latter because the refusal is not justified, to use Feinberg's terminology, because "Daddy knows best"<sup>42</sup> but because the gambler knew best when they executed the self-exclusion agreement.

The second criticism of paternalism which self-exclusion purports to avoid is that it relies on an indefensible value-judgement. Forced exclusion, for example, relies on the paternalist's affirmative judgement that the excluded party should not gamble. The rightness of this judgement, critics argue, is indemonstrable and, ultimately, beyond the paternalist's limited knowledge. As Mill famously wrote, "[s]peaking generally, there is no one so fit to conduct any business, or to determine how or by whom it shall be conducted, as those who are personally interested in it."<sup>43</sup> Mill argues, in essence, that the prospective paternalised person is always a better judge of what is right for them than the paternalist. This is a highly suspect assumption. The above discussion of *akrasia* highlights one way in which a problem gambler is, at least, not a very good judge of the right course of action. The akratic gambler is driven by what Loewenstein calls 'visceral' influences, which impede their ability to make a rational decision.<sup>44</sup> In addition to visceral, emotional influences on decision-making, a wealth of research in behavioural economics has also demonstrated the pervasive role of cognitive biases in impeding rational decision-making. A previously dominant assumption in economic theory was the existence of a *homo economicus*, an agent who acts perfectly rationally and self-interestedly, based on complete information. More recently, however, a wealth of research in behavioural economics has severely undermined this assumption.<sup>45</sup> Rather, the rationality of decision-making can be severely disrupted by various cognitive biases, including status quo bias, the framing effect, and the use of representative heuristics.<sup>46</sup> This,

---

41. Joel Feinberg, *The Moral Limits of the Criminal Law: Harm to Self*, vol 3 (Oxford University Press, 1986) at 23-24.

42. *Ibid* at 23.

43. John Stuart Mill, *On Liberty* (Batoche Books, 2001) at 100.

44. See George Loewenstein, "Out of Control: Visceral Influences on Behavior" (1996) 65:3 *Organizational Behavior and Human Decision Processes* 272.

45. See Colin F Camerer & George Loewenstein, "Behavioural Economics: Past, Present, Future" in Colin F Camerer, George Loewenstein & Matthew Rabin, eds, *Advances in Behavioural Economics* (Princeton University Press, 2004) 3; Daniel Kahneman, Paul Slovic & Amos Tversky, eds, *Judgement Under Uncertainty: Heuristics and Biases* (Cambridge University Press, 1982); Daniel Kahneman, *Thinking, Fast and Slow* (Farrar, Straus & Giroux, 2011).

46. Research in this area is voluminous and is impossible to recount in its totality here. See for example Donald A Redelmeier, Paul Rozin & Daniel Kahneman, "Understanding Patients' Decisions: Cognitive and Emotional Perspectives" (1993) 270:1 *JAMA* 72; Amos Tversky & Daniel Kahneman, "The Framing of Decisions and the Psychology of Choice" (1981) 211:4481 *Science* 453; Daniel Kahneman & Amos Tversky, "Prospect Theory: An Analysis of Decision Under Risk" (1979) 47:2 *Econometrica* 263; Hideaki Takeuchi et al, "Framing Effects on Financial and Health Problems in Gambling Disorder" (2020) 110 *Addictive Behaviors*, DOI: <https://doi.org/10.1016/j.addbeh.2020.106502>; James J Choi et al, "Defined Contribution Pensions: Plan Rules, Participant Choices, and the Path of Least Resistance" (2002) 16 *Tax Policy and the Economy* 67; Brigitte C Madrian &

again, has significant application to problem gamblers. Inasmuch as problem gamblers are susceptible to akratic, visceral urges, they also exhibit a marked susceptibility to cognitive bias.<sup>47</sup> Although the ‘best judge’ argument against paternalism is undermined by the presence of rationality-subverting visceral urges and cognitive biases, this does not necessarily justify paternalism. Those in positions of authority, including would-be paternalists, are also imperfect judges. Research has demonstrated that civil servants, asylum judges, emergency managers, and holders of other similar positions are also susceptible to a range of cognitive biases.<sup>48</sup> Though Bastiat did not have access to this research, he may as well have when he wrote, “[s]ince the natural tendencies of mankind are so bad that it is not safe to allow them liberty, how comes it to pass that the tendencies of organizers are always good?”<sup>49</sup> The answer, of course, is that the tendencies of paternalists will not always be good or, at least, grounded in an accurate appreciation of the facts. The limitations of prospective paternalists surpass mere cognitive bias, however. Rather, the paternalist’s legitimacy is undermined by their lack of absolute knowledge of objectively good ends or, at least, the indemonstrability of those ends. To be sure, the relative susceptibility of paternalists and problem gamblers to the framing effect can be ascertained and compared, and the better judge discovered. Likewise, the accuracy of their respective knowledge of the odds of roulette can be demonstrated. The paternalist’s claim to knowledge of the objective rightness or wrongness of gambling (whether generally or in the case of a particular gambler) cannot be demonstrated in a way which would satisfy the evidentiary criteria of the liberal state.<sup>50</sup> The indemonstrability of

---

Dennis F Shea, “The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior” (2001) 116:4 *The Quarterly Journal of Economics* 1149; Amos Tversky & Daniel Kahneman, “Belief in the Law of Small Numbers” (1971) 76:2 *Psychological Bulletin* 105.

47. See Maria Ciccarelli et al, “Decision Making, Cognitive Distortions and Emotional Distress: A Comparison between Pathological Gamblers and Healthy Controls” (2017) 54 *Journal of Behavior Therapy and Experimental Psychiatry* 204; Tony Toneatto, “Cognitive Psychopathology of Problem Gambling” (1999) 34:11 *Substance Use & Misuse* 1593; Gabrielle Atkins & Louise Sharpe, “Cognitive Biases in Problem Gambling” (2003) 15:2 *Gambling Research: Journal of the National Association for Gambling Studies* 35; Tony Toneatto et al, “Cognitive Distortions in Heavy Gambling” (1997) 13:3 *Journal of Gambling Studies* 253; Erica E Fortune & Adam S Goodie, “Cognitive Distortions as a Component and Treatment Focus of Pathological Gambling: A Review” (2012) 26:2 *Psychology of Addictive Behaviors* 298.
48. See Nicola Bellé, Paola Belardinelli & Paolo Cantarelli, “Prospect Theory Goes Public: Experimental Evidence on Cognitive Biases in Public Policy and Management Decisions” (2018) 78:6 *Public Administration Review* 828; Daniel L Chen, Tobias J Moskowitz & Kelly Shue, “Decision Making Under the Gambler’s Fallacy: Evidence From Asylum Judges, Loan Officers, and Baseball Umpires” (2016) 131:3 *Quarterly Journal of Economics* 1181; Patrick S Roberts & Kris Wernstedt, “Decision Biases and Heuristics Among Emergency Managers: Just Like the Public They Manage For?” (2019) 49:3 *American Review of Public Administration* 292.
49. Frédéric Bastiat, *The Law* (Ludwig von Mises Institute, 2007) at 46.
50. This formed the basis of one of Locke’s arguments in favour of religious toleration. Absolute sovereigns, though superior in power, are nonetheless equal in nature to other people. They thus can no more determine with surety the correct path to salvation than any other person, and therefore have no standing to interfere with each citizen’s religious beliefs. See John Locke, *A Letter Concerning Toleration* (Prometheus Books, 1990) at 37.

objectively good ends is why Le Grand and New reject what they call ends-related or moral paternalism.<sup>51</sup> Because objectively good ends cannot be supported by empirical evidence, ends-related paternalism will invariably rely on the value judgements of the paternalist.

