

#MeToo, Social Norms, and Sanctions*

KATHARINA BERNDT RASMUSSEN

Philosophy, Institute for Futures Studies, Stockholm

and

NICOLAS OLSSON YAOUZIS

Philosophy, University of Stockholm

I. INTRODUCTION: #metoo

In October 2017, following the Harvey Weinstein scandal, US actress Alyssa Milano called upon victims of sexual harassment to unite on social media. Since then, millions of women worldwide have used the hashtag #metoo, a slogan originally coined by activist Tarana Burke in 2006, to push for an end to sexual harassment and violence. Among them were collectives of women from various professions, who set out to highlight workplace harassment, revealing anonymized stories ranging from apparently ‘innocent’ sexist comments to full-blown sexual assaults. The general aim of the #metoo movement was thus both to reveal the scope of the problem and to advocate change.

The initial #metoo movement, in many countries, was followed by a further chain of events: public or powerful figures were ‘outed’ as sexual harassers, by traditional media or on public social media accounts. Consequently, many of them were publicly shamed, fired, or otherwise socially sanctioned. The French hashtag #BalanceTonPorc (‘Expose Your Pig’) indeed explicitly encouraged women to name and shame their harassers.

In France—as in many other countries across the globe—these events moreover sparked legal efforts to introduce or improve anti-harassment legislation, as well as court cases in which some harassers were dealt prison sentences. These legal decisions were greeted with equal parts disbelief and joy by activists, indicating how unexpected such outcomes are.¹

*This is the accepted version of the following article: “#MeToo, Social Norms, and Sanctions”, Katharina Berndt Rasmussen & Nicolas Olsson Yaouzis, *The Journal of Political Philosophy*: Early View, pp.1–23, Copyright © 2019, John Wiley & Sons Ltd, which has been published in final form at <https://doi.org/10.1111/jopp.12207>.

The authors wish to thank Erik Angner, Gustaf Arrhenius, Paul Bowman, Tim Campbell, Åsa Carlson, Erik Carlson, Anders Herlitz, Karim Jebari, Magnus Jedenheim-Edling, Jens Johansson, Björn Lundgren, Ed Page, Jessica Pepp, Folke Tersman, and other participants of the research seminar in practical philosophy, Philosophy Department, Uppsala University, and of the PPE seminar, Institute for Futures Studies, Stockholm, as well as Bob Goodin and two anonymous referees of this journal, for helpful comments on earlier drafts of this article. Berndt Rasmussen gratefully acknowledges financial support by the Swedish Research Council and by the Marianne and Marcus Wallenberg Foundation. Olsson Yaouzis gratefully acknowledges financial support by FORTE, the Swedish Research Council for Health, Working Life and Welfare.

¹For a recent assessment of some of the global effects of #metoo so far, see e.g.

<<https://foreignpolicy.com/2019/03/07/metoo-global-impact-international-womens-day>>.

These developments—the increase in naming and shaming and the perceived changes in court sentencing practices—have also led to a backlash against the entire movement. Critics mainly worry that innocent people may be found guilty, be it by public ‘witch hunts’ or by what are perceived as politicized judges or juries. In the wake of #metoo, people have become more likely to see false accusations as a greater problem than unreported assaults.² There has also been a rise of defamation suits, or threats thereof, in response to the naming and shaming.

In sum, while the consciousness-raising efforts of the initial #metoo movement—which refrained from outing perpetrator—were generally accepted as necessary for change, the subsequent public naming and shaming of suspected perpetrators has been widely renounced. Activists, on the other hand, have pointed out that previous consciousness-raising efforts (from the 1970s feminist movement to recent campaigns such as #EverydaySexism, #YesAllWomen, and #NotOkay) have had little effect, as the problem persists.

The #metoo movement, the subsequent media drives, and current backlash have rattled the system of social norms concerning harassment and retribution. This article aims to illuminate these recent events with the help of a game-theoretical framework of social norms. Our analysis shows that, contrary to prevailing public opinion, there are reasons to accept not only the consciousness-raising activities, but also sanctioning—such as naming and shaming. Our guiding questions are: why is sexual harassment still a problem when we have been aware of it at least since the 1970s? What are the mechanisms behind sexual harassment? And what are the practical implications from such an improved understanding?

In Section II, we review the philosophical discussion on sexual harassment and situate our article against this background. In Section III, we introduce our game-theoretical model and put it to work. We model the effect of social norms on different types of norm-followers in a dynamic simulation of workplace sexual harassment, and relate this to both the recent #metoo movement and the philosophical work concerning the topic. Section IV discusses practical implications of our analyses and provides some suggestions about what can be done to eliminate sexual harassment.

II. PHILOSOPHICAL BACKGROUND

Although the phenomenon has presumably been around throughout the history of human kind, the term ‘sexual harassment’ was first coined in the 1970s by separatist women’s groups in the USA.³ In 1979, Catherine MacKinnon analysed the concept within a framework of philosophy of law, drawing on and renegotiating legal practice. MacKinnon’s work has spawned a large field within philosophy of law dealing with sexual harassment: its definition,

²For some backlash statistics, see e.g. < <https://www.economist.com/united-states/2018/10/20/measuring-the-metoo-backlash>>.

³See Crouch 2001; Farley 1978; MacKinnon 1979; Siegel 2003.

its moral wrongness, and legal remedies.⁴ The legal philosophical debate has since been revolving around two main questions: what is sexual harassment, that is, which actions fall under the concept? And how can these actions be handled by the law?

The concept sparked interest even outside of the philosophy of law. Philosophers of language and epistemologists, specifically those prone to an ethical turn within their fields, have used it to illustrate their theories. Pioneering this development, Miranda Fricker analyses the long-time lack of a concept for the distinctive and shared social experiences of sexual objectification and exploitation as a form of epistemic injustice.⁵ Others have contested this analysis, or further refined the sense of injustice at play in sexual harassment.⁶ One of the most recent additions to this field employs the #metoo movement itself, as a case study illustrating the mechanisms of epistemic injustice.⁷ The guiding question here is what kind of wrong or injustice does sexual harassment constitute?

