

# Censorship and Two Types of Self-Censorship<sup>\*†</sup>

Philip Cook<sup>‡</sup>

Conrad Heilmann<sup>§</sup>

## Abstract

We propose and defend a distinction between two types of self-censorship: public and private. In public self-censorship, individuals restrain their expressive attitudes in response to public censors. In private self-censorship, individuals do so in the absence of public censorship. We argue for this distinction by introducing a general model which allows us to identify, describe, and compare a wide range of censorship regimes. The model explicates the interaction between censors and censees and yields the distinction between two types of self-censorship. In public self-censorship, the censee aligns her expression of attitudes according to the public censor. In private self-censorship, the roles of censor and censee are fulfilled by the same agent. The distinction has repercussions for normative analysis: principles of free speech can only be invoked in cases of public self-censorship.

## Introduction

Flemming Rose, the culture editor of the Danish newspaper *Jyllands-Posten*, defended his newspaper's publication of caricatures of Mohammed as an act of defiance against

---

<sup>\*</sup>This working paper was previously entitled 'Two Types of Self-Censorship: Public and Private' which is now the title of a substantially different paper of ours that is forthcoming in *Political Studies* in 2012. The present working paper contains a more general discussion of aspects of censorship which is reflected in the new title.

<sup>†</sup>Thanks to Richard Bradley, Franz Dietrich, Markus Grottke, Paul Kelly, Andrew Lewis, Christian List, Alice Obrecht, Fabienne Peter, Juliane Riese, Kai Spiekermann, Chris Thompson, and Andrew Williams as well as seminar participants at the Warwick Centre for Ethics, Law and Public Affairs and the LSE Political Theory Seminar for helpful discussions and comments. We are grateful to Jon Parkin and the Morrell Centre for Toleration at University of York for the initial invitation to present at the 2008 Morrell Trust conference on self-censorship. Thanks also to the Philosophy Programme and the Social and Political Theory Programme at the Research School for Social Sciences at the Australian National University in Canberra for hosting both authors as visitors in July and August 2008 where work on this paper began.

<sup>‡</sup>Department of Politics and International Relations, University of Leicester, Email: pac20@leicester.ac.uk

<sup>§</sup>Department of Philosophy, Logic and Scientific Method, London School of Economics (LSE), Email: c.heilmann@lse.ac.uk

increased self-censorship by artists and publishers on issues relating to Islam.<sup>1</sup> This increased self-censorship, Rose claimed, was a response to intimidation from Muslims in Europe. Whilst regretting the violence and offense caused by the publication of the caricatures, Rose argued that principles of free speech justified taking a stand against self-censorship caused by intimidation. Rose's defense coheres with important liberal objections to self-censorship caused by threats and intimidation. However, as culture editor of a newspaper, Rose also points out that suppression of certain material is required by morality and taste. *Jyllands-Posten* does not publish images of dead bodies, and swear words are usually edited out of copy. Rose again appeals to what many would regard as considered convictions about the appropriateness of self-censorship in matters of taste, civility, and morality. As Jytte Klausen puts it in her discussion of the Danish cartoon incident, "[s]elf-censorship may be caused by a credible fear of retaliation and bodily harm, but it may also follow out of respect for other people's religious beliefs or from a desire not to hurt people's feelings."<sup>2</sup>

Whilst the relationship between justifiable and unjustifiable self-censorship raises important normative questions, self-censorship also raises a prior, more fundamental question: does self-censorship always require a censoring agent that exists independently of the censee? Are the suppression of attitudes in response to intimidation and the suppression of attitudes in response to aesthetic values both instances of self-censorship? In this paper we argue that while they are both instances of self-censorship, they are different *types* of self-censorship. The purpose of this paper is therefore to distinguish two types of self-censorship: public and private.<sup>3</sup> Public self-censorship describes a range of individual reactions to a public censorship regime. Self-censorship thus understood means that individuals internalize some aspects of the public censor and then censor themselves. Private self-censorship is the suppression by an agent of their own attitudes where a public censor is either absent or irrelevant. Private self-censorship therefore involves an intrapersonal

---

<sup>1</sup>Flemming Rose, "Why I Published Those Cartoons", *The Washington Post*, 19th February 2006

<sup>2</sup>Jytte Klausen, *The Cartoons that Shook the World*, (London: Yale University Press, 2009), p. 16

<sup>3</sup>This issue has been mentioned in passing without a full analysis of its implications. For instance, Randal Marlin states "[o]rdinary censorship typically provides us with a duality: a censoring authority and the censored communicator. How then, where censor and censored are one and the same person, does the notion of censorship get a foothold? We must suppose a duality within that person." Randal Marlin, "The Muted Bugle: Self-Censorship and the Press" in Allan C. Hutchinson and Klaus Petersen, *Interpreting Censorship in Canada*, (Toronto: University of Toronto Press, 1999), p. 291. Whilst Marlin's insight identifies the question, he offers no discussion of this proposed duality upon which self-censorship is based. The issue is also raised by Mark Cohen: "...is censorship only performed by a third party, or can it also take the form of self-suppression?", Mark Cohen, *Censorship in Canadian Literature*, (London: McGill-Queen's University Press, 2001), p. 9. Cohen argues that self-censorship is a form of censorship because an individual can internalize public forms of censorship. Like Marlin, Cohen also fails to distinguish public from private self-censorship, and consequently how private self-censorship should be understood in the absence of an external censor.

conflict between the actual expressive attitudes held by an agent and the set of permissible expressive attitudes that they endorse.

Current debates about censorship refer to a distinction between governmental and non-governmental restrictions on speech, and in particular whether it is meaningful to speak of non-governmental censors such as churches, corporations, or even social norms.<sup>4</sup> However, whether censorship is restricted to legal-political power or extended to include social and economic power, censorship in this debate refers exclusively to a censor that exists independently from a censee. Self-censorship is therefore understood as a censee's response to this externally existing censor. Yet, as suggested by Klausen, there is an important class of cases where individuals censor themselves in the absence of an *external* censor. Our paper presents a model that identifies this class of cases, and argues that individuals can engage in a process that constitutes self-censorship in the absence of an external censor.

Censorship and self-censorship seem to fit Bernard Williams's notion of *thick evaluative concepts* which include both descriptive and evaluative dimensions (where censorship is usually evaluated negatively).<sup>5</sup> Our model allows us to separate the descriptive and the evaluative dimensions of censorship, and to identify and describe different types of censorship regimes. Whilst our paper is primarily concerned to establish this novel descriptive distinction between public and private self-censorship, our analysis also has evaluative implications. Our model reveals that the agents and processes involved in public and private self-censorship are substantively different, and can therefore be subjected to different normative principles. In particular, we suggest that principles of free speech may not apply to the case of private self-censorship, because whilst an instance of censorship, the absence of an external censor makes the censorship non-coercive.