If self-exclusion is to be prized for its basis in the gambler's autonomy, it must avoid these objections. The objection from condescension is presumably mitigated to some degree by the problem gambler's prior indication at  $t^1$  that they ought not gamble at  $t^2$ . Those enlisted at  $t^2$  to prevent the self-excluder from gambling are not telling the self-excluder that "they are too stupid to run their [life]" but are instead saying that the self-excluder is wise enough to run their life and did so at  $t^1$ .<sup>52</sup> On the other hand, the objection from value-partiality—and with it the self-exclusion problem—remain. Those who would prevent a  $t^2$  self-excluder from gambling are faced with the choice of whether or not to refuse service.<sup>53</sup> If self-exclusion is to regard the self-excluder's autonomy, those enforcing it must act on the self-excluder's preferences and not their own. However, the self-excluder has two conflicting preferences. The  $t^1$  preference is for the gambler to be refused service at  $t^2$ , while the  $t^2$  preference is for the gambler to be allowed to gamble at  $t^2$ . If self-exclusion is to exist at all, the  $t^1$  preference must, of course, take precedence and the  $t^2$  preference must be discarded. How can this precedence be

---

51. See Julian Le Grand & Bill New, *Government Paternalism: Nanny State or Helpful Friend?* (Princeton University Press, 2015) at 102-03. Le Grand and New thus favour means-related paternalism, where the paternalist may paternalise, based on empirical evidence, the means by which an agent pursues the agent's chosen ends. Of course, the demand for empirical evidence is also a value judgement. As Ronald Dworkin observes, arguments within a practice can only be engaged by arguments that are also within that practice. In other words, scientific arguments can only be engaged by scientific arguments and moral arguments by moral arguments. See Ronald Dworkin, "On Interpretation and Objectivity" in Ronald Dworkin, ed, *A Matter of Principle* (Harvard University Press, 1985) 167 at 170-76. Demanding that a moral argument be scientifically or empirically demonstrable is as futile as demanding that a scientific argument be morally demonstrable. Indeed, the assumption that a paternalist should not interfere with an agent's chosen ends, when both have equally limited knowledge, is itself a moral or value judgement unsupported by empirical evidence.

52. This, however, may not be satisfactory. While preventing the  $t^2$  gambler from gambling may not condescend on the grounds that the gambler cannot make a wise decision about whether or not to gamble, it may condescend on the grounds that the gambler is unable to adhere to their prior decision. It also presumably condescends insofar as it precludes the self-excluder from updating their preferences. One challenge to the notion that paternalism (or, by extension, the  $t^2$  portion of a self-exclusion agreement) is condescending is the universality of human fallibility. See Sarah Conly, *Against Autonomy: Justifying Coercive Paternalism* (Cambridge University Press, 2012) at 2. This, however, is insufficient to extinguish the objection from condescension. The father who tells his child to go to bed because "Daddy knows best" is, of course, fallible; he may drink too much or drive too fast. Likewise, the paternalist who prohibits a gambling addict from gambling is also fallible. However, these acts are still condescending because the relationships are non-reciprocal. The child and the gambling addict cannot paternalise the fallibilities of the father and paternalist, respectively. In any case, the potential condescension of self-exclusion agreements is not the emphasis of this article and I will leave this issue open.

53. "Those who would prevent a  $t^2$  self-excluder from gambling" ought to be interpreted broadly. The operator and its staff are tasked in the immediate term with refusing service and, if applicable, removing the self-excluder. Their motivation, however, derives at least in part from a desire to avoid punishment. As such, I mean here not just operators and their staff but the entire structure which supports the enforcement of self-exclusion, including legislators, regulators, courts, etc.

justified, however? It cannot, for the reasons I have outlined, be because the  $t^1$  preference is ‘correct.’ Another justification might be an arbitrary favour for earlier decisions over later ones. This is also unfeasible because, as much as it would enable the operation of self-exclusion, it would equally justify the enforcement of agreements by gamblers never to self-exclude, perhaps included in the terms and conditions of gambling contracts. An ostensibly more meritorious justification is to find some way in which the  $t^2$  decision fails to satisfy the requirements of an autonomous decision. In other words, if the  $t^2$  decision-maker is not fully capable to decide to gamble, the  $t^2$  decision can be discarded and the  $t^1$  decision acted on.

Capacity is, after all, the “gatekeeper of autonomy,”<sup>54</sup> and even Mill, the arch-libertarian of the nineteenth century, withheld autonomy from those not “in the maturity of their faculties” and those who are “delirious, or in some state of excitement or absorption incompatible with the full use of the reflecting faculty.”<sup>55</sup> The visceral urges which bear on the akratic problem gambler could easily be seen as making their decision to gamble less-than-rational, less-than-autonomous, or less-than-capable. Nonetheless, the concept of legal capacity is fraught with complexity and is ever-changing.<sup>56</sup> If capacity is to be equated, rudimentarily, with “understanding,”<sup>57</sup> then its applicability to Gambling Disorder is contestable. GD sufferers are not *tout court* incapable in a general sense. A recent study, for example, found that GD sufferers are highly susceptible to the effect of positive framing in the context of financial decision-making.<sup>58</sup> With regard to healthcare-related decisions, however, sufferers were no more susceptible to the framing effect than the control group. On the other hand, a functional approach to capacity, as adopted in the *Mental Capacity Act 2005* in England and Wales, does not demand general incapacity.<sup>59</sup> Rather, a person may lack capacity in certain spheres and enjoy it in others. A GD sufferer may therefore lack capacity to enter gambling contracts but have full capacity to enter a mortgage contract or choose whether or not to undergo medical treatment. The question of whether a problem gambler has capacity, however, is only relevant to justifying forced exclusion. In the realm of self-exclusion, it is beside the point. If the  $t^2$  decision is made by an incapable agent and therefore non-autonomous, then the  $t^1$  decision is irrelevant and the incapable gambler ought to be prevented from gambling whether they expressed a contrary prior preference or not. If, as I have

---

54. Mary Donnelly, *Healthcare Decision-Making and the Law: Autonomy, Capacity and the Limits of Liberalism* (Cambridge University Press 2010) at 90.

55. Mill, *supra* note 43 at 14, 88. Mill did not use the word ‘autonomy’ but, rather, ‘liberty’. O’Neill speculates that this is a deliberate attempt on Mill’s part to distance his theory from that of Kant. See Onora O’Neill, *Autonomy and Trust in Bioethics* (Cambridge University Press, 2002) at 30. Nonetheless, Mill’s theory is at least an unmistakable precursor to later procedural theories of autonomy, which I will discuss later.

56. For a recent consideration of the evolution of capacity law and the normative ideas which underlie it, see Mary Donnelly, “Changing Values and Growing Expectations: The Evolution of Capacity Law” (2017) 70:1 *Current Leg Probs* 305.

57. Donnelly, *supra* note 54 at 95-96.

58. See Takeuchi et al, *supra* note 46.

59. See Mary Donnelly, “Capacity Assessment under the *Mental Capacity Act 2005*: Delivering on the Functional Approach?” (2009) 29:3 *LS* 464.

already said, the allure of self-exclusion is its basis in autonomy, the  $t^1$  preference is crucial and cannot be rendered irrelevant to the refusal of service at  $t^2$ .

Before proceeding to the next portion of this article, it is worth restating what has been discussed thus far. Self-exclusion, as a Ulysses pact, has obvious appeal in addressing problem gambling. The problem gambler cannot, in many cases, abstain from gambling by sheer force of will. Rather, rationality-impugning forces may cause the gambler's commitment to their abstinence to waver. In anticipation of their future wavering, the gambler can enlist others to prevent them from renegeing on their earlier commitment. If self-exclusion is meritorious in a way that is distinct from forced exclusion, however, it must be because it emphasises the self-excluder's autonomy. If this is the case, the  $t^2$  refusal of service cannot be based on a value judgement of any party other than the self-excluder, nor can it be because the  $t^2$  decision to gamble is *prima facie* discardable. Simply put, the  $t^2$  refusal of service can *only* occur because of the  $t^1$  decision to self-exclude.