Philosophers focusing on applied philosophy have discussed the normative implications of the phenomenon. Jennifer Saul urges philosophers to refocus on the question of what we should do about sexual harassment, instead of getting stuck in debates about its nature and wrongness. Moreover, she cautions against limiting this discussion to legal and institutional remedies, since this risks creating the impression that those—contested and oftentimes inefficient—remedies are all we can and should rely on. Her suggestion is to focus more on informal ways of addressing the problem, and particularly on bystanders' responsibilities to intervene and report.⁸ The question is thus what should we all do about sexual harassment?

Our article is a piece of applied philosophy. Ultimately, we want to take on Saul's challenge to make progress on the question of what to do about sexual harassment. In the course of our arguments, we will consider the roles that informal rules and procedures, as well as bystanders, can play. We will not address the conceptual, legal, and moral issues directly. Instead, we will be working with a stipulated, clear-cut case of workplace sexual harassment. We will consider the social norms at play in such a case, and their consequences.

Before proceeding, let us say something about why game theory is well equipped for improving our understanding of what can be done about sexual harassment. After all, just as it has been objected that game theory cannot properly account for conventions and norms, it can be objected that it cannot properly account for phenomena such as sexual harassment. Margaret Gilbert, for example, has argued that David Lewis's game-theoretic account of conventions⁹ should be dismissed because it fails to capture 'the way people

⁴See Anderson 2006, for a helpful taxonomy of different theories of the nature and wrongness of sexual harassment within philosophy of law. Cf. Crouch 2001, ch. 5.

⁵Fricker 2007.

⁶See e.g. Anderson 2017; Dotson 2012; Mason 2011.

⁷Jackson 2018.

⁸Saul 2014. For a discussion of situational factors that make different types of bystander interventions more or less likely, see Bowes-Sperry and O'Leary-Kelly 2005.

⁹Lewis 1969.

think and talk about convention in the course of their everyday life'. Gilbert's main point is that an adequate theory of conventions should capture its phenomenology: it should help us 'understand the terms in which people understand their own lives—the terms in which they act'.¹⁰

We concede that if we want to *understand* how people think of, for example, conventions, social norms, and sexual harassment, then it is important to provide a theory that captures its phenomenology. However, our main interest is not to understand how ordinary people think and talk about such phenomena. Our goal is to examine what, if anything, can be done to get rid of sexual harassment. For this purpose, a rational reconstruction of what a social norm is may suffice.¹¹ In so far as the reconstruction allows us to make good predictions about what will happen when we make social interventions, it does not matter whether the reconstruction provides understanding of how people think of their social norms.

In other words, we do not claim that our analysis provides a descriptively accurate picture of the psychological and sociological processes in play.¹² Rather, we propose an idealized, but predictively useful game-theoretic model of sexual harassment in the workplace. We focus on the workplace, since it constitutes a limited environment with clearly identifiable 'players' with intuitive sets of strategies and payoff structures. Moreover, as described, the #metoo movement is strongly (though not exclusively) connected to women's professional environment.

III. A MODEL OF SEXUAL HARASSMENT

The purpose of the model is to point out that some seemingly innocuous assumptions have surprising implications with respect to the persistence of sexual harassment. Like all models, ours involves simplifications.¹³ We assume that the workplace consists of two types of players, *men* and *women*, who play a sequential game consisting of two stages. The first stage consists of the men waking up in the morning, deliberating about whether to *adopt a professional jargon* or *adopt a macho jargon*. The men make these decisions simultaneously and independently of each other. Those who adopt a professional jargon will, during the work day, take care not to make, for example, abusive jokes and comments about women.

¹⁰Gilbert 2008, p. 6. See also Anderson (2000), who argues that game theory's 'as if' is ill equipped to account for social norm following.

¹¹See Bicchieri 2005, p. 3, for a justification of providing a rational reconstruction of social norms.

¹²We will not make any assumptions about the psychological, sociological, or other background of the perpetrators of harassment. For a research overview on characteristics of men who sexually harass, see e.g. Pina et al. 2009. For a discussion of various social meanings and functions of sexual harassment, see e.g. Siegel 2003. For an overview of the personal costs to victims of sexual harassment, see e.g. O'Hare and O'Donohue 1998, p. 562.

¹³See Gibbard and Varian's (1978) classic justification of models and simplifications.

The men who adopt the macho jargon, on the other hand, may make sexist comments, jokes, and so on, that can lead to other forms of sexual harassment.¹⁴

If nobody adopts the macho jargon, the game ends after the first stage. However, if someone adopts the macho jargon during the first stage, the game proceeds to a second stage where the women, at the end of the work day, deliberate about whether to *remain silent*, *speak up*, or *sanction*. Just like the men, the women are assumed to make their decisions simultaneously and independent of each other. Women who remain silent do not make their dissatisfaction with the use of the macho jargon public. Women who speak up make their dissatisfaction public (for example, engage in consciousness raising about sexual harassment), but do *not* direct this towards the co-worker who used the macho jargon. Finally, women who sanction *both* make public their dissatisfaction and take action against the co-worker who adopted the macho jargon (for example, report him, or name and shame him on social media).¹⁵

The women's strategies will be introduced one by one as parts of different scenarios. Scenario 1 assumes that women have no choice but to remain silent. Subsequently, in Scenario 2, we introduce the alternative of *speaking up*. The set of the women's alternatives will be completed when we add the strategy *sanction* in Scenario 5.