The paper proceeds as follows. In the first section, we introduce a general model that allows us to identify, describe, and compare a wide range of censorship regimes and we give examples of its application. In the second section, we show how the model yields the distinction between public and private self-censorship. In the third section we defend the distinction against two objections that attempt to deflate self-censorship to either one of the two types and show how private self-censorship elucidates existing conceptions of censorship by Hobbes, Locke, Mill and Scanlon. The fourth section demonstrates the

---

<sup>4</sup>For an influential statement of this distinction see Fredrick Schauer, *Free Speech*, (Cambridge: Cambridge University Press, 1982), pp. 122-123. A similar argument is made by Judith André, "Censorship: Some Distinctions" in *Philosophical Issues in Journalism*, edited by Elliot D. Cohen (New York: Cambridge University Press, 1992): 77-84, p. 77, and Judith André, "Poole on Obscenity and Censorship", *Ethics*, 93 (1982): 496-500, p. 499. An alternative view is presented by Lawrence C. Soley, *Censorship Inc.*, (New York: Monthly Review Press, 2002) and Mark Cohen, *Censorship in Canadian Literature*.

<sup>5</sup>Bernard Williams, *Ethics and the Limits of Philosophy*, (Cambridge: Cambridge University Press, 1985), p. 141

usefulness of the distinction for normative analysis. We conclude that the complex phenomena of censorship and self-censorship cannot be suitably discussed without recognizing the two types of public and private self-censorship identified in this paper.

# 1 Modeling Censorship Regimes

In section 1.1, we propose that censorship regimes can be characterized as a process by which censoring agents establish, justify, and enforce a goal of fit between a conception of permissible expressive attitudes and attitudes actually expressed by censees. In section 1.2, we apply the model to the incidents surrounding publication of the caricatures of Mohammed by the *Jyllands-Posten* which reveals the distinction between public and private self-censorship.

## 1.1 The Model

Our model of censorship regimes has five elements: (1) a goal of fit between a conception of permissible expressive attitudes and attitudes actually expressed; (2) the content of the conception of permissible expressive attitudes; (3) the enforcement of the fit according to the conception of permissible expressive attitudes; (4) the censor and censee; and (5) their interaction. Note that these elements of the model do not constitute necessary and sufficient conditions for censorship to exist. Rather, we take these as the five dimensions on which particular conceptions or accounts of censorship regimes will provide specific and distinctive content. We consider each of these features in turn.

**The Goal of Fit** One feature common to censorship regimes is the goal of creating a fit between a conception of permissible expressive attitudes and attitudes actually expressed. In other words, the goal of fit answers the question “why is censorship established?” Conceptions of censorship can answer this question differently, referring to considerations such as national security, public order, or democratic freedom. This dimension of the model makes it possible to describe the fundamental aims of censorship regimes.

**The Content of the Conception** Censorship regimes include an account of the content of their conception of permissible expressive attitudes. The content of the conception of publically permissible expressive attitudes answers the question “what is censored?” Censorship regimes may justify which expressive attitudes are permitted and which are not on grounds such as prudential or epistemic benefit, or moral right and wrong.

**The Enforcement of the Fit** Another feature common to censorship regimes is the enforcement of the goal of fit according to the conception of permissible expressive attitudes. This dimension of our model answers the question “how and on what grounds is censorship enforced?” Different censorship regimes will vary according to how they seek to enforce their goal of a fit between a conception of permissible expressive attitudes and attitudes actually expressed; for instance, by using their power, their authority, or by appealing to the content of the conception of permissible expressive attitudes.

**Agents** An important feature of a censorship regime is the nature of the censor and censee. This dimension answers the question “who is the censor and the censee?” Censors enforce their goal of fit between permissible expressive attitudes and the attitudes expressed by means such as their power, their authority, or their normative viewpoint. Censees express their attitudes under these conditions.

**Interaction between Censors and Censees** In addition to the four elements introduced, we need an understanding of the censorship process and the response of the censees. This final dimension of our model answers the question “how do censor and censee interact?” Intuitively, we can describe a number of ways in which censees can respond to censorship, including various degrees of opposition, compliance, brainwashing, or self-censorship. Here, we give a more fine-grained analysis of these broad intuitive categories in two steps. Firstly we discuss the possible degrees of success in the suppression of attitudes, and secondly we discuss the different dimensions of our model to which individuals respond.

(i) *Degrees of Success of Censorship Regimes.* Censorship can be more or less successful. In the perfect case, it succeeds in “brainwashing” all censees such that it creates a perfect fit between the permitted expressive attitudes and attitudes actually expressed by agents. Yet, many censorship regimes do not aim to change the private attitudes of censees, as long as the censored attitudes are not expressed or not acted on. In order to understand better the interaction between censors and censees, it is natural to suppose that censees can have two different kinds of attitudes: those that are privately held, and in addition those that are expressed publically. Further suppose, for simplicity, that the censee’s attitudes to the content of the conception of permitted expressive attitudes can take one of three values: acceptance, opposition or indifference. Crucially, in order to analyze the degree of success in the suppression of attitudes, the values of acceptance, opposition and indifference can characterize both the privately held and the publically expressed attitudes of censees. Depending on how her private and public attitudes change

in response to the censorship, any censee can respond with “perfect alignment”, “perfect non-alignment”, and various forms of “weak alignment” between their privately held and publically expressed attitudes. The possible combinations are listed in the following table.

Public Attitude Private Attitude	Public Opposition	Public Indifference	Public Acceptance
Private Opposition	Perfect alignment (1.)	Weak “pragmatic” alignment (4.)	Perfect non-alignment (8.)
Private Indifference	Weak “idealistic” alignment (6.)	Perfect alignment (2.)	Weak “pragmatic” alignment (5.)
Private Acceptance	Perfect non-alignment (9.)	Weak “idealistic” alignment (7.)	Perfect alignment (3.)

Table 1. Degrees of alignment between a censee’s privately held and publically expressed attitudes

The following list considers each cell in the table, giving examples for a censorship regime in which the censor is public and the censored agents are private individuals:

- Perfect alignment (1. 2. and 3. in Table 1) describes coherence between private and public attitudes. In these cases, the censored agent responds to the censorship announcement such that there is no incoherence in the posterior attitudes:
  1. Full public and private opposition, for example an active dissenter.
  2. Full public and private indifference, for example an individual that is not interested at all in politics or society in a dictatorship.
  3. Full public and private acceptance, for example a loyal party member in a one-party system.
- Weak “pragmatic” alignment (4. and 5. in Table 1) describes an incoherence between private and public attitudes that results in being more in alignment with the censorship in the public attitude than in the private attitude:
  4. Public indifference with private opposition, for example a cautious individual.
  5. Public acceptance with private indifference, for example an even more cautious and anxious individual.
- Weak “idealistic” alignment (6. and 7. in Table 1) describes an incoherence between private and public attitudes that results in being less in alignment with the censor in the public attitude than in the private attitude:

6. Public opposition with private indifference, for example an individual with no strong private views, but a strong sense that the censor's attitude is publicly offensive to some other individuals.
  7. Public indifference with private acceptance, for example a party member in a one-party system who privately endorses the censor's view on the grounds of ideology yet is unsure about whether such attitude should be voiced publicly.
- Perfect non-alignment (8. and 9. in Table 1) describes a maximal incoherence of private and public attitudes:
    8. Public acceptance with private opposition, for example a very cautious individual that fears consequences for showing even public indifference towards the censor.
    9. Public opposition with private acceptance, for example someone who opposes the censoring even if agreeing with the content of the conception of permissible expressive attitudes.