### III. Self-exclusion and the inadequacy of procedural autonomy

A central assumption of this article is that self-exclusion is prized because it values the problem gambler's autonomy, in contrast to paternalistic forced exclusion, which does not. The long and complex history of the term 'autonomy' is by now familiar to many and would be impossible to recount in full here. Etymologically, the term derives from the Ancient Greek *auto*, meaning self, and *nomos*, meaning law. To be autonomous in its simplest form, then, means to make one's own law, to be governed by rules one has developed for oneself. As Gerald Dworkin notes, the term originally applied to Greek city-states that made their own laws free of external interference, before its "natural extension" to individuals as being autonomous when they are self-determining.<sup>60</sup> Despite its common root, however, the term now encompasses a variety of different and often conflicting meanings. Returning to Gerald Dworkin, he writes that, "the only features held constant from one author to another are that autonomy is a feature of persons and that it is a desirable quality to have."<sup>61</sup> One useful way to divide the vast range of theories is into the categories of substantive, which place direct constraints on the content of decisions for them to count as autonomous, and procedural, which require only that decisions, whatever their content, are arrived at in a procedurally independent way. In this article, I take autonomy to refer to a procedural theory. First, I believe simply that this is what is instinctively understood as the meaning of autonomy by anyone not steeped in its voluminous theoretical treatment, including gambling operators and their staff, gambling regulators, legislators, and prospective self-excluders themselves.

---

60. Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge University Press, 1988) at 12-13 [Dworkin, *Theory and Practice*].

61. *Ibid* at 6. Cf O'Neill, *supra* note 55 at 22. O'Neill doubts that even this broad assertion is universally held.

Self-exclusion is a fundamentally practical exercise involving two parties, the self-excluder and those who would enforce self-exclusion at  $t^2$ . As such, its theoretical underpinnings must be such that they are mutually understood by both parties. Second, a substantive theory of autonomy that considers the  $t^2$  decision to gamble as necessarily heteronomous nullifies the  $t^1$  decision in a way I have earlier rejected.<sup>62</sup> If the  $t^2$  decision is declared heteronomous on a substantive account, the prospective gambler is excluded for reasons other than their  $t^1$  decision, which deprives self-exclusion of its apparently unique value. Finally, although this article is not intended as a refutation of substantive autonomy, questions linger about whether substantive theories can be properly categorised as autonomy at all. Specifically, the value judgements inherent in these theories may be entirely alien to the agent whose autonomy they purport to represent. In the extreme, we see theories such as Babbitt's.<sup>63</sup> She argues that the liberal view, which bases an agent's autonomy on their "untransformed" self, is inadequate in cases where an agent does not have a self which would choose a "full sense of autonomy," even in idealised conditions.<sup>64</sup> Rather, Babbitt argues that the agent's sense of self must be transformed into one which would endorse a full sense of autonomy, a contention with all the disconcerting implications of Rousseau's plan for the citizens of his envisioned state.<sup>65</sup> What Babbitt advocates cannot, of course, properly be called autonomy, as it proposes the destruction of the agent's self (*auto*) and the imposition of another's law (*nomos*) on the agent. Any theory of autonomy in this vein is, in fact, anathema to autonomy in its proper form and can be more accurately described as heteronomy or xenonomy.<sup>66</sup>

Procedural autonomy is content-neutral.<sup>67</sup> It does not place direct constraints on the substance or content of decisions, requiring only that an agent's decisions "express her will, and they express her will because the dispositions which they

---

62. Kant coined the term "heteronomy" to mean actions driven by alien causes and which are characteristic of the sensible world of "beings lacking reason" who are motivated by "natural necessity." Immanuel Kant, *Groundwork for the Metaphysics of Morals*, ed by and translated by Allen W Wood (Yale University Press, 2002) at Ak 4:441-46. Conversely, Kant considers autonomous action to be that which is "effective independently of alien causes determining it" and which is characteristic of the noumenal, ideal world. Even removed from the context of Kant's transcendental idealism, heteronomy is typically used as the opposite of autonomy. *Ibid* at Ak 4:447 [emphasis removed].

63. See Susan E Babbitt, "Feminism and Objective Interests: The Role of Transformation Experiences in Rational Deliberation" in Diana Tietjens Meyers, ed, *Feminist Social Thought: A Reader* (Routledge, 1997) 368 at 373.

64. *Ibid*.

65. See Jean-Jacques Rousseau, *The Social Contract* (Oxford University Press, 1994) at 76. Rousseau's citizens are to be surreptitiously influenced by a figure he calls the "legislator," who must "feel himself capable of changing, so to speak, the nature of man; of transforming each individual" (*ibid*).

66. Or, more plainly, indoctrination or brainwashing.

67. See John Christman, "Procedural Autonomy and Liberal Legitimacy" in James Stacey Taylor, ed, *Personal Autonomy: New Essays on Personal Autonomy and its Role in Contemporary Moral Philosophy* (Cambridge University Press, 2005) 277 at 280. Friedman, in fact, considers the term "content-neutral" more perspicuous than "procedural" and uses the former rather than the latter. Marilyn Friedman, *Autonomy, Gender, Politics* (Oxford University Press, 2003) at 19.

put into effect are hers in some important sense.”<sup>68</sup> The demand that the dispositions which inform a decision are the agent’s own gives effect to the most literal meaning of autonomy, that of *self-imposed* law. It disqualifies actions which are motivated by dispositions arising from others, a category which Gerald Dworkin takes to include hypnosis, manipulation, coercion, and subliminal influence.<sup>69</sup> Insofar as procedural autonomy requires freedom from others’ influence, it echoes the negative liberty of Berlin’s dichotomy, which he defines simply as being free from interference by others with the range of activities that are within one’s capacity.<sup>70</sup> Nonetheless, mere freedom from external interference fails to fully explain the demands of procedural autonomy. Indeed, this sort of *de facto* freedom would fail to preclude an agent’s life being a turbulent series of disjointed and whimsical decisions.<sup>71</sup> In other words, it would focus narrowly on the *auto* and remain wholly ignorant of the *nomos*; it would promote an anarchy of the self rather than a governed self. As such, procedural conceptions of autonomy tend to demand not just freedom from interference but a degree of *de facto* self-government to which *de jure* liberty is blind.<sup>72</sup> The agent is generally not considered to be autonomous simply because they act free from external constraints, but because they, in some sense, ‘own’ or ‘identify with’ their decision. Traces of this requirement are evident in Mill’s *On Liberty*, where he bemoans that individuals ask themselves, “What is usually done by persons of my station and pecuniary circumstances?” rather than “What would suit my character and disposition?”<sup>73</sup> The question of determining whether a person’s decisions are truly reflective of their “character and disposition” has abided since Mill first identified it. One attempt to resolve this issue is with a structural, or hierarchical, approach to preferences.<sup>74</sup> For Gerald Dworkin, autonomy involves not just first-order preferences but also “preferences about . . . preferences, a desire not to have or act upon various desires.”<sup>75</sup> Autonomy, then, constitutes

---

68. Neil Levy, “Autonomy and Addiction” (2006) 36:3 *Canadian Journal of Philosophy* 427 at 429.

69. See Dworkin, *Theory and Practice*, *supra* note 60 at 18.

70. See Isaiah Berlin, “Two Concepts of Liberty” in Nigel Warburton, ed, *Philosophy: Basic Readings* (Routledge, 1999) 180.

71. As Spinoza wrote, “If all things therefore were dependent on the inconstant will of one person there would be nothing fixed.” Baruch Spinoza, *A Treatise on Politics*, translated by William MacCall (Hollyoake & Co, 1854) at 53.