A. SCENARIO 1: NO NORM AGAINST SEXUAL HARASSMENT

The first scenario models the historical situation before the term 'sexual harassment' was coined. Men are assumed to deliberate between adopting a macho and a professional jargon, while women are assumed to automatically remain silent.

The reason why women do not have the choice of speaking up may be that they lack the hermeneutical resources to make sense of their experiences, and thus face hermeneutical injustice.¹⁶ Alternatively, women's attempts to speak up may be silenced in the dominant discourse, since *men* ignorantly or wilfully lack this concept.¹⁷ Or, as far as a successful instance of speaking up involves being treated as a competent user of the concept of sexual harassment, women may also be described as remaining silent if their use of the concept is seen as misguided, such that they are subjected to conceptual competence injustice.¹⁸ This may be the case when their complaints are dismissed with the response 'he was just joking' or 'you should take it as a compliment'.

¹⁴We assume that men choose between adopting different jargons since this seems less far-fetched than assuming they decide whether or not to commit sexual harassment. Also, research shows that among the strongest organizational risk factors for sexual harassment are (1) an unprofessional environment (e.g. 'where there is swearing and drinking on the job') and (2) sexist attitudes ('women are viewed as sex objects and as inferior to men'); O'Hare and O'Donohue 1998, pp. 567–568. The strategy *adopt macho jargon* can thus be understood as active participation in causing or maintaining these risk factors. Finally, macho jargon may itself constitute sexual harassment. International surveys show that among the most commonly reported forms of sexual harassment are sexual jokes and remarks, sexual looks and gestures; see e.g. Pina et al. 2009, p. 128.

¹⁵See Section IV for different ways of sanctioning.

¹⁶Fricker 2007.

¹⁷Mason 2011.

¹⁸Anderson 2017.

In Scenario 1, it is assumed that all men have identical preference orderings. Consider, for example, Caesar, one of the male employees, who wakes up in the morning deliberating what jargon to adopt. Caesar wants to develop his career, but does not feel completely comfortable using a macho jargon. This may be because, unlike his parents' generation, Caesar's generation does not consider traditional macho traits like strength, aggressiveness, and womanizing to be virtues. Thus, adopting the macho jargon gives rise to a cognitive dissonance that Caesar would rather be without. If the choice was between an outcome where everyone adopts the macho jargon and an outcome where everyone adopts the professional jargon, Caesar would prefer the latter. However, Caesar also knows that because his superiors belong to an older generation, they value traditional macho virtues. Therefore, he believes that if he adopts a professional jargon while everyone else adopts the macho jargon, his superiors will view him as weak or aloof and probably not promote him when the time comes. Caesar also realizes that if everyone else adopts a professional jargon, then he can catch his superiors' attention by adopting a macho jargon and thereby increase the probability of being promoted.

If all men have the same preference ordering as Caesar, then the men's payoffs are represented by the matrix in Table 1. The matrix shows that Caesar is trapped in a *prisoner's dilemma* and that, no matter what everyone else does, Caesar is better off adopting the macho jargon.

		Other men	
		Macho	Professional
Caesar	Macho	2, 2	4, 1
	Professional	1, 4	3, 3

Table 1. The harasser's dilemma

Note: For our purposes, it is sufficient that these numerical values represent an ordinal ranking of the options for Caesar and the Others, respectively, ranging from 4 (best) to 1 (worst).

It is easy to verify that the strategy profile where everyone adopts a macho jargon is the unique *Nash equilibrium*.¹⁹ That is, when everyone adopts the macho jargon, everyone plays their best response to everyone else's chosen strategy. Furthermore, because adopting the macho jargon strictly dominates adopting the professional jargon, a rational man will adopt the macho jargon no matter what he believes the other players will do. So, in Scenario 1, we can expect everyone to adopt the macho jargon.

¹⁹Formally, let s_i be i 's chosen strategy and $s_{-i} = \{s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_m\}$ be the set of everyone else's chosen strategies. Then the strategy profile $\mathbf{s} = \langle s_i^*, s_{-i}^* \rangle$ is a Nash equilibrium if, and only if, for all players, i , $u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*)$.

B. SCENARIO 2: PROFESSIONALISM AS A SOCIAL NORM

It is common to think that *social norms* help groups avoid suboptimal outcomes.²⁰

Therefore, suppose that the discovery of the concept of ‘sexual harassment’ also results in the discovery of the social norm, ‘at work, adopt a professional jargon’. Scenario 2 examines whether the addition of this norm affects how the men behave in the workplace. We apply Bicchieri’s game-theoretical account of social norms,²¹ and suggest that the group of men can be separated into four categories with respect to their attitudes regarding the rule ‘at work, adopt a professional jargon’.

The first group are the *Saints*, who consider adopting a professional jargon in the workplace to be a *moral norm*, in the sense that they are *unconditional* followers of the rule.²² An example of what most of us consider a moral norm in this sense is the rule, ‘do not murder’. In this context, Saints are supposed to be unconditional adopters of the professional jargon. Thus, Saints refrain from adopting a macho jargon, even if they believe that everyone else would adopt it. The second group, *Sexists*, are the exact opposite of the Saints: they consider the adoption of the macho jargon to be a moral norm. That is, they adopt it no matter what they believe others will do.

However, unconditional norm followers such as the Saints or Sexists will not be the focus of our discussion. We will, rather, focus on the third and fourth group of men who consider the rule a *social norm*. What distinguishes those who view the rule as a social norm from those who view it (or its opposite) as a moral norm is that the former are *conditional rule followers*. So, the third group are the *Normies*, who see the adoption of a professional norm at work as a *legitimate* social norm. This means that Normies prefer to follow a rule if, and only if,

[*Empirical expectations*:] they believe that enough others follow the rule, and

[*Normative expectations*:] they believe that enough others expect them to follow the rule.²³

In condition 2, the term ‘expect’ denotes a normative belief. Think of a mother telling her child, ‘I *expect* you to behave at granny’s’. The mother may not have the empirical belief that the child will actually behave at granny’s, but rather wants to communicate that the child ought to behave. Similarly, according to Bicchieri, ‘enough others expect them to follow the rule’ should be interpreted as ‘enough others believe that they ought to follow the rule’.²⁴

²⁰That is not to say that all social norms are beneficial. See e.g. Bicchieri and Fukui 1999, on the persistence of unpopular social norms.