This taxonomy expresses to what extent a censee's attitudes are in line with those required by the censor and how much her privately held and publically expressed attitudes differ. This allows us, for instance, to distinguish different ways in which censees oppose (1, 6, 9), comply (3, 5, 8) or remain indifferent (2, 4, 7) with censorship. As such, the description above allows for quantifying degrees of success in the suppression of censee's attitudes.

(ii) *Qualitative Dimensions of Censorship.* In addition to this description of the different degrees to which the censee's expressive attitudes can fit the set of permitted attitudes, there is a qualitative dimension to censees' responses to public censorship. The extent to which censees align their truly held and expressed attitudes with the censor can occur for very different reasons. Recall the enforcement dimension in the model of censorship regimes which describes how exactly a censor aims to enforce the required fit between permitted and actually expressed attitudes. We characterize such enforcement along the different dimensions of power (e.g. the ability to suppress expression of attitudes, such as closing plays), authority (e.g. legitimacy to suppress the expression of attitudes on grounds such as rule of law), external justification (e.g. a censor's appeal to values such as public order or human rights that are independent of their particular power and authority to censor), and content (the censor's appeal to the truth or validity of the content of the conception of permissible attitudes).

In a maximally successful censorship regime, not only would all truly held and expressed attitudes of all censees be in alignment with the censor's attitudes (i.e. a case

of perfect alignment of private and public acceptance (3. in Table 1). All censees would also hold such attitudes in positive regard on any dimension of the enforcement of these attitudes: that means a censor would be feared for his power and respected for his authority and hence also agreed with in the content of the attitude, and the censoring of the attitude. However, such a case seems highly unlikely. We can therefore ask to which of the dimensions of power, authority, external justification, and content censees respond, and then ask whether they are responding affirmatively or negatively. Do censees express attitudes in alignment mainly because they are afraid of the power; do they align because they cannot bring themselves to question the authority of a dictator; do they align because they share an ideology with the censor; or do they simply agree on the content of the censored attitude? The reverse is also possible, i.e. individuals could mainly oppose a censorship regime and hold opposite expressed attitudes because of an abuse of power, or the assumed authority, justification or content of a censored attitude spurs them into open opposition. In short, using the model of censorship introduced here, we can describe an individual's response to censorship as any combination of the nine degrees of success (Table 1) and the dimensions of enforcement (power, authority, justification, content) that we can ascribe to a censorship regime.

## 1.2 Applying the Model

The model of censorship regimes describes the goal, the content of conception of attitudes, the means of enforcement, the agents within a censorship regime, and the responses it can induce, and thus allows for a range of substantively different conceptions of censorship. In order to demonstrate the efficacy of our model, we consider the incidents surrounding the publication of the caricatures of Mohammed by the Danish newspaper *Jyllands-Posten*. We show how the various dimensions of our model apply, and how this application reveals different instances of censorship and resultant self-censorship. The interaction between censors and censees in the Danish cartoon incident is complex, as there were multiple agents interacting. Application of our model brings to light the distinction between public and private self-censorship.

We identify three different censorship regimes that are associated with the Danish cartoon incident: (1) the public censorship regime constituted by Danish law, regulating the permissible expressive attitudes by newspapers, (2) the public censorship regime constituted by those members of the Muslim community who opposed publication of the caricatures and who wished to regulate issues relating to Islam in the public sphere, (3) the self-censorship regime by the editors of *Jyllands-Posten*, regulating what kinds of content they want to exclude from publishing.



Firstly, consider the public censorship regime constituted by Danish law.

**The Goal of Fit** The *Jyllands-Posten*, like all publications, exists within a legal system which promotes a fit between this conception of permissible attitudes and the attitudes actually expressed by Danish newspapers. The fundamental aim of this goal of fit is to maintain peace and order in Denmark.

**The Content of the Conception** The content of the conception of permissible expressive attitudes is constituted by the moral, political, and social values of Danish society, such as freedom, equality, and democracy. This gives substantive content to which speech acts are impermissible.

**The Enforcement of the Fit** The goal of a fit between Denmark’s conception of permissible expressed attitudes and the attitudes actually expressed in Danish newspaper is enforced legally through section 77 of Denmark’s constitution which protects freedom of speech and the freedom of the press, and sections in the Danish penal code which forbid blasphemy (§140),<sup>6</sup> racism (§266b), and libel (§267).

**Agents** The *Jyllands-Posten* is the censee. It is subject to a public censorship regime by the Danish legal system.

**Interaction between Censors and Censees** The *Jyllands-Posten* was accused of breaking various laws governing speech acts. The Danish Attorney General judged that there were no charges to answer under blasphemy and race hate laws, and Danish civil courts dismissed complaints of defamation.<sup>7</sup> According to our model, the degree of alignment between attitudes held by the censee (the *Jyllands-Posten*), and the attitudes they actually expressed was “perfect alignment of acceptance” (3. in Table 1) because they accepted the political and legal framework of censorship that allowed publication of the cartoons. Additionally we might describe the quality of the alignment as based on an acknowledgment of the power of the Danish courts (the courts have the power to prevent publication), respect for the authority of the courts (assuming that the editors regard the constitution and courts as fair and legitimate and therefore respect the court procedures in principle), consensus on the moral values that the Danish constitution endorses (values that underlie the particular legal and political institutions of Denmark such as human rights and equality), and agreement that publication of the cartoons did not violate laws of blasphemy and defamation.

---

<sup>6</sup>In practice, the blasphemy law has become moribund and decisions about the publication of religious images are governed by individual discretion, and not criminal law.

<sup>7</sup>Klausen, *The Cartoons that Shook the World*, pp. 194-196

Thus our model identifies this as a *public censorship regime*. Given that conflict was largely absent from the relationship between the Danish law and the *Jyllands-Posten*, it does not seem plausible to describe the censee's reaction to the public censorship regime as self-censorship of any kind.

Secondly, consider the public censorship regime constituted by those members of the Muslim community who objected to the publication of the caricatures, and who wished to regulate the expression of attitudes about Islam and Mohammed in the public sphere.