72. On this distinction, see Feinberg, *supra* note 41 at 62-68.

73. Mill, *supra* note 43 at 57.

74. For two significant accounts, see Harry G Frankfurt, “Freedom of the Will and the Concept of a Person” in Harry G Frankfurt, ed, *The Importance of What We Care About: Philosophical Essays* (Cambridge University Press, 1988) 11; Gerald Dworkin, “Autonomy and Behavior Control” (1976) 6:1 *Hastings Center Report* 23 [Dworkin, “Behavior Control”]. Dworkin revisits and amends his hierarchical theory of preferences in Dworkin, *Theory and Practice*, *supra* note 60 at 14-20. The hierarchical approach bears some similarity to the view of *akrasia* offered by Heather & Segal, “Understanding Addiction”, *supra* note 30. Indeed, they refer to Frankfurt’s hierarchical theory and acknowledge its similarity to their view of *akrasia* (*ibid* at 449).

75. Dworkin, *Theory and Practice*, *supra* note 60 at 15.

a second-order capacity of persons to reflect critically upon their first-order preferences . . . and the capacity to accept or attempt to change these in light of higher-order preferences and values.<sup>76</sup>

The hierarchical theory of preferences Dworkin advances here is not the only such theory or even, as I have observed in the footnotes, the only such theory advanced by Dworkin himself. Nonetheless, the nuances which separate these various theories are largely irrelevant here and I provide Dworkin's above theory for exemplary purposes only. What is relevant is the distinction between first- and second-order preferences and the requirement that the latter be regarded in some sense for an action to be autonomous. The structural approach appears, at first glance, to offer a solution to the self-exclusion problem. The  $t^1$  decision to self-exclude can be neatly equated with a second-order preference not to gamble and the  $t^2$  decision to gamble can then be categorised as a heteronomous first-order preference because it fails to cohere with or otherwise properly regard the superior second-order preference.<sup>77</sup> This avoids the issue of value-partiality I objected to in other theories because both the first- and second-order preferences emanate from the agent and not another party.

Hierarchical theories of preferences are not without issue, however. First, as much as the validity of first-order preferences must be ascertained with reference to second-order preferences, the validity of the second-order preferences must also be determined. This can presumably only be accomplished with reference to third-order preferences, which must be measured against even higher-order preferences. This introduces the problem of "infinite regress."<sup>78</sup> Writing in a different but still relevant context, Elster provides the following example to question the reality of a simple two-preference model:

I wish to eat cream cake because I like it. I wish that I didn't like it, because, as a moderately vain person, I think it is more important to remain slim. But I wish I was less vain. (But do I think that only when I wish to eat cake?)<sup>79</sup>

A simple two-preference model fails to explain the predicament of the prospective cake eater. The first-order preference is to eat the cake, which is heteronomous because of the second-order preference to remain slim. However, this is, in turn, heteronomous because of a third-order preference to reject vanity. Elster seems to suggest this third-order preference is unreliable because it is possibly

---

76. *Ibid* at 20. In "Behavior Control", Dworkin argued that autonomy required congruence between second-order identifications and first-order motivations, a state that he called "authenticity." Dworkin, "Behavior Control", *supra* note 74 at 24. By *Theory and Practice*, he had disavowed this view in favour of considering autonomy to be the *capacity* to "raise the question of whether I will identify with or reject the reasons for which I now act." Dworkin, *Theory and Practice*, *supra* note 60 at 15.

77. Dworkin endorses the idea that his theory justifies Ulysses pacts, writing, "In limiting [Odysseus'] liberty, in accordance with his wishes, we promote, not hinder, his efforts to define the contours of his life." Dworkin, *Theory and Practice*, *supra* note 60 at 15.

78. Laura Waddell Ekstrom, "A Coherence Theory of Autonomy" (1993) 53:3 *Philosophy and Phenomenological Research* 599 at 601-02.

79. Jon Elster, *Nuts and Bolts for the Social Sciences* (Cambridge University Press, 1989) at 37.

situational and self-deceptive. (Although this may be framed as a fourth-order preference to make consistent, rational decisions.) Frankfurt, another proponent of a hierarchical model, acknowledges this potential shortcoming but proposes that a non-arbitrary tourniquet can be placed on the infinite regress.<sup>80</sup> The infinite regress can, he says, be terminated when the agent identifies “decisively” with a first-order preference and this “commitment ‘resounds’ throughout the potentially endless array of higher orders.”<sup>81</sup> He provides the example of a person who “without reservation or conflict, wants to be motivated by the desire to concentrate on his work.”<sup>82</sup> Because his second-order volition to be so motivated is “decisive,” he can “properly insist that this question concerning a third-order desire does not arise.”<sup>83</sup> Whether this is truly non-arbitrary, however, is deeply contestable.<sup>84</sup> If Frankfurt is suggesting that the force of identification is sufficient to forestall subsequent orders of preference, it must be asked how force is to be measured. The force of a preference can only be mitigated by other preferences, not by factors internal to that preference. For example, two students may have respective second-order preferences to achieve good results. The force of these respective preferences must be assumed to be equal. If the force of one student’s preferences is sufficient to preclude higher-order preferences, it is because that student also has a first-order preference to study diligently. If the preferences of the other student are not sufficiently forceful, it is because that student has a first-order preference to procrastinate. Inevitable arbitrariness arises in determining how many orders of preference must be coherent before further investigation becomes unnecessary. Even if this is not what Frankfurt suggests, a greater problem remains. The fact that a commitment must resound “throughout the potentially endless array of higher orders” admits that these higher orders exist. The fact that an agent’s lower-order preference can resound so endlessly simply speaks to an agent whose preferences are so consistent that they *would* remain entirely monolithic through infinite orders. If agents that are non-fragmented in this way exist, they are surely in a minority. In any case, I do not believe that such an agent—invariably impervious to inertia, lethargy, distraction, etc.—exists.

Gerald Dworkin adopts a different tack to Frankfurt in dealing with the problem of infinite regress.<sup>85</sup> He claims to analyse autonomous *persons* rather than autonomous *acts*. As such, a person is autonomous if they have the capacity for critical reflection and exercise it, so long as this exercise is itself free from third-party interference and the person possesses the requisite identification with their first-order preference. There is, Dworkin says, no “conceptual necessity” for questioning whether second-order preferences are themselves justified by still

---

80. See Frankfurt, *supra* note 74 at 21-22.

81. *Ibid* at 21.

82. *Ibid*.

83. *Ibid*.

84. See for example Adrian MS Piper, “Two Conceptions of the Self” (1985) 48:2 *Philosophical Studies* 173 at 176.

85. See Dworkin, *Theory and Practice*, *supra* note 60 at 19-20.

higher-order preferences.<sup>86</sup> In fact, Dworkin abrogates the question. Of course, his theory for gauging the general autonomy of a person may have some broad value in determining if, for example, a person with anterograde amnesia is autonomous.<sup>87</sup> The self-exclusion problem, however, contains far more nuance than simply determining whether a person is *tout court* autonomous. As discussed previously, there is reason to believe that the irrationality of problem gamblers is domain-specific.<sup>88</sup> Adopting a Dworkinian approach to autonomy might categorise a problem gambler as an autonomous person but ignore that, with regard to gambling, they cannot or do not engage in critical reflection.

Even leaving aside these issues with the respective theories of Frankfurt and Dworkin, the hierarchical approach is beset by an even more fatal issue: its impracticability. As an abstract principle of philosophy, the hierarchical approach might be meritorious, notwithstanding the problem of infinite regress. As a practical solution to the self-exclusion problem, however, it is not. It might be tempting to assume that the  $t^1$  decision (or any other binding decision) is rational and/or representative of a higher-order preference and that the  $t^2$  decision (or any other bound decision) is irrational and/or representative of a mere first-order preference. This, however, is a third-party assumption which is not necessarily correct. A precommitting decision may, as Elster observes, be no more rational or dispassionate than the later decision the precommitment seeks to restrain. Elster gives the example of covenant marriage, a form of marriage which is more difficult to enter and dissolve than ordinary marriages. An engaged person might be motivated by a number of non-rational or first-order preferences to opt for a covenant marriage, such as intense infatuation or a reluctance to imply to their prospective spouse that their commitment is not absolute.<sup>89</sup> In this instance, a first-order preference unsupported by second-order reflection might bind a later first-order preference (to divorce) which is supported by a second-order preference (self-respect or the value of marital faithfulness) when the person's spouse commits infidelity. This danger is also present in the case of self-exclusion. A gambler might self-exclude, for example, because of the admonitions or threats of divorce from their spouse, a first-order decision that jars with their second-order preference to be more independent, to live a libertine life of indulgence, or simply to be a gambler. This heteronomous decision would bind the agent's later autonomous decision to gamble. As deplorable as this person's skewed priorities are, it is nonetheless the case that these priorities are their own. The point here is not that this is a regular occurrence but that, on the rare occasion it does occur, those enlisted to realise the Ulysses pact cannot be expected to discern the complex motivations which are obscured by the ostensibly simple self-exclusion agreement. In these instances, self-exclusion would become a

---

86. *Ibid* at 20.