²¹Bicchieri 2005, 2016.

²²Note that this is a stipulative definition of ‘moral norm’ used in the context of norms.

²³Bicchieri 2005, p. 11.

²⁴Ibid., p. 14. Her term ‘normative expectation’ is thus best understood as elliptical for ‘empirical expectations about others’ normative beliefs’.

A good example of what many of us consider to be a legitimate social norm is ‘when asked, pay your taxes’. We are willing to follow this rule as long as we believe enough others follow it, and as long as we believe enough others believe that we ought to do so. However, were we to discover that few others follow the rule, we would stop following it, since we do not want to be ‘suckers’. Furthermore, if we were to discover that few others expect us to follow it, we would also stop following it, because doing so involves a significant cost.

The fourth group is the *Jerks*, who are also conditional norm followers, but consider the social norm *illegitimate*. This means that Jerks want to follow a rule if, and only if,

[*Empirical expectations*:] they believe that enough others follow the rule, and

[*Normative expectations with sanctions*:] they believe that enough others want them to follow the rule *and* are willing to sanction their behaviour.

The difference between Normies and Jerks is that Normies, but not Jerks, may be willing to follow the social norm even if they come to believe that nobody is willing to punish their transgression. For example, suppose that a Jerk and a Normie both believe that everyone else pays their taxes and believes that everyone else expects them to pay their taxes. This is enough to get the Normie to pay her taxes, but, unless the Jerk also believes that most others are willing to punish him for not paying, he will be unwilling to pay his taxes.

Note that if the population is believed to be composed mostly of Sexists, then neither Normies nor Jerks are willing to follow the norm, since they do not want to end up as ‘suckers’. Conversely, if the population is believed to be composed of mostly Saints, then the empirical-expectation condition will be satisfied for both Normies and Jerks. This means that they do not have to worry about ending up as suckers. However, it is important to remember that increasing the number of Saints in the population does not guarantee that Normies and Jerks begin to follow the social norm. Normies only conform if they also believe that enough others expect them to conform, and Jerks only conform if they believe that enough others are willing to sanction their behaviour. From now on, we ignore Saints and Sexists and focus on social norm followers: that is, Normies and Jerks.

Note that individual Normies and Jerks may differ with respect to the conditions on which they prefer to conform to the rule. Some want to conform only if they believe that almost everyone else conforms, whereas for some it is enough that only some others conform. Furthermore, some Normies demand that almost everyone expects them to conform, whereas for others it is enough that just a few expect them to conform. Finally, some Jerks demand harsh sanctions, whereas for others light sanctions suffice.

Initially, we assume that all men are Normies. That is, all men consider the rule ‘at work, use professional jargon’ to be a legitimate social norm. We also assume that all Normies are identical with respect to the conditions on which they prefer to conform to the rule. That is, all men prefer to conform to the rule if they believe that some proportion (d_{emp}) of all men conform, and if they believe that at least some proportion (d_{norm}) of all women want them to conform. To simplify, it is also assumed that all men have identical

beliefs about the proportion of other men who conform to the social norm (b_{emp}), and about the proportion of women who want them to conform (b_{norm}).

These simplifications allow us to say that all Normies, i , prefer to follow the social norm R ='at work, use professional jargon' on the condition that,

Empirical expectations: i believes that enough other men conform to R , i.e., $b_{emp} \geq d_{emp}$, and

Normative expectations: i believes that enough women expect i to conform to R , i.e., $b_{norm} \geq d_{norm}$.

When a social norm is followed, mixed-motive games, such as prisoner's dilemmas, are transformed into coordination games.²⁵ To see how this works, consider again Caesar, who is made aware of the existence of the social norm to use a professional jargon at work. If Caesar considers the norm legitimate and believes 1) that other men will adopt a professional jargon, and 2) that women want him to use a professional jargon, then his preference ordering changes to the one depicted in Table 2.

		Other men	
		Macho	Professional
Caesar	Macho	Second best	Third best
	Professional	Worst	Best

Table 2. The harasser's preference ranking with a professionalism norm

Assume that everyone else, in line with Caesar's empirical expectations, adopts a professional jargon. In this case, the cost of violating the norm that is followed by everyone else exceeds the benefit of gaining a career boost. Therefore, Caesar prefers to adopt a professional jargon when everyone else adopts a professional jargon. If, contrary to his expectations, everyone adopts the macho jargon, then Caesar prefers to adopt the macho jargon. His preference for following the norm is, after all, conditional on everyone else following it. Were he to discover that he is the only one who follows the norm, he would feel like a 'sucker' and prefer to switch to the macho jargon. This distinguishes Caesar from a Saint, who is an unconditional follower of the rule 'at work, adopt a professional jargon'. If Caesar had considered it a moral norm, he would have preferred to adopt a professional jargon even when others adopt a macho jargon.

So, when Normies believe that enough others conform to the norm and when they believe that enough women expect them to conform, they view their choice between adopting a macho and a professional jargon as a coordination game. The coordination game is illustrated in Table 3. Since adopting the professional jargon is the best response to

²⁵Ibid., p. 26.

everyone else adopting the professional jargon, the outcome where all Normies adopt a professional jargon is an equilibrium of this game.