**The Goal of Fit** The goal of fit between a conception of permissible expressive attitudes and attitudes actually expressed concerns what might be described as a soteriological understanding of the effect of representing Mohammed visually. In other words, those who objected to the publication of the images of Mohammed claimed it was sinful and affected a right relationship between persons and God which might prevent salvation. So the goal of fit is to promote a right relationship between persons and God that ensures salvation.

**The Content of the Conception** The content of the conception of permissible expressive attitudes concerns the permissibility of representing Mohammed visually, and an interpretation of the Koran and other religious teachings (*hadith*) that prescribe rules on such visual representations.

**The Enforcement of the Fit** The enforcement of this fit was conducted through a range of social, political, and economic actions by various Muslim groups, including attempts at international denunciation at the UN,<sup>8</sup> an economic boycott of Danish goods in certain Muslim countries,<sup>9</sup> and most controversially actual and threatened violence to employees of *Jyllands-Posten*<sup>10</sup> and other Danish citizens.<sup>11</sup>

**Agents** The censees are the *Jyllands-Posten* and any other newspaper or individual with access to publishing. The censors are those parts of the Muslim community that have threatened (and/or carried out) violence, legal action and public pressure with regards to the publishing of the cartoons.

**Interaction between Censors and Censees** In this instance, the attitudes held by the *Jyllands-Posten* were in opposition to the perceived censorship by the Muslim community, and the attitudes they actually expressed were intended to voice that

---

<sup>8</sup>Ibid., pp. 63-82

<sup>9</sup>Ibid., pp. 72-76

<sup>10</sup>Ibid., p. 58

<sup>11</sup>Ibid., p. 46

opposition (1. in Table 1). The qualitative basis for such “perfect alignment of opposition” was the *Jyllands-Posten*’s contestation of the Muslim community’s perceived power to enforce censorship, their rejection of the authority of this group of Muslims to create and enforce a conception of permissible expressive attitudes, and their rejection of a theological framework in which religious offense restricts the rights of others to free expression.<sup>12</sup> Jytte Klausen points out: “The cartoons were printed around the margin of an essay, and the headline of the essay was ‘Mohammed’s Ansigt’ which means the face of Mohammed. [...] it was a cartoon editorial, it was a provocation. [...] They deliberately wanted to break a taboo and that’s why it was printed.”<sup>13</sup>

However, the broader public reaction to the publication of the cartoons was of course characterized by violent objection. Many publishers and news organizations refrained from publishing the caricatures. Yale University Press, the publisher of Jytte Klausen’s *The Cartoons that Shook the World*, declined to publish the cartoons on the pragmatic ground of safety and security.<sup>14</sup> In this case we can see that Yale University Press are in a condition of perfect non-alignment between private opposition to the public censor prohibiting the publication and public acceptance of that attitude by not publishing (i.e. 8. in Table 1). We may assume that this perfect non-alignment is based on a pragmatic acknowledgment that public opposition threatens violence, a principled rejection of the authority of any violent objectors to stifle free speech, indifference or perhaps disagreement that there is a theological basis for objecting to the caricatures, and indifference that the content of the caricatures is offensive.<sup>15</sup> This analysis suggests that Yale University Press reluctantly responded to this public censorship regime with a type of self-censorship. As we will discuss in the next section, this non-alignment amounts to *public self-censorship* because the non-alignment is in response to an external censor, which in this case is a perceived community that threatens violence in response to the expression of an attitude it opposes.

Thirdly, consider the self-censorship regime by the editors of *Jyllands-Posten*, regulating what kinds of content they refrain from publishing.

**The Goal of Fit** The proprietors and editors of *Jyllands-Posten* establish a goal of fit between permissible expressive attitudes and the attitudes actually expressed to

---

<sup>12</sup>Ibid., pp. 20-27

<sup>13</sup>Jytte Klausen, “See No Evil”, *Index on Censorship* 38 (2009), pp. 77-78. Available online: <http://ioc.sagepub.com/cgi/content/abstract/38/4/74>.

<sup>14</sup>Ibid, p. vi

<sup>15</sup>Indifference here specifies what might in fact be a combination of attitudes such as that the caricatures are not offensive in content to most Muslims but are very offensive to some.

promote the existence and influence of the *Jyllands-Posten*. The goal of publishing a newspaper that exercises influence is itself informed by the moral, political, commercial, aesthetic, and professional goals of the publishers and editors.

**The Content of the Conception** The proprietors and editors institute a conception of those attitudes it is permissible to express in the newspaper (constrained by more general laws and norms), that define the format and content of the paper, and which they believe best serves its goal of commercial, journalistic, and political viability. These various goals may not be perfectly compatible, for example: material they consider morally unacceptable might sell more newspapers. But whilst the newspaper might have plural and conflicting goals, it nonetheless has a determinate conception of the attitudes it deems permissible to express in the *Jyllands-Posten*. For example, Jytte Klausen points out that the *Jyllands-Posten* is broadly center-right in its political support and editorial, and that it is the largest selling Danish broadsheet newspaper.<sup>16</sup>

**The Enforcement of the Fit** The newspaper enforces these rules and the content of its own conception through its employment practices, disciplinary procedures, and editorial decisions.

**Agents** The *Jyllands-Posten* is both the censor and the censee.

**Interaction between Censors and Censees** The individual journalists on the *Jyllands-Posten* may have views that conflict with the conception of attitudes permitted by the newspaper, for example they may take a different political line on a given issue, or believe the newspaper's copy editing suppresses good writing. We might describe journalists who suppress such disagreements as in a condition of holding private opposition to the policies and practices of the newspaper but expressing public indifference or acceptance. Our model describes such journalists as in a condition of either weak "pragmatic" alignment (such as 4. and 5. in Table 1) or perfect non-alignment (8. and 9.) of their own attitudes. The individuals might be responding merely to the power of editors to hire and fire the journalists or to the authority of the editors.

Our model also helps explicate the censorious dimensions of the notorious decision of the *Jyllands-Posten* to decline to print unsolicited caricatures of Jesus on the grounds that they would cause offense to readers.<sup>17</sup> In the case of the suppression of the cartoons of

---

<sup>16</sup>Klausen, *The Cartoons that Shook the World*, p. 11

<sup>17</sup>"Danish Paper Rejected Jesus Cartoon", *The Guardian*, 6th February, 2006

Jesus, the *Jyllands-Posten* was acting as a censor over the attitudes of the cartoonist who submitted the unsolicited images of Jesus. Such suppression occurs within a domain of discretion created by Danish law and norms (expression of such images is permitted and not required, therefore discretion is created). The decision of the editors to suppress such publication is therefore made in the absence of a public censorship regime. This amounts to *private self-censorship* as it belongs to a class of cases of censorship where suppression of the expression of attitudes takes place in the absence of a public censor.