87. Anterograde amnesia limits the sufferer's ability to form new memories after its onset. I have in mind a sufferer whose case is so severe that they can never remember their reasons for any of their actions, which would make second-order reflection impossible.

88. See Takeuchi et al, *supra* note 46.

89. See Elster, *supra* note 26 at 19-20.

vessel for the preservation of heteronomy, rather than a vessel for the realisation of autonomy.

The hierarchical approach cannot, I have argued, resolve the self-exclusion problem and, as such, I will now turn to an alternative approach to determining identification with a decision within a procedural framework of autonomy: the historical approach. Christman, an important proponent of a historical approach, contends that autonomy relies on “what the agent thinks about the *process* of coming to have the desire, and whether she resists that process when (or if) given the chance.”<sup>90</sup> On this conception, the autonomy of a person’s decision is not contingent on some other, possibly heteronomous second-order preference. Rather, it relies on a historical construction of the self, in which one has developed in a procedurally independent way. Christman later adds the condition of non-alienation. To be alienated is to “experience strong negative affect relative to that characteristic—to disapprove in some manner—and to resist whatever motivational force it may have.”<sup>91</sup> In essence, the agent is non-autonomous if aspects of their character would produce feelings of alienation and be rejected by the agent who critically reflected on *how they came about*. While it is not worthwhile to engage with all the minutiae of Christman’s theory here,<sup>92</sup> I will now argue that it is beset by issues which make it equally as unable to solve the self-exclusion problem as the hierarchical approaches discussed earlier. One forceful challenge to Christman’s theory comes from Oshana, who questions whether alienation, as described by Christman, represents a failure of autonomy.<sup>93</sup> She outlines the case of a reluctant careerist, who leaves their child with a childminder while they pursue full-time employment. This desire, Oshana illustrates, arises from an envious motivation to emulate the luxurious lives and conspicuous consumption of others. The reluctant careerist does not revise or reject that lifestyle but does abhor its origin in a selfish desire. Oshana argues that Christman must consider this agent non-autonomous because of the sense of alienation which arises from how their lifestyle began. She refutes the charge of heteronomy, however. The reluctant careerist may have made a choice with which they are uncomfortable, but they did not suffer from any cognitive failure in doing so. Oshana perhaps mischaracterises Christman’s point, however. Christman’s threshold for non-alienation is quite high. He contends that the agent must feel “no affinity” with the relevant aspect of themselves.<sup>94</sup> The reluctant careerist, however, plainly has some affinity with their lifestyle; after all, they continue in it. On an all-things-considered judgement, the agent might feel distaste for some element of their motivation but ultimately has concluded that they identify with their choice.

90. John Christman, “Autonomy and Personal History” (1991) 21:1 *Canadian Journal of Philosophy* 1 at 10 [emphasis in original]. For an alternative historical approach, see Alfred Mele, *Autonomous Agents: From Self-Control to Autonomy* (Oxford University Press, 1995).

91. Christman, *supra* note 67 at 279.

92. For an updated, comprehensive account of Christman’s theory, see John Christman, *The Politics of Persons: Individual Autonomy and Socio-Historical Selves* (Cambridge University Press, 2009).

93. See Marina Oshana, *Personal Autonomy in Society* (Ashgate, 2006) at 37.

94. Christman, *supra* note 67 at 280.

A more forceful challenge to Christman's account lies in alienating aspects which cannot be undone. *In extremis*, we might look to natural limitations. As Nozick writes, "I may voluntarily walk to someplace I would prefer to fly to unaided."<sup>95</sup> If I feel alienated by the inability to fly that comes with my birth as an ordinary human being, am I heteronomous? Plainly, this is untenable. Leaving aside this fantastic example, we might take the victim of a road traffic accident who has had a leg amputated as a result. The person might feel alienated by their accident, which certainly irretrievably limits their scope of choice in many or most aspects of their life. Must it be concluded that this person, on Christman's understanding, has permanently and irreversibly been rendered heteronomous? Physical disability obviously poses unique challenges for theories of autonomy and therefore may be inappropriate grounds for objection to the historical approach. Given that the focus of this article is Gambling Disorder, what of psychiatric disorders? Take the example of a person suffering from Avoidant Personality Disorder (AVPD) as a result of non-ideal childhood experiences.<sup>96</sup> The sufferer was too young to express any actual resistance to the traumatic experience which led to their AVPD and, as such, heteronomy cannot be ascertained on these grounds. Like the amputee, the sufferer's traumatic experience cannot be undone, though its effects perhaps can. However, even if the person undergoes treatment to address their AVPD, their subsequent choices are those of a person who has undergone an autonomy-distorting event and psychiatric treatment to address it, not those of a person who has not undergone an autonomy-distorting event. Their choices can never be fully extricated from their traumatic experience. Does a historical approach therefore demand that a person who undergoes an unresisted, traumatic experience be forever categorised as heteronomous?

Christman, to his credit, addresses this concern. He writes that an assessment may need to be hypothetical, in the case of an agent who did not actually resist the development of a desire as it happened but "*would have* done so under conditions that make this possible."<sup>97</sup> Here, however, Christman veers inevitably into the substantive. Procedural autonomy is distinguished by its indifference to the content of decisions. If the autonomy or heteronomy of a decision is to be determined not by an agent's actual circumstances and characteristics but by a hypothetical judgement, this can surely only be determined by whether a hypothetical agent would make a decision with the specific content at hand. This is an inevitably partial judgement and does not emanate, as procedural autonomy demands, from the actual agent. Whatever of the substantivity of Christman's theory, its applicability to the self-exclusion problem is precluded by the same issue as hierarchical theories: its impracticability. For a historical approach to justify

---

95. Robert Nozick, *Anarchy, State, and Utopia* (Basic Books, 1974) at 262.

96. See David C Rettew et al, "Childhood Antecedents of Avoidant Personality Disorder: A Retrospective Study" (2003) 42:9 *Journal of the American Academy of Child and Adolescent Psychiatry* 1122.

97. Christman, *supra* note 90 at 11 [emphasis in original].

self-exclusion, it would have to hold the  $t^1$  decision to self-exclude as the historical, non-alienated and therefore autonomous preference and the  $t^2$  decision to gamble as alienated and therefore heteronomous. Of course, a gambling operator and its staff lack the relevant information to make this determination. In the case of a hypothetical judgement, they are even less able to determine what the self-excluder before them *would have done*. (In fact, no one is so able.)

In sum, neither a hierarchical nor a historical approach to preferences can resolve the self-exclusion problem. As a matter of abstract philosophy, either approach might theoretically justify self-exclusion by considering the  $t^1$  decision autonomous and the  $t^2$  decision heteronomous. In practice, however, the opacity of the self-excluder's thought processes makes it impossible for operators and their staff—as enlistees to enforce the Ulysses pact—to reliably make this judgement. It must then be assumed, for practical purposes, that both the  $t^1$  and  $t^2$  decisions are equally autonomous.