		Other men	
		Macho	Professional
Caesar	Macho	3, 3	2, 1
	Professional	1, 2	4, 4

Table 3. The harasser's coordination game with a professionalism norm

C. SCENARIO 3: PROFESSIONALISM AS A SOCIAL NORM—A DYNAMIC MODEL

The third scenario introduces the women's second strategy, *speaking up*, and transforms the game into a dynamic one. The game is now assumed to be played for an indefinite number of rounds. In the first stage of each round, each man decides, based on his expectations about the other men and women, what jargon to adopt. In the second stage of the round, if any men have adopted the macho norm, each woman decides whether to speak up or to remain silent. At the end of each round, the players update their expectations based on their observations of the other players' behaviour. Once the expectations have been updated, the next round begins.

There are many possible rules for updating one's expectations. A man could, for example, change his expectation about the proportion of men who will adopt a professional jargon next round to reflect his most recent observation; or he could change it to reflect the mean of the most recent 10 observations; or update it with the help of Bayes's theorem, and so on. Here it will be assumed that the agents use a simple and general rule for updating their expectations. They increase their expectation of the proportion of others who will follow the norm next round, if their observation in this round exceeded their expectations. They lower their expectations for the next round if the observation fell short of their expectations in this round. Finally, they retain their expectation if their observation matched their expectations. It is also assumed that, if increased, the new expectation will not exceed their observation, and if decreased, the new expectation will not fall below their observation.

If men never make mistakes, and if men's normative expectations about the proportion of women who believe men ought to adopt a professional jargon do not change, then once a population of Normies has adopted the professional jargon, they will stick to it indefinitely. After all, if they do not make mistakes, then their observation during the initial round will match their expectation that everyone will adopt the professional jargon. Therefore, nobody changes his expectation about the proportion of men who will adopt the professional jargon in the second round. Because no mistakes are made, the same reasoning shows that men will not change their beliefs after round 2, and therefore they will also adopt the professional jargon in round 3. And so on, ad infinitum.

Now, suppose that men sometimes ‘tremble’, in the sense that they intend to adopt the professional jargon, but instead adopt the macho jargon by mistake.²⁶ This may happen because they make an honest mistake or because they fall back into old habits.²⁷ Suppose that at some round, t , men expect all other men to adopt the professional jargon. Suppose also that a very small proportion of men adopt the macho jargon by mistake. At the end of the round, men will update their empirical expectations about what men will do next round. Suppose that the new expectation will not dissuade men from following the social norm next round. That is, suppose that the new expectation is above the threshold value, d_{emp} . Thus, all men prefer to adopt the professional jargon next round and, if no mistakes are made, everyone adopts the professional jargon at $t + 1$. So, at the end of the next round, men will adjust their expectations upwards again and, as long as there are no mistakes in the coming rounds, the expectation will recover and steadily approach 1.

The dynamic is illustrated by Figure 1. The dashed line represents men’s empirical expectations and the solid line represents the proportion of men who adopt the professional jargon. At $t = 2$, a proportion of men adopt the macho jargon by mistake. This causes men to lower their expectations about how many men will adopt the professional jargon at $t = 3$. However, since no more mistakes are made, the empirical expectations quickly recover—shown by the dashed line approaching 1.²⁸

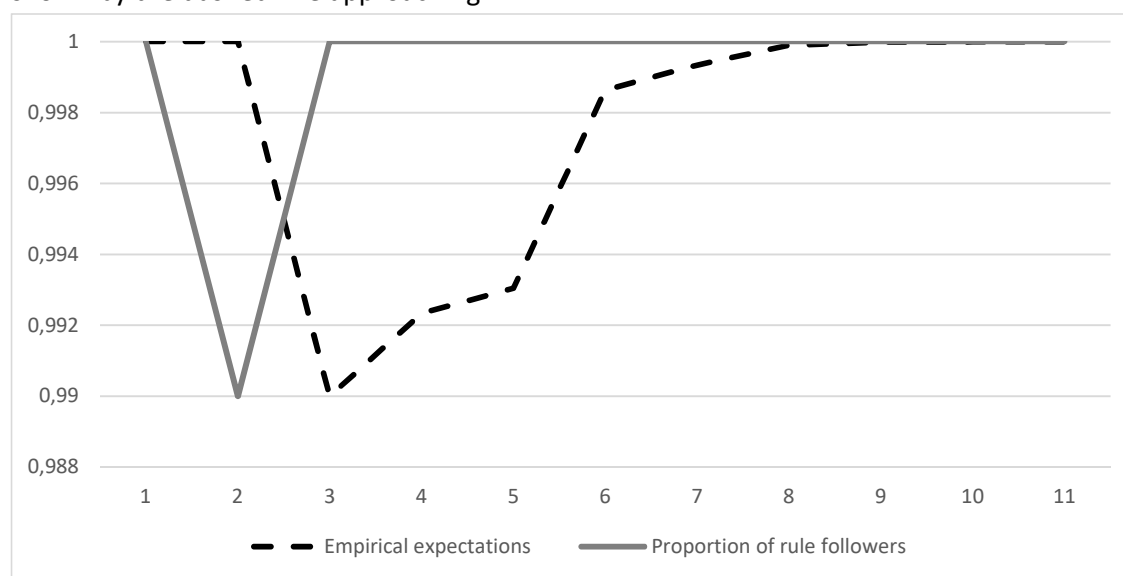


Figure 1. Men’s recovering empirical expectations after a ‘tremble’

²⁶The concept of ‘trembling hands’ was introduced by Selten 1975.

²⁷To simplify, we assume that under normal circumstances nobody makes mistakes. However, it is easy to extend the model so that, under normal circumstances, a certain proportion of the men—it may be different individuals at each period—adopt the macho jargon by mistake. As long as the level of ‘continuing trembles’ (as a reviewer neatly called them) is relatively low, the following analysis will change only nominally. See n. 33 below.