## 2 Public and Private Self-Censorship

Applying the model, we have identified two different types of self-censorship, characterized by the presence or absence of an external public censor. We now consider the difference between these two types of self-censorship in more detail.

### 2.1 Public Self-Censorship

There are various ways in which an individual may respond to public censorship. Consider again the censorship model outlined in section 1. In order to distinguish public self-censorship from other kinds of responses agents may have to public censorship regimes, one can combine the nine dimensions of attitude alignment with the descriptions of how censorship is actually enforced. In a second step, one can then identify the combinations which may most plausibly be described as a response of public self-censorship when facing a public censor.

Most conceptions of public self-censorship will center on those cases in which there is some conflict between private and public attitudes, such as weak “pragmatic” alignment (4. and 5. in Table 1), weak “idealistic” alignment (6. and 7.) or perfect non-alignment (8. and 9.). This captures the intuition that public self-censorship involves a deliberate effort on the part of an individual to resolve a conflict of attitudes between herself and the censor by taking into account both idealistic and/or pragmatic concerns. It seems particularly plausible to reserve the notion of public self-censorship for the pragmatic responses of an individual to public censorship regimes, such as the responses of weak “pragmatic” alignment (4. and 5.) and instances of “pragmatic” perfect non-alignment (8.) of attitudes as outlined by our model above. To restrict public self-censorship to a conflict between private and public attitudes seems especially plausible conceptually if individuals respond to the power of the censor rather than to its authority or normative justification. The more pragmatic dimensions of the enforcement of censorship capture the idea that public self-censorship is the result of a deliberative effort in which an individual

endorses both their private attitudes and censor’s point of view. By distinguishing different ways to enforce the goal of fit of permissible expressive attitudes and expressed attitudes on the basis of the power, authority, or external justification of the censor, our model can describe the process by which an individual responds to different elements of a censorship regime, and how this may affect a censee. For example, whilst “pragmatic” alignment seems a plausible response to censorship enforced by power, more principled alignments seem likely when the censee is responding to the authority or external justification of the censor. Censees are much more likely to also privately adopt the claims of the censor on these grounds than that of power alone.

In the light of these distinctions, consider again the case of Yale University Press and their decision not to publish the cartoons in Klausen’s book. We can see that our model helps specify both the nature of the censor (those members of the Muslim community who violently objected to publication of the caricatures), and the effect of the censor on the alignment of the attitudes of the censee (public self-censorship). We describe this as a form of public self-censorship, as the censee has decided to suppress their original attitude of wishing to publish the cartoons in light of the public censorship regime. Consequently, our model allows for specification of a type of self-censorship in which the censee responds to a censor that is external to the censee on several dimensions of our model.

## 2.2 Private Self-Censorship: by Proxy and by Self-Constraint

Censorship regimes may be instigated by a range of different kinds of censors, acting for different kinds of reasons. Here, we consider one particularly interesting variance of censoring agents, namely cases where the roles of censor and censee are fulfilled by the same agent.

Private self-censorship captures the idea of personal restraint resulting in the suppression of privately held attitudes; that is, as an individual’s self-imposed suppression of the expression of their own attitudes. In this context, private self-censorship is a requirement to adhere to certain standards, derived from, for instance, a set of norms, moral considerations, or decency. Private self-censorship thus understood means that individuals take some idea or consideration to overrule other attitudes they might have. Here, we consider two levels for establishing such private self-censorship: either *by proxy*, or *by self-constraint*. Firstly, private self-censorship can be established by an individual’s internalization of some external set of values, such as the norms of an association. We describe this as private self-censorship by proxy. Secondly, private self-censorship can be established by an individual’s suppression of attitudes in the absence of an explicitly external or public influence, such as when an individual adopts a personal set of values that

constrain the expression of their attitudes. We describe this as private self-censorship by self-constraint.

Consider the example of Anne, Brian and Chris who are three work colleagues. Anne and Brian both dislike Chris but say nothing. Chris moves to another job. Brian starts to insult Chris in private to Anne. Anne shares all of Brian's attitudes to Chris, and knows that their private conversation could never get back to Chris or anyone else, and therefore there could be no punishment for expressing the same attitudes as Brian. However, Anne does not wish to express these attitudes.

The example of the colleagues can represent an instance of private self-censorship where Anne is suppressing her attitudes by considering a conception of constraints on personal behavior. If the sources of those conceptions are external to Anne, for instance she wishes to abide by a code of conduct or a social norm, then we can describe this as private self-censorship by proxy. If the motivations are much more private, for instance that she likes to lead a life without talking behind people's backs out of a private sense of proper personal conduct, then we can describe this as private self-censorship by self-constraint. The difference can be spelt out in more detail using our model of censorship, beginning with private self-censorship by proxy:

**Self-Censorship by Proxy and the Goal of Fit** A goal of fit between a conception of permitted expressive attitudes and attitudes actually expressed is constituted by an external set of values or public institution. This goal is adopted and endorsed privately by an agent. An agent adopts the goal of achieving a fit between the attitudes they actually publically express, and the conception of permitted expressive attitudes constituted by the external set of values.

**Self-Censorship by Proxy and the Content of the Conception** An agent adopts and endorses privately the content of the conception of permitted expressive attitudes of a public agent. The agent therefore endorses the content of the set of external values, whether that is a set of social norms or political credos.

**Self-Censorship by Proxy and the Enforcement of the Fit** An agent enforces the fit between the conception of publically permitted expressive attitudes and the attitudes they actually express, independently of whether the external set of values is itself enforced by some other agent.

**Self-Censorship by Proxy and the Censoring Agent** Whilst adopting the goal of fit, endorsing the content of the conception of publically permitted expressive attitudes, and enforcing the goal of fit, the agent is regarding their own private attitudes from the point of view of the external set of values. The agent is therefore acting

as the proxy of some public institution or set of independently existing values by establishing a private censorship regime which can be described by our model.

**Self-Censorship by Proxy and Interaction between Censor and Censee** The conflict between the expressive attitudes regulated by the censoring agent and those of the censored agent can be described by the nine degrees of alignment (Table 1) and various qualitative aspects.

Note that as in the application of our model to public censorship, the different elements just mentioned do not constitute necessary and sufficient conditions of private self-censorship by proxy – they are rather the dimensions on which we might want to describe those censorship regimes. We next set out the second kind of private censorship regime: private self-censorship by self-constraint.

**Self-Censorship by Self-Constraint and the Goal of Fit** An individual adopts a goal of fit between a conception of attitudes they allow themselves to express and the attitudes they actually express absent reference to any external censorship regime. The conception of attitudes they allow themselves to express may derive from a set of moral commitments, personal political or theological principles, or pragmatic commitments they endorse. The goal of fit element of self-censorship by self-constraint points to the fact that an individual sets themselves the goal of making their expressed attitudes conformable to a set of constraints on those attitudes.