#### IV. Proposed solutions to the self-exclusion problem

Thus far, I have argued that the self-exclusion problem cannot be resolved by considering the  $t^2$  decision heteronomous because of a failure of rationality or, more generally, capacity. This approach would deprive the  $t^1$  decision of any gravitas, which would fail to justify self-exclusion as a distinct, autonomy-based alternative to paternalistic, forced exclusion. Hierarchical and historical approaches to autonomy purport to solve this issue by providing an avenue by which the  $t^2$  decision can be categorised as heteronomous relative to the agent's own preferences. At the theoretical level, these approaches' merit is contestable. As regards the self-exclusion problem, their impracticability makes them irrelevant. As such, the self-exclusion problem remains, and we are left with two conflicting preferences of presumptively equal autonomy. I will now consider two attempts to resolve the issue of conflicting, autonomous preferences.

Davis succinctly explains the issue with what I have called the self-exclusion problem, although he does not refer to it as such or write in the context of problem gambling.<sup>98</sup> He explains that self-binding in this way conflicts with what he calls the “Current Preference Thesis”, i.e., that respecting an individual's autonomy means only respecting their *current* preferences.<sup>99</sup> Accepting the Current Preference Thesis at face value would make self-exclusion impossible. At  $t^2$ , the pathological gambler's preference would be to gamble, this ought to be respected and the self-exclusion agreement revoked or ignored. Davis' solution to the existence of conflicting preferences—at  $t^1$  and  $t^2$ —is to identify a third preference, which he calls the “resolution preference” and which is a preference expressed by the individual on how to resolve a foreseen conflict of

---

98. See John K Davis, “Precedent Autonomy and Subsequent Consent” (2004) 7:3 Ethical Theory and Moral Practice 267.

99. *Ibid* at 267.

preferences.<sup>100</sup> This initially appears appealing in resolving the self-exclusion problem. The gambler has two conflicting preferences at different points in time: at  $t^1$ , their preference is not to gamble once  $t^2$  arrives and, at  $t^2$ , their preference is to gamble at  $t^2$ . Their resolution preference is the self-exclusion agreement, which was presumably entered into in foresight of a future desire to gamble and which states that, in that case, the gambler ought to be refused service. Davis' apparent solution evokes the hierarchical approaches discussed earlier. Rather than trying to determine which of the  $t^1$  and  $t^2$  decisions is the higher-order preference, however, both are cast as first-order preferences. These, then, are both subordinate to the second-order, resolution preference. The resolution preference demands that, in the case of conflict, the  $t^1$  preference ought to prevail. On the one hand, Davis' approach avoids the impracticability of hierarchical approaches. Those enlisted to enforce the Ulysses pact need not identify which preference belongs to a higher order. Rather, the decisive resolution preference is explicitly set out by the self-exclusion agreement. On the other hand, it does not avoid the problem of infinite regress. A resolution preference is no more unimpeachable than a second-order preference in a hierarchical approach and an agent may therefore develop conflicting resolution preferences. Imagine a  $t^2$  gambler who is told that their resolution preference demands a refusal of service because their  $t^1$  preference to self-exclude must prevail. There is nothing to stop the gambler from simply saying that they now have a new resolution preference in favour of  $t^2$ . These conflicting resolution preferences would now necessitate recourse to an even higher-order resolution preference about conflicting resolution preferences, which could itself be confronted by a conflicting preference, and so on. Davis says that, in the case of an agent with conflicting resolution preferences, "it is impossible to respect his autonomy—but that is his problem, not ours."<sup>101</sup> This flippant dismissal, however, raises the question of why facilitating a problem gambler avoiding their vice is 'our problem' at all and why we have bothered to come even this far.

Another potential solution is advanced by Levy, who writes specifically in the context of addiction.<sup>102</sup> He dismisses synchronic approaches—Frankfurt's in particular—and adopts a diachronic view of autonomy as the agent's *extended will* across time.<sup>103</sup> According to Levy, humans typically unify themselves as they mature. He equates this with altering one's discount curve for future goods until it approaches the ideal, so that one begins to value future goods appropriately in relation to immediate interests and maintain a unified will accordingly.<sup>104</sup> Addicts, he says, are a partial exception to this; they are more susceptible to rapidly changing preferences (the very concept of self-exclusion exemplifies

---

100. *Ibid* at 268-69.

101. *Ibid* at 275.

102. See Levy, *supra* note 68.

103. *Ibid* at 437-40. In this regard, Levy explicitly draws (*ibid* at 428 and 436) on Christman's account, *supra* note 90.

104. See Levy, *supra* note 68 at 440.

this), they are unable to apply their extended will over time and are, supposedly, more ‘fragmented’ than non-addicts. He says,

[A]ddicts experience preference reversals which are sharper and less controllable than those to which non-addicts are subject. An unwilling addict is unwilling because, though she chooses to consume, her preference is temporary, and does not reflect her will.<sup>105</sup>

Levy separates one’s (extended) will from temporary preferences.<sup>106</sup> Insofar, then, as he bases autonomy on historical personal development, Levy’s theory is susceptible to the same charges of impracticability I earlier levied against Christman’s. A gambler who enters into a self-exclusion agreement undoubtedly will have a history of gambling, one which may even outweigh in duration their period of desiring abstinence. How, then, are an operator and its staff to determine which of the  $t^1$  or  $t^2$  preferences is temporary and which reflects the agent’s will over time? If the will is simply identified with abstinence because it is preferable, this is plainly a substantive claim. Levy identifies this challenge and, drawing on Bratman,<sup>107</sup> proposes to “identify the agent herself with the part-self which is the product of her own strong-arm tactics.”<sup>108</sup> By this, he means that the individual’s ‘true self’ is the one that is forward-planning, and that this true self’s desire to abstain from their vice ought to be regarded as their extended will. This is nonetheless unsatisfying as a procedural theory of autonomy. Identifying the true self with forward planning and distinguishing this true self’s extended will from temporary preferences would probably justify self-exclusion. However, Bratman, from whose theory Levy admits to drawing heavily, holds that the true self’s acceptance of a desire is vindicated when the individual has a self-governing policy which views that desire as providing a justifying reason according to practical reasoning. From here, we can easily move to Finnis’ assertion that having a coherent plan of life forms an essential aspect of practical reasoning and therefore a good life.<sup>109</sup> My intention is not to disagree with the importance Levy, Bratman, and Finnis place on planning, coherence, and self-governance. However, these are all plainly substantive assertions. In the case of a self-excluded gambler, the assumption that long-term planning is necessarily represented by the addict’s self-denial rather than, say, a lifelong commitment to libertinism is a substantive condition. This substantive condition gives preference to one value, self-denial, over another, indulgence, based on reasons which do not emanate from the agent and therefore cannot be considered a procedural approach.<sup>110</sup>

---

105. *Ibid.*

106. Here, Levy essentially though inexplicitly identifies *akrasia* and its particular applicability to addiction.

107. See Michael E Bratman “Reflection, Planning, and Temporally Extended Agency” (2000) 109:1 *Philosophical Review* 35.