²⁸In the simulation illustrated in Figure 1, it is assumed that the probability for trembling is 0.025 and that, in the case of a tremble, 0.01 of the male population adopts the macho jargon. Men are assumed to use the belief-revision algorithm described above.

Next, assume that men form their normative expectations concerning how many women believe that men ought to adopt a professional jargon based on the proportion of women who speak up *in case* there are men who adopt the macho jargon. If nobody adopts the macho jargon, then men do not expect women to speak up. However, if someone adopts the macho jargon, then men expect women who oppose the use of the macho jargon to speak up. So again, we suppose men use the simple rule for updating their beliefs. If more women than expected speak up, men increase their normative expectation. If fewer women than expected speak up, they lower their expectation. If the proportion of women who speak up match their expectation, they retain it for the next round. We also assume that, if increased, the new expectation will not exceed the observed proportion and, if lowered, the new expectation will not fall below the observed proportion.

Suppose that, in the initial round, men expect all women to believe that men ought to adopt the professional jargon. First, note that because men's normative expectations can only change if some men actually adopt the macho jargon, they will never change if nobody does so. Second, if women speak up every time men adopt the macho jargon, men will retain their normative expectation that all women believe that men ought to adopt the professional jargon. Therefore, if women speak up following a tremble, this dynamic will resemble the one shown in Figure 1.

However, now consider *Ada*, one of the women, who has the choice between speaking up and remaining silent. Ada knows that it is important for women to speak up in order to remind men that they ought to adopt the professional jargon. Therefore, she prefers the outcome where everyone speaks up to the outcome where everyone remains silent. However, Ada is also painfully aware of what can happen to women if they speak up. The costs associated with speaking up may consist of worsened career opportunities due to being considered 'difficult' or being branded a 'feminist killjoy'. Therefore, the worst alternative for Ada is to speak up when everyone else remains silent. Furthermore, Ada realizes that if everyone else speaks up, men will be reminded that they ought to adopt the professional jargon. So, if everyone else speaks up, Ada will be able to reap the benefits while avoiding the risks of speaking up. Ada's situation is effectively identical to Caesar's in the no-norm situation: whatever the other women do, Ada is better off remaining silent. The situation is represented by Table 4.

		Other women	
		Remain silent	Speak up
Ada	Remain silent	2, 2	4, 1
	Speak up	1, 4	3, 3

Table 4. The dilemma of the harassed

Now suppose, as above, that a very small proportion of men tremble, and adopt the macho jargon by mistake. As we saw, as long as there are no more mistakes, men's *empirical*

expectations will soon recover and approach 1. The problem is that men's *normative expectations* will not recover, according to the following dynamic.

The dilemma of the harassed entails that no women will speak up when men tremble. Since the observed proportion of women who speak up falls short of men's expectations, men will lower their expectation about the proportion of women who believe that men ought to adopt the professional jargon. Suppose that the new expectation remains above the threshold d_{norm} . This means that during the next round, all men will (try to) adopt the professional norm. However, men will only upwardly adjust their normative expectation if, *in case* some men adopt the macho jargon, more women speak up than they expect. But, because of the dilemma of the harassed, no women will ever speak up, so the men's normative expectations will never recover.

Consequently, the next time some men adopt the macho jargon by mistake, men will be given further reasons to decrease their expectations about the proportion of women who believe men ought to adopt the professional jargon. Assuming that men continue to occasionally tremble, their expectation about women will eventually drop below the threshold d_{norm} . Once below the threshold, men, being conditional norm followers, no longer want to conform to the social norm. Caught once again in the harasser's dilemma, everyone adopts the macho jargon. In other words, when the game is repeated and mistakes are made, the equilibrium where everyone adopts the professional jargon becomes unstable.

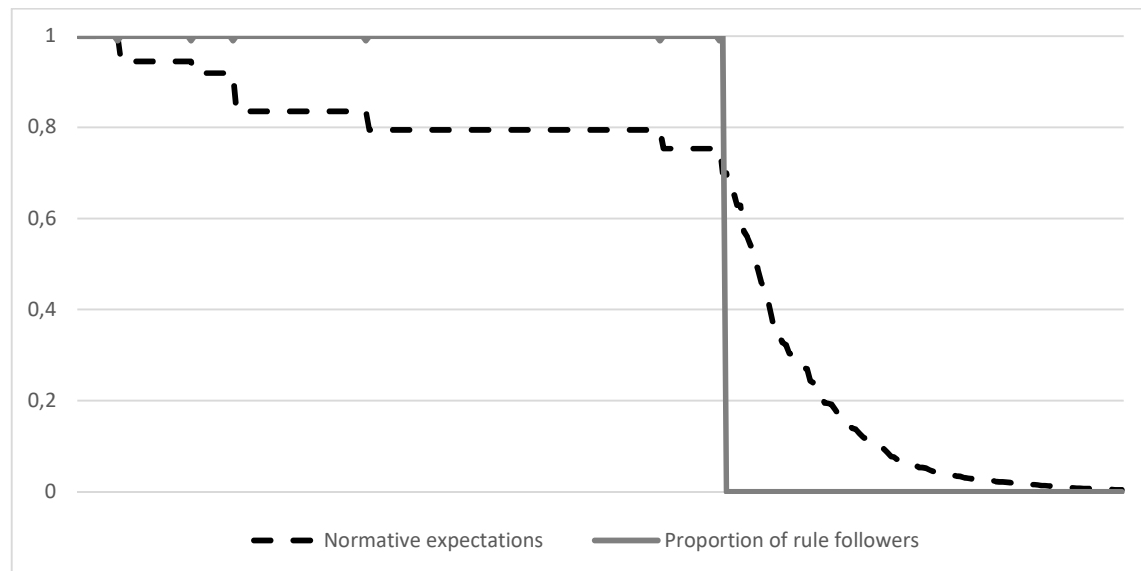


Figure 2. Men's normative expectations with a professionalism norm (no speaking-up norm)