**Self-Censorship by Self-Constraint and the Content of the Conception** An agent adopts and endorses the content of a set of constraints on expressed attitudes, independently of any external censorship regime. Self-censorship by self-constraint requires that the individual finds the basis for such a conception of permissible expressive attitudes in herself.

**Self-Censorship by Self-Constraint and the Enforcement of the Fit** An individual enforces the goal of fit according to the authority, external justification or even power of the conception of permissible expressive attitudes. Self-censorship by self-constraint means that the individual will establish the resources of enforcement herself. For instance, an individual might use her willpower, meditation, or prayer to find the strength to enforce the goal of fit. Or the individual will convince herself of the authority or the normative superiority of the content of the conception.

**Self-Censorship by Self-Constraint and the Agent of the Private Censor** An individual suppresses their own public expression of attitudes from the point of view of



a conception of permissible expressive attitudes that does not reduce to an external agent of a public censor. Such a point of view could be a set of moral constraints such as deontological or consequentialist impartiality. Or it could be a set of constraints that is grounded in tradition or a strong conception of shared interests such as a republican notion of “common good”. These points of view are based on a set of moral, prudential, or political principles, and are not constituted by any particular agent or institution.

### **Self-Censorship by Self-Constraint and Interaction between Censor and Censee**

The suppression of expressive attitudes of the private censee by the private censor can be analyzed according to the nine degrees of success (Table 1) and the qualitative dimensions of enforcement. This enables characterization of the internal dynamics of attitudes within the individual.

The two types of self-censorship identified by our model allow for analysis of a wide range of possible conceptions of self-censorship, including those in which it is primarily a private notion and those in which self-censorship is viewed in relation to public censorship regimes.

## **3 Defending Two Types of Self-Censorship**

In this section, we address some possible objections to our distinction between public and private self-censorship. In section 3.1, we deal with the objection that public self-censorship will always be reducible to an instance of private self-censorship. In section 3.2, we consider the converse objection, namely that whenever we talk about private self-censorship, it just amounts to discussing a reaction to public censorship. In section 3.3 we show that private self-censorship elucidates existing conceptions of public censorship by Locke, Mill and Scanlon.

### **3.1 Is Self-Censorship Always Private?**

We now consider a “deflationary” objection against public self-censorship. According to this view, all discussion of self-censorship as a response to public censorship should refer only to self-censorship as a private censorship regime: while the self-censorship might indeed be induced by a public censorship regime, it is best understood without reference to it.

Such a view however, renders inapplicable a particularly useful part of our model, namely the description of censorship outcomes according to the alignment of attitudes

between the censee and the censor. As mentioned in Section 1, the different degrees of success a censor can have in manipulating the censor are important to distinguish particular conceptions of self-censorship. For instance, one might want to defend the view that self-censorship is akin to successful brainwashing. Then, one will have to endorse the view that self-censorship is associated with perfect alignment of attitudes between censor and censee. To take a different example, one might want to defend self-censorship as a particularly skillful opposition tactic, where individuals maintain their private attitudes, but in public always act as if they are in full support of the censor’s attitudes. Then, one will have to endorse the view that only specific instances of “pragmatic” alignment (namely 4. and 5. in Table 1) and perfect “pragmatic” non-alignment (namely 8. in Table 1) of attitudes between censor and censee constitute outcomes that are self-censorship.

It is of course possible to maintain the view that such different degrees of success are really just the outcomes of a much more complicated self-censorship process which is best described in the context of an entirely private censorship regime. The latter view is consistent with our model: one could view self-censorship in response to public censorship as always instigating a private censorship regime which then generates different levels of alignment with the public censor. But upholding that the censorship regime is entirely private in all cases makes it impossible to discuss self-censorship in relation to public censorship and the outcomes it yields. This would preclude any normative or conceptual discussion of the effect public censorship can have on individuals in a society. While it might be opportune to discuss self-censorship in this way in some cases, we find that restricting all discussion to such cases would limit the understanding of self-censorship considerably.

### **3.2 How Public is Private? – On the Existence of Self-Censorship by Self-Constraint**

Here we consider the converse objection: that any instance of self-censorship can be adequately described as a response to a public censorship regime. We defend the existence of the second type by returning to the discussion of the two different kinds of private censorship regime identified in section 2.2. Whilst both private self-censorship by proxy and private self-censorship by self-constraint fulfill the general model of censorship that we outlined in section one, they differ from the conceptions of public censorship through the absence of an external censoring agent.

Maintaining that there is only the first type of self-censorship amounts to claiming that there is in fact only one kind of private self-censorship: self-censorship by proxy, such that any instance of self-censorship must reduce to an agent acting on behalf of or from

the standpoint of some external agent. The latter, in turn, could then be reduced to self-censorship of the first type, namely self-censorship as a response to a public censorship regime. We accept that what might at first appear to be self-censorship by self-constraint may be an outcome of a certain process of self-censorship by proxy. However, the ability of an agent to adopt a standpoint from which they can assess and regulate the expression of their own attitudes does not depend on that standpoint always reducing to an external *agent*.

In order to exemplify our notion of self-censorship by self-constraint in section 2.2, we gave an example in which Anne suppresses the expression of attitudes on the grounds of a private sense of how to properly conduct her life. In the proxy case, the individual was taking the point of view of the external agent, but in the self-constraint case, the individual is taking a point of view external to their private subjective attitudes, but not external to their own agency. According to the objection, Anne’s self-suppression can be traced to some public conception, such as social conventions regarding public decency. Even if this is the case, it does not follow that Anne’s suppression is ultimately traceable to an external *agent*: for an individual to reason or act from the standpoint of some set of claims, is not necessarily to reason or act from the point of view of some external agent. It is indeed possible to consider the content of an agent’s propositions without considering their external agency. Thus, the objection seems to rest on conflating agency with standpoint. Hence reducing all cases of self-censorship to public self-censorship would limit our understanding of the processes by which agents can regulate their behavior.

### 3.3 Self-Censorship by Self-Constraint and Public Censorship by Proxy

The basis on which individuals may self-censor in the absence of an external agent provides us with an important development in the understanding of public censorship regimes. That is to say, there are some kinds of public censorship that are conceivable in terms of a conception of self-censorship by self-constraint. In those cases, the public censor is acting as a proxy for a set of self-constraints on individuals.