108. Levy, *supra* note 68 at 443-44.

109. See John Finnis, *Natural Law and Natural Rights* (Oxford University Press, 2011) at 103-05.

110. Levy does admit that there are “weak substantive conditions” in his otherwise procedural theory. Levy, *supra* note 68 at 444. Whether weak substantive conditions erode the value of a procedural account or not, I believe the identification of the agent’s ‘strong-arm tactics,’

## V. Why self-exclusion is permissible and voluntary slavery is not

Thus far, I have demonstrated the failure of a procedural theory of autonomy to justify self-exclusion. A content-neutral, procedural theory demands that autonomous preferences emanate from the agent and are not based on the substantive preferences of others. As I have argued, procedural attempts to resolve the self-exclusion problem invariably fail either at the theoretical level—because they introduce substantive conditions or are beset by the problem of infinite regress—or at the practical level—because of their impracticability. I will now turn to the final point of my argument, which is to make explicit what has so far been implicit: self-exclusion can only be justified within a normative framework that considers abstinence from gambling as preferable to participation in gambling. To this end, I will compare self-exclusion to a topic which has troubled procedural theorists of autonomy for centuries, the “riddle” of voluntary slavery.<sup>111</sup>

By voluntary slavery, I simply mean a contract in which one party agrees to become the slave of another in exchange for consideration of some sort. Voluntary slavery contracts bear some resemblance to self-exclusion. The person decides at  $t^1$  to become the slave of another person or to have themselves excluded from gambling, respectively. At  $t^2$ , the person articulates a desire contrary to their  $t^1$  decision—to be manumitted or to gamble—and is kept from fulfilling this desire because of the prior decision.<sup>112</sup> As a matter of procedural autonomy, voluntary slavery contracts are presumably permissible as long as the prospective slave is not coerced or manipulated. Nonetheless, I assume that, while self-exclusion grows in popularity, voluntary slavery is widely considered abhorrent and unacceptable. How can this different treatment be justified, however? For Mill, voluntary slavery is unacceptable because of its totality:

But by selling himself for a slave, he abdicates his liberty; he foregoes any future use of it beyond that single act. He therefore defeats, in his own case, the very purpose which is the justification of allowing him to dispose of himself. He is no longer free. . . . The principle of freedom cannot require that he should be free not to be free.<sup>113</sup>

The slave, he says, foregoes any future exercise of freedom. This, of course, is distinguishable from self-exclusion, which precludes only a narrow range of activities. Is it true, however, that slavery represents such a uniquely total relinquishment of future autonomy that it is to be distinguished from all else? Block points out that voluntary slavery agreements do not, in fact, involve the total

---

instead of their propensity for indulgence, as their true self to be so normative as to exceed a weak notion of substantive autonomy.

111. Feinberg, *supra* note 41 at 71.

112. The enforcement of a voluntary slavery contract differs from that of a self-exclusion agreement because it involves the protection of the rights of another party, the slaveholder. I will not pursue this point, however. The comparison between voluntary slavery and self-exclusion here is confined to the normative permissibility of entering into these respective agreements.

113. Mill, *supra* note 43 at 94.

transfer of autonomy.<sup>114</sup> The voluntary slave transfers control over their body, giving their owner the right to enact violence on them if they seek to defy the owner's commands and thereby foregoing bodily autonomy. The slave has not, however, transferred their will. The master can force the slave to toil in a mine for eighteen hours per day, seven days per week. The master can even compel the slave to act as though they are happy with this state of affairs. The master cannot, however, compel the slave to actually be happy. The will is, as a matter of fact, inalienable.<sup>115</sup> Acknowledging that voluntary slavery does not therefore involve a uniquely total transfer of autonomy, its difference from self-exclusion appears to be merely a matter of degree.

If voluntary slavery is distinguishable from self-exclusion only by the extent to which it limits autonomy and not some *a priori* difference, how can their different treatment be justified? One approach might be to consider autonomy as an abstract principle to be maximised. This abstract principle of autonomy, seeking maximisation, will be privileged over any individual act of autonomy. Therefore, the potentially autonomous act of agreeing to a contract for voluntary slavery is void *ab initio* because it extinguishes the potential for a greater realisation of autonomy (perhaps measured in number of autonomous acts performed) in the future. This view certainly coheres with Mill's utilitarian conception of autonomy, which differs markedly from deontological views of autonomy as a right. Rather, Mill defends autonomy for its instrumental, rather than intrinsic, value. He argues that vindicating autonomy contributes to utility by diffusing the potential sources of improvement, writing, "the only unfailing and permanent source of improvement is liberty, since by it there are as many possible independent centres of improvement as there are individuals."<sup>116</sup> This he calls the "progressive principle."<sup>117</sup> The autonomy maximisation model, then, might strike

---

114. See Walter Block, "Toward a Libertarian Theory of Inalienability: A Critique of Rothbard, Barnett, Smith, Kinsella, Gordon, and Epstein" (2003) 17:2 *Journal of Libertarian Studies* 39 at 47.

115. Cf Block, *ibid* at 46-47. Block believes it is possible to rid oneself of one's will by undergoing a lobotomy. He also cites with approval Kinsella's conjecture that there may, in the future, be a 'zombie pill' available which would allow a person to abrogate their own will. See N Stephan Kinsella, "Inalienability and Punishment: A Reply to George Smith" (1999) 14:1 *Journal of Libertarian Studies* 79 at 89. These arguments are questionable. First, a lobotomy may well deprive a person of their will, but it does not transfer that will to another in its entirety. The recipient of a lobotomy may be highly suggestible, but this suggestibility is not total, nor does it mean the person's pre-lobotomy capacity for wilful decisions is now controlled by another. For example, a previously highly-skilled mathematician who has been lobotomised will probably be incapable of performing the calculations they could pre-lobotomy rather than being able to perform those same calculations automatically at the initiation of another person. Second, a 'zombie pill' is, at present, science fiction. Block assumes that such a pill is an inevitability and therefore its present unavailability is "merely a technical matter, not one pertaining to philosophy." Block, *supra* note 114 at 46. First, such a pill is not an inevitability. Second, Block's philosophy being a political one, technical matters should be considered highly relevant to philosophy if he aspires to see his political philosophy implemented in practice. As an aside, even if a zombie pill were available, transferring one's will to another should be prohibited on normative moral grounds.

116. Mill, *supra* note 43 at 65.

117. *Ibid*.

down voluntary slavery contracts on the grounds that the autonomy realised in their creation is eclipsed by the future autonomy they extinguish.

Nonetheless, this model contains troubling implications. Firstly, it would compel individuals to engage in the paralysing enterprise of having to calculate in advance whether each of their actions would lead to maximal autonomy. Faced with any choice, the individual would have to predict which option would create the most opportunities for autonomous choice, then which choice presented by each of those options would do the same, and so on *ad infinitum*. This issue extends to a meta-level, too. The time it would take to make these calculations carries significant opportunity costs. A person would have to forego opportunities for autonomous action in favour of taking time to calculate maximal utility. These lost opportunities would also have to be factored into calculations, which would increase the time dedicated to calculation, resulting in a never-ending cycle. Secondly and relatedly, the autonomy maximisation model suffers from humans' inherently limited knowledge. If I knew I would die tomorrow, it is entirely probable that entering into a voluntary slavery contract today would be the most utility-maximising decision. I might, for example, leave the proceeds of the contract to my children, ensuring their long-term financial security. It must be remembered that prohibiting a voluntary slavery contract also limits my autonomy insofar as I cannot enter into one. If I died tomorrow, limiting my autonomy to sign a voluntary slavery contract today might represent an overall detriment to the utility of the final two days of my life. The same is true of self-exclusion. A problem gambler determined to overcome their addiction might derive no utility from gambling today and significant utility from signing a self-exclusion agreement. Prohibiting self-exclusion would therefore have a net negative effect on that gambler's utility today. If the gambler dies at midnight, without having yet suffered even a momentary relapse in their addiction, signing the agreement is certainly the utility-maximising choice. Of course, most people will not die tomorrow. Perhaps the utility maximisation justification of Mill's opposition to voluntary slavery is feasible on a societal level. Some, in very unique circumstances, may lose utility to the prohibition on voluntary slavery. Society in general might, however, yield greater utility by having more people innovating through autonomous action and thereby maximising societal utility. Mill does seem to consider this societal aspect. He claims that competition between the progressive principle and custom is the "chief interest of the history of *mankind*."<sup>118</sup> Calculating utility at the societal level does not escape the challenge of limited knowledge, however. A voluntary slave's master might put them to work on a project of such utility that it outweighs the utility lost through legalised voluntary slavery. Imagine a skilled medical researcher whose talent surpasses every other in their generation. The researcher, however, chooses to pursue a career as a jazz musician, despite their middling musical talent. Falling into significant debt in their chosen pursuit, a wealthy philanthropist

---

118. *Ibid* [emphasis added].

induces the researcher-musician to sign a voluntary slavery contract and immediately puts the researcher to work developing a cure to a devastating and widespread plague. The utility of this cure would surely be greater to society than that lost through prohibiting voluntary slavery, not least in that it would preserve the future autonomy of those who would otherwise have died.