This phenomenon is illustrated by Figure 2, which shows how the effect on men's normative expectations aggregates when women fail to respond to trembles. The solid line represents the number of men who adopt the professional jargon and the dashed line represents men's expectation about the proportion of women who believe men ought to adopt the professional jargon. Each notch in the solid line represents a round where a small

proportion of men tremble and adopt the macho norm by mistake. Since women do not speak up, men's normative expectations will keep on dropping until they fall below the threshold point (here assumed to be 0.7). Once below the threshold, all men adopt the macho jargon. Men's expectation about the proportion of women who believe men ought to adopt the professional jargon, represented by the dashed line, will then approach zero.²⁹

To recapitulate, the third scenario models the situation after the introduction of a social norm of using a professional jargon. It is assumed that public scandals and awareness campaigns cause men to change their expectations both about the proportion of other men who adopt a professional jargon *and* about the proportion of women who believe men ought to adopt a professional jargon. However, as women find themselves in a dilemma of the harassed, they will not speak up. Therefore, the model predicts that men's normative expectations will decrease over time as more and more mistakes are made. Being conditional norm followers, they will fall back into the macho jargon.

D. SCENARIO 4: SOCIAL NORMS TO BREAK THE SILENCE

The obvious solution is to introduce a norm of speaking up. It could be argued that one of the consequences of #metoo was to introduce a norm that women should publicly oppose the macho jargon. Here, we simplify, and assume that the social norm involves the behaviour rule 'whenever you observe the use of macho jargon, speak up'.

Suppose that women view this as a legitimate social norm, and that they prefer to follow the norm on the condition that 1) they expect enough other women follow this rule, and 2) they expect that enough other women believe that women ought to follow this rule.³⁰ Now, suppose that in addition to introducing a norm of speaking up, #metoo convinced women that enough other women follow the norm of speaking up, and that enough other women believe that women ought to follow it as well.³¹

Just as the professionalism norm transformed the harasser's dilemma into a coordination game, so the norm of speaking up transforms the dilemma of the harassed into a coordination game. If Ada views the norm as legitimate, believes that enough other women follow the norm and that other women believe Ada ought to follow the norm, then Ada prefers to follow the norm. This means that, although she is unwilling to be the only

²⁹In the simulation illustrated in Figure 2, it is assumed that the probability for trembling is 0.025, that in the case of a tremble, 0.01 of the male population adopts the macho jargon, and that the threshold point for men's willingness to adopt the professional jargon is 0.7. Women are modelled as not speaking up. Men are assumed to use the belief-revision algorithm described above.

³⁰Note that it is possible to extend the model to allow for women being game theoretical equivalents of Saints, Jerks, and Sexists as well.

³¹'Enough other women' obviously does not imply 'all other women'—women may be conditional norm followers for rather low threshold values, meaning that a proportion of 'just a few' might be enough to make them willing to speak up. Then, it suffices that #metoo succeeds in raising some women's expectations above these threshold levels, making them willing to speak up. This is compatible with the fact that not *all* women speak up.

person who speaks up when everyone else remains silent, she does prefer to speak up *if* others speak up. The payoff structure for Ada and other women like her is shown in Table 5.

		Other women	
		Remain silent	Speak up
Ada	Remain silent	3, 3	2, 1
	Speak up	1, 2	4, 4

Table 5. Coordination game for the harassed with a speaking-up norm

Since no individual norm-following woman becomes better off by unilaterally changing her strategy, the outcome where everyone speaks up is an equilibrium. Assuming that women consider the social norm of speaking up to be legitimate, that they believe that enough other women will conform, and believe that enough other women believe that women ought to conform, then women will indeed speak up.

Let us now turn to the repeated game. Suppose that men have become aware of the norm of adopting a professional jargon *and* that women have become aware of the norm of speaking up. Furthermore, suppose that, initially, men expect all other men to adopt a professional jargon and expect that all women believe men ought to follow this norm. Suppose also that women expect enough other women to follow the norm of speaking up and expect enough other women to believe that women ought to follow the norm.

Now, suppose some small proportion of men tremble in round t and adopt a macho jargon. Following the tremble, women will speak up. At the end of the round, men will lower their expectations about the proportion of men who will adopt a professional jargon in the next round, but, unlike the situation where there was no social norm of speaking up, men's expectations about the proportion of women who believe men ought to adopt a professional jargon remains unchanged. As long as men's expectations remain above the threshold values d_{emp} and d_{norm} , men will (try to) adopt a professional jargon next round. Therefore, as long as there are no more trembles, men's empirical expectation will eventually recover and approach 1.

This is represented by Figure 3, where the solid line represents the proportion of men who adopt the professional jargon and the (barely visible) dashed line represents men's normative expectations. Each notch in the solid line represents a round where some small proportion of men adopt the macho jargon. The dashed line shows that men's expectations about the proportion of women who believe men ought to adopt a professional jargon does not change when the social norm to speak up is followed.³²

³²In the simulation illustrated in Figure 3, it is assumed that the probability for trembling is 0.025 and that, in the case of a tremble 0.01, of the male population adopts the macho jargon. Women are assumed to always speak up if a man adopts a macho jargon. Men are assumed to use the belief-revision algorithm described above.