To elucidate this kind of public censorship by proxy, consider the difference in the public censorship regimes described by Hobbes, Locke, Mill, and Scanlon. In Hobbes’s case, the entire content of the conception of publically permissible attitudes is given by the sovereign, in virtue of their sovereignty.<sup>18</sup> This is a clear instance of public censor-

---

<sup>18</sup>Hobbes’s discussion of censorship is well known for granting the Sovereign power and authority to restrict the publication of books and the performance of public speeches. Hobbes’s argument for such restrictions is based on the interest of the members of the commonwealth in maintaining peace and preventing war: “... it is annexed to the Sovereignty, to be Judge of what Opinions and Doctrines are averse, and what conducing to Peace; and consequently, on what occasions, how farre, and what, men are to be trusted withall, in speaking to Multitudes of people; and who shall examine the Doctrines of all bookes

ship where the censor is an agent external to the censees, in the form of the sovereign. However, not all conceptions of public censorship ascribe the content of a conception of publically permissible expressive attitudes to an external agent. For example, in Locke's case, the content of the conception of publically permissible expressive attitudes is given by positive law which relies on humans to comprehend and apply natural law according to reason.<sup>19</sup> Similarly, in Mill's case, the content the conception of publically permissible expressive attitudes is determined by the principle of aggregated utility. Again, particular political, philosophical, and judicial agents will actually calculate and judge the content of this principle, but this is merely incidental and instrumental to the nature of the existence of the principle, which is independent of any particular person's conception of it.<sup>20</sup> Farther, Scanlon has defended a "Millian" narrow conception of public censorship on non-consequentialist grounds.<sup>21</sup> Scanlon's argument is derived from the audience's interest in being able to make decisions freely and independently and as such does not reduce to the will or interests of an external agent as censor. Hence, in certain kinds of public censorship regimes, such as those proposed by Locke, Mill and Scanlon, the content of a conception of permitted expressive attitudes can be established independently of a public agent.

Whilst the content may be determined absent a public censor, is enforcement of the conception of permitted expressive attitudes also possible absent a public censor? Our model suggests: yes. Normative reasons may have authority based on the shared claims of individuals. Deontological impartialist conceptions of morality, such as Scanlon's, argue that principles of right and wrong are authoritative if they cannot be reasonably rejected by anyone motivated to agree on principles to regulate behavior that are acceptable to all similarly motivated. Consequentialist impartialist conceptions of morality, such as Mill's, argue that principles of right and wrong are authoritative if they treat each person's utility equally. In other words, the constraints on individual behavior are enforced on the basis of a normative authority that is constituted by shared authority relations between persons, and not in reference to an interest or value that is external to the shared authority relations between persons. Therefore, the enforcement of a conception of publically permissible

---

before they be published." Thomas Hobbes, *Leviathan*, (Cambridge: Cambridge University Press, 1996), Bk. 2, Ch. 18, p. 124

<sup>19</sup>John Locke, *A Letter Concerning Toleration*, edited by Susan Mendus and John Horton, (London: Routledge, 1991)

<sup>20</sup>J. S. Mill, *On Liberty*, edited by John Gray, (Oxford: Oxford University Press, 2008), ch. 2. A similar argument is presented by Alexander Meiklejohn, *Political Freedom*, (New York: Harper, 1960) who defends the first amendment to the Constitution of the United States as an absolute prohibition on censorship justified because of the necessity for free speech in a society in which citizens are responsible for democratic decisions.

<sup>21</sup>See T. M. Scanlon, "A Theory of Freedom of Expression", *Philosophy and Public Affairs*, 2 (1972), pp. 204-226, and T. M. Scanlon, "Freedom of Expression and Categories of Expression", *University of Pittsburgh Law Review*, 40 (1978-9), pp. 519-550

expressive attitudes can proceed in the absence of an external public authority.

In other words, self-censorship by self-constraint allows us to see that conceptions of public censorship that are based on second-personal normative authority are only conceptions of public censorship derivatively. They are primarily conceptions of public censorship by proxy. Therefore, it is important to refute the reductivist view of private self-censorship as being solely a private censorship regime by proxy because this not only retains a conceptually distinct conception of self-censorship by self-constraint, but this refutation also allows us to distinguish conceptions of public censorship that rely on non-agent based content and enforcement.

## 4 Two Types of Self-Censorship and Normative Analysis

In this section, we show how our distinction between public and private self-censorship can contribute to normative discussions of censorship and free speech. We firstly revisit the events Mohammed caricatures case, and then consider conceptions of free speech.

### 4.1 Private and Public Self-Censorship in the Danish Cartoon Incident

Our model provides a unique explanation of the different kinds of agents involved in the Danish cartoon incident, and the variety of possible interactions between them. This unique account of the agents and their interactions reveals the different normative principles that apply, and to whom they apply. Consider two very different responses to the publication of the cartoons by the *Jyllands-Posten*.

The first response, voiced by the Danish government and many sympathetic newspapers, was that the newspaper was justified in publishing the cartoons because principles of free speech permit causing offense, and that it is impermissible to censor speech through threats of violence. Our model reveals that this response recognizes the public censorship regime of Danish law as legitimate, and the public censorship regime instituted by parts of the Muslim community as illegitimate. As shown in section 1.1, the *Jyllands-Posten* was aligned with the Danish law, yet was in conflict with parts of the Muslim community. Hence, normative evaluation of the *Jyllands-Posten*'s actions as a response to censorship requires differentiating the two public censorship regimes with which the newspaper was interacting. Such differentiation is required because supporters of the *Jyllands-Posten* would regard the Danish censorship regime as legitimate, and the censorship regime instituted by that part of the Muslim community who objected as illegitimate. Therefore, our model specifies that the normative issue at stake is a conflict between two opposing censorship regimes, and their competing claims to legitimacy.

The second response regards the *Jyllands-Posten*'s actions as morally wrong as they caused offense, hurting the feelings of members of a minority religious group who suffer many injustices, caused outrage and festered division between cultures. This view can be explained by our model as a negative normative assessment of their private self-censorship regime. For instance, one could argue that their private censorship regime was illegitimate because it discriminated against the feelings of the Muslim community unfairly by publishing the caricatures of Mohammed but censoring the caricatures of Jesus. Note how our model enables fine-grained distinctions within an analysis of a private self-censorship regime: one could either criticize the *Jyllands-Posten* in its role as a censor or as censee. As censor, the *Jyllands-Posten* institutes a censorship regime by specifying a goal of fit, a conception of permissible attitudes, and enforcing this fit over the expression of its own attitudes. If the *Jyllands-Posten*'s censorship regime endorsed a conception of permissible attitudes that was discriminatory towards Muslims, it is morally culpable as private self-censor. As censee, the *Jyllands-Posten* responds to this privately instituted censorship regime through the publication of its newspaper. If the *Jyllands-Posten*'s discriminatory approach to printing the cartoons of Mohammed and cartoons of Jesus was a result of failing to apply a non-discriminatory private censorship regime, then it is morally culpable as a censee.