It is worth restating that procedural theories of autonomy demand content-neutrality. A procedural approach therefore cannot object to voluntary slavery as impermissible because it is, for example, repugnant to human dignity. Nor, as I have argued, can a procedural approach refute voluntary slavery based on what I have called a utility maximisation model. The only remaining avenue, then, is a consistent preference for later decisions over prior ones. This approach, which would preclude Ulysses pacts entirely, is something Hobbes considered to be an essential aspect of sovereignty and therefore, presumably, autonomous self-sovereignty.<sup>119</sup> There is some ostensible merit to this view. Notwithstanding *akrasia*, the later actor does have access to information which might have been unforeseen to their earlier self. A voluntary slave might rue their naïve earlier decision when confronted with the grim reality of slavery. Similarly, a self-excluded gambler may not have foreseen just how intense their craving would be or the embarrassment they feel upon being refused service in front of members of their community. A Hobbesian approach to self-sovereignty certainly provides grounds for prohibiting voluntary slavery. However, it also makes Ulysses contracts of any sort impossible; Ulysses, Uncle Bill, and the self-excluded gambler must all have their earlier commitment disregarded and be left to choose the Sirens' enticements, cocaine, and gambling, respectively. Alternatively, of course, the  $t^1$  decision might instead be consistently preferred. This would allow for Ulysses pacts, including self-exclusion. However, within a content-neutral, procedural framework, it would also allow for voluntary slavery contracts. Robert Nozick, one of the few theorists to take this content-neutrality to its logical extreme, allows for voluntary slavery contracts, writing,

[The question] is whether a free system will allow him to sell himself into slavery. I believe that it would. (Other writers disagree.) It also would allow him permanently to commit himself never to enter into such a transaction.<sup>120</sup>

This typifies the issue with a content-neutral approach that allows for precommitment; as much as it permits precommitment of one sort, it must also allow precommitment of the opposite sort. A person can sell themselves into slavery or commit never to do so. Likewise, a problem gambler can commit to self-exclusion but can also commit never to self-exclude, a commitment which the gambler may be induced to make in exchange for special offers or which might simply become a standard clause in gambling contracts. This brings me to the main point

119. See Thomas Hobbes, *Leviathan* (Penguin Books, 1985) at 313. Though, of course, Hobbes had no interest in self-sovereignty or autonomy but, rather, totalitarian absolute monarchy.

120. Nozick, *supra* note 95 at 331. Nozick, like most libertarians, did not use the word 'autonomy', but his approach to liberty as negative freedom is nonetheless useful here.

of this article. I contend that self-exclusion is justifiable for the same reason that voluntary slavery is not—a normative position. In order for self-exclusion to be theoretically justifiable, it cannot be based on a content-neutral, procedural theory of autonomy. Rather, it must be based on an affirmative position—commonly acknowledged from the outset by all parties—that abstinence from gambling is preferable to indulgence.

## VI. Conclusion

In this article, I have argued that a procedural approach to autonomy cannot resolve the self-exclusion problem. For theoretical reasons, such as the problem of infinite regress, or reasons of impracticability, it is impossible to justify a preference for the  $t^1$  decision to self-exclude over the conflicting  $t^2$  decision to gamble only through a content-neutral recourse to the self-excluder's autonomous will. Rather, self-exclusion can only be based on a normative position that, at least in the case of the self-excluder, abstinence from gambling is preferable to participation therein. In order to justify operators and their staff acting on the  $t^1$  decision and refusing service to self-excluders, this normative position must form the explicit basis of all self-exclusion law and policy.

The inadequacy of procedural autonomy as a theoretical basis for self-exclusion leaves open numerous potential alternatives. I have argued that self-exclusion must rest on a normative position, a contention that some may take as a cue to suggest that a substantive theory of autonomy is a suitable basis. For the reasons I have outlined earlier, substantive theories of autonomy must be approached with great suspicion, lest they inappropriately (or disingenuously) affix the term 'autonomy' to something altogether different.<sup>121</sup> Alternatively, it may be the case that self-exclusion does not have the unique value I have assumed. This is, of course, a possibility. For example, Kotter *et al.* found that self-exclusion and forced exclusion have similar rates of success in promoting abstinence and reducing gambling.<sup>122</sup> If exclusion policies are based on a consequentialist desire to prevent problem gamblers from gambling, and not a deontological regard for autonomy, self-exclusion may simply exist to allow problem gamblers to self-identify for exclusion, which they might be subjected to without their consent at a later stage anyway. Self-exclusion would be deprived of its unique value, whereby a self-excluder is prevented from gambling at  $t^2$  only because of their  $t^1$  decision. Rather, the self-excluder's consent would be irrelevant and they would be excluded because they are a problem gambler in the context of a gambling policy which seeks to funnel all problem gamblers toward exclusion. Apart from its disregard for autonomy and probable disingenuity, the

---

121. To be sure, the example I provided of Babbitt, *supra* note 63, is an extreme one. Nonetheless, these issues pervade many, if not all, substantive theories. For a critique of strong substantive theories of autonomy on epistemological grounds, see Elizabeth Sperry, "Dupes of Patriarchy: Feminist Strong Substantive Autonomy's Epistemological Weaknesses" (2013) 28:4 *Hypatia* 887.

122. See Kotter, Kräplin, & Bühringer *supra* note 4.

consequentialist view fails to explain the absence of forced exclusion as a possible alternative to self-exclusion in, for example, the UK. The lack of a mechanism for excluding problem gamblers who do not self-exclude surely refutes the view that self-exclusion exists only to more easily identify some prospective excludees.

The cause of providing a theoretical basis for self-exclusion might be more fruitfully pursued in the sphere of paternalism. Recent decades have witnessed a renewed interest in paternalism and seen novel approaches advanced. These include Sunstein and Thaler's influential 'libertarian paternalism', which seeks to capitalise on individuals' cognitive bias to 'nudge' them toward particular decisions but purportedly preserves their autonomy to resist these nudges.<sup>123</sup> Earlier, I identified two criticisms of paternalism: the objection from condescension and the objection from value-partiality. As I have argued, self-exclusion must rely on a normative judgement. If it is to exist, value-partiality is therefore an inevitability. Notwithstanding this unavoidable aspect, a novel approach to paternalism might provide a theoretical basis for self-exclusion that mitigates or avoids condescension and gives adequate weight to the self-excluder's exercise of autonomy at  $t^1$ . For now, however, I will settle for having accomplished the modest goal of demonstrating that procedural autonomy cannot be the theoretical basis for self-exclusion. Discourse around self-exclusion must therefore move beyond procedural autonomy and consider alternative theoretical bases.

---

**Acknowledgements:** I would like to thank Prof. Mary Donnelly and Dr. Fidelma White for their invaluable comments on earlier drafts of this article.

**Bernard Long** is an Assistant Lecturer and PhD candidate at University College Cork's School of Law. His ongoing doctoral research focuses on the theory and practice of gambling self-exclusion agreements. His research interests include gambling law and policy, legal theory, and political philosophy. Email: [bernard.long@ucc.ie](mailto:bernard.long@ucc.ie)

---

**123.** See Cass R Sunstein & Richard H Thaler, "Libertarian Paternalism Is Not an Oxymoron" (2003) 70:4 U Chicago L Rev 1159. Sunstein and Thaler write that "respect for autonomy is adequately accommodated by the libertarian aspect of libertarian paternalism" (*ibid* at 1167). Cf Gregory Mitchell, "Libertarian Paternalism Is an Oxymoron" (2005) 99:3 Nw UL Rev 1245. Mitchell objects to the 'manipulation' inherent in libertarian paternalism. For another novel approach to paternalism, see Colin Camerer et al, "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism'" (2003) 151:3 U Pa L Rev 1211.