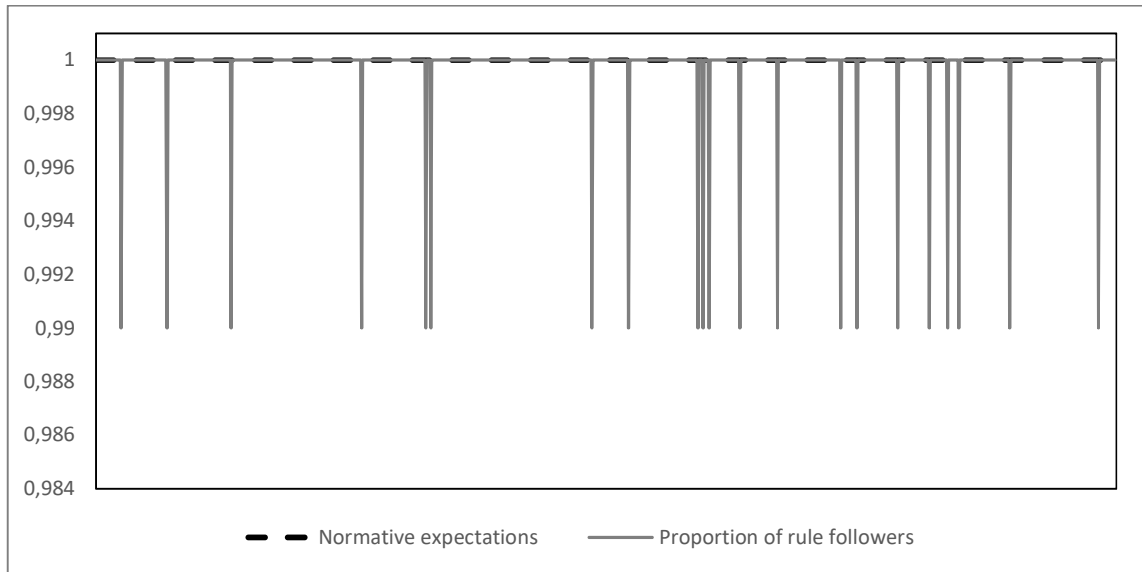


Figure 3. Men's normative expectations with a professionalism norm and a speaking-up norm

There is no risk, then, that men's normative expectations deteriorate below the threshold. Furthermore, if for some reason they were to drop, they would soon recover. This could happen, for example, because some small proportion of women tremble and fail to speak up when some small proportion of men adopt the macho jargon.³³

Under these assumptions, the equilibrium—men adopt a professional jargon and women speak up in case someone adopts a macho jargon—is stable. In other words, there is little risk that small deviations will cause men to fall back into the macho jargon. It is, therefore, tempting to conclude that the #metoo movement will succeed in virtue of reminding men of the social norms of professionalism and creating a speaking-up norm among women.

However, previous consciousness-raising campaigns, from the 1970s feminist movement onwards, have not been observed to produce this effect. The main point of the next subsection is to show why consciousness raising is not enough to secure universal adoption of a professional jargon. Using a broader range of our model's analytical tools, we show that the above equilibrium becomes unstable if we relax the assumption that all men are Normies, and introduce *Jerks*.

³³The model can be extended to allow for 'continuing trembles'. Suppose that, in the normal run of things, a proportion of all men may adopt the macho jargon by mistake. It may or may not be the same men in each period. In this case, men's empirical expectations will adapt to the level of continuing trembles. If there are relatively few trembles, men will continue to adopt the professional norm, even though they expect that some other men will adopt the macho jargon. Furthermore, if some few additional men tremble at a time period, then expectations will drop, only to slowly recover to the level of the continuing trembles.

E. SCENARIO 5: JERKS AND SOCIAL NORMS

Recall that, unlike Normies, Jerks do not consider the social norm ‘when at work, adopt a professional jargon’ legitimate. Following Bicchieri’s model, Jerks are willing to conform to the norm if and only if they expect enough other men to conform, *and* they expect that enough women are willing to punish them if they do not conform.³⁴

Suppose that a proportion of all men are Jerks, and that all men have a belief, b_{sanc} , about the proportion of women willing to punish them if they adopt the macho jargon. Let d_{sanc} be the threshold value under which Jerks are unwilling to conform to the social norm.

We can now say that all Jerks, k , prefer to follow the social norm R =‘at work, use professional jargon’ on the condition that

[*Empirical expectations*:] k believes that enough other men conform to R , i.e., $b_{emp} \geq d_{emp}$, and

[*Normative expectations with sanctions*:] k believes that enough women are willing to punish k for not conforming to R , i.e., $b_{sanc} \geq d_{sanc}$.

As long the two conditions are satisfied, Jerks behave just as Normies do. Thus, if 1) all men expect that enough other men will adopt the professional jargon, 2) Normies expect that enough women believe that men ought to conform to the norm, and 3) Jerks believe that enough women will punish them if they adopt the macho norm, then the situation where everyone adopts the professional jargon is an equilibrium.

However, as before, suppose that the game is repeated. Furthermore, assume that, initially, men expect all other men to adopt a professional jargon, expect that all women believe that men ought to adopt a professional jargon, and expect that all women are willing to punish them if they adopt a macho jargon. From this, it follows that all men (attempt to) adopt a professional jargon in the first round of the game.

What happens if some small proportion of men tremble at round t and adopt a macho jargon? If there is a social norm among women to speak up, then women speak up. This causes men to retain their expectation that all women believe that men ought to conform to the professional norm. Unfortunately, speaking up is not enough to dissuade Jerks from adopting a macho jargon. They must be made to expect that they will be punished if they fail to conform.

Will women be willing to punish Jerks? Prospects seem dim: punishing transgressors is probably at least as costly as speaking up. Speaking up involves condemning the use of a macho jargon in general terms. Punishing transgressors involves further action, towards a specific individual transgressor—who, in the case of workplace harassment, is a colleague, boss, or client. In addition to the cost of being branded as difficult or a feminist killjoy, women who punish transgressors may be considered harmful to their workplace, dangerous if they disrespect due process, or even liars. Moreover, a woman who goes public about an incident of sexual harassment she herself has experienced runs the risk of being subjected to

³⁴Bicchieri 2005.

‘slut shaming’. Therefore, women who contemplate punishing transgressors find themselves in yet another prisoner