The distinction between private and public self-censorship is particularly useful in analyzing the more recent controversies in the Danish cartoon case. In February 2010, a settlement between the newspaper *Politiken* which had reprinted the cartoons and eight organizations representing 94,923 of the Prophet Mohammed's descendants was reached, with the latter dropping legal action against the newspaper. As part of the settlement, the newspaper *Politiken* issued a statement in which they apologized for the offense caused by their reprinting the Cartoons in question.<sup>22</sup> This statement was met by harsh criticism from Danish media, politicians and also the *Jyllands-Posten*: "*Politiken* has betrayed the battle for freedom of speech. They've given up and bowed to threats. That is, of course, disgraceful," its Editor-in-Chief Mikkelsen said in response. The three censorship regimes discussed earlier help to analyze these developments. Firstly, *Politiken*'s statement frames the problem in terms of apologizing for failed private self-censorship. Secondly, the statement arose in the context of a lawsuit that can be described as being part of the public censorship regime in which parts of the Muslim community participate. Thirdly, the *Jyllands-Posten*'s reaction draws on free speech, framing the problem in terms of the perceived legitimacy of the Danish public censor and the perceived illegitimacy of the public censorship regime by parts of the Muslim community. This suggests that the

---

<sup>22</sup> "Politiken Settles Mohammed Cartoon Issue", *Politiken*, 26th February 2010. Available online: <http://politiken.dk/newsinenglish/article911102.ece>

public debate about the Cartoon incident can be seen as disagreements about which of the censorship regimes involved are legitimate and which are illegitimate.

## 4.2 Free Speech

As mentioned in the introduction, our model enables us to separate the descriptive and evaluative dimensions of censorship. The preceding discussion has focused on how our model helps clarify varieties of censorship and self-censorship regime, and in particular how our model reveals a distinction between public and private self-censorship. We now consider the implications of the descriptive role of our model for normative discussions. Prominently, Fredrick Schauer argues that public censorship is characterized by the discretionary power of governments to control speech.<sup>23</sup> Schauer's conception of censorship leads to him to deny the existence of private and public self-censorship: because individuals are in a condition of necessary choice about their reception of speech acts, "[t]his additional dimension of private suppression as an act of speech, or at least a corollary to it, sharply distinguishes private from government censorship, and makes the notion of private self-censorship almost self-contradictory."<sup>24</sup> Our model disputes this elimination of private and public self-censorship and provides support for those who accuse Schauer of holding a too restrictive view of censorship. Indeed, Schauer has been criticized for excluding non-governmental agents such as corporations, churches, norms and cultures that can also exercise non-necessary restrictions on speech acts, and therefore act as censors.<sup>25</sup> Our model showed that censorship processes can produce many kinds of formal and informal agents who can act as censors such as constitutions, governments, public opinion, and social norms. For example, our model was able to explain the response of those Muslims who objected to the caricatures as a public censorship regime, because they constituted a goal of fit, a conception of permissible expressive attitudes, and sought to enforce this conception on the *Jyllands-Posten*. Our model therefore clearly allows for both governmental and non-governmental public censorship.

The ability of our model to describe a broader set of non-governmental public censorship regimes has important normative implications for the application of principles of free speech. It describes a wider set of censorship regimes; should principles of free speech similarly widely? Schauer argues that principles of free speech only apply to public censorship regimes because censorship does not exist outside relations between governments and in-

---

<sup>23</sup>Fredrick Schauer, *Free Speech*, pp. 122-123

<sup>24</sup>Ibid. However, Schauer goes on to grant the possibility of (public) self-censorship when a publisher faces risks of liability, i.e. when there is the possibility that there will be legal consequences to publishing a statement (pp. 170-171).

<sup>25</sup>For example, see Soley, *Censorship Inc.*, and Mark Cohen, *Censorship in Canadian Literature*.

dividuals. Our model has revealed that censorship relationships do exist beyond relations between government and individuals (namely our broader conception of public censorship, and our identification of public and private self-censorship). However, whilst our model does identify a wider set of censorship regimes, this does not imply that principles of free speech apply to them all.

Schauer maintains that principles of free speech are inapplicable to self-censorship because principles of free speech do not apply to the necessary choices of individuals. In contrast, our model can extend the application of principles of free speech to *public self-censorship*. Thus, if we have a legitimate public censorship regime, then agents who align the expression of their attitudes with it are behaving justifiably. Conversely, if a public censorship regime is normatively problematic, for instance one that intimidates and threatens the censees, then opposition to it is preferable to public self-censorship. In addition, the degree of justifiability can be explained by the nine degrees of success of censorship regimes specified by our model: the more successful at enforcement an illegitimate public censorship regime, the less praiseworthy the individual's alignment.

When considering *private self-censorship* we suggest that principles of free speech are inapplicable, but on different grounds from Schauer. Principles of free speech relate to coercive relationships between agents: namely a public censor and a censee. Our model suggests that analysis of the legitimacy of censorship regimes should focus on the actions of the censor. We agree with Schauer that it is entirely appropriate to consider such public censorship regimes under principles of free speech. The wrong that is involved in violations of free speech must include, *inter alia*, the wrong of coercing an agent to suppress their speech against their will or interests. Can such a wrong (namely a wrong of coercion) be performed by an individual on the expression of their own attitudes? We argue: no. Private self-censorship does not involve an interpersonal relationship between agents as censor and censee; private self-censorship consists in an intrapersonal relationship within an agent where they act as both censor and censee. It is impossible for an individual to coerce themselves by restricting their own freedom of speech through private self-censorship, because coercion must involve the restriction of the action of one agent by another. Coercion therefore requires an interpersonal relationship. Private self-censorship, as our model has shown, is an intrapersonal relationship. It could be objected that the model indeed introduces two agents within one individual, the censoring agent and the censored agent, and that their relationship can be understood as interpersonal. Even though we find this implausible, accepting this objection does not pose a problem: both the censoring and the censored agent are part of a private individual. Therefore, the public dimension is absent from private self-censorship and the coercion that may be involved between the two agents in the individual cannot be analysed by principles of free speech. Our model consequently



allows for both a broader description of censorship and self-censorship regimes, and consequently a clearer specification of the agents within such regimes. An important normative implication is that principles of free-speech only apply within a narrow set of interpersonal censorship relationships, and therefore that principles of free speech are inapplicable to intrapersonal private self-censorship.

## Conclusions

We have argued for a distinction between two types of self-censorship: public and private. Applying the general model of censorship regimes which we have introduced in this paper, we find that the two types imply different notions of censor, censee and their interaction: in public self-censorship, the censor is a public agent, such as a government or public authority, and the censees are private individuals or corporations. By contrast, in private self-censorship, censee and censor are the same agent, such that the censorship process involves the suppression of attitudes within one individual. We have distinguished two conceptions of private self-censorship, namely by proxy and by self-constraint, to highlight the degrees to which motives for private self-censorship can be external to the agent. We find that those distinctions are relevant for the normative analysis of complex cases of censorship and self-censorship, and the interaction between censors and censees. We argued that principles of free speech are not applicable for cases of private self-censorship. Thus, the distinction between public and private self-censorship put forward enables clarification of the complexities of censorship and self-censorship, and the normative and descriptive problems associated with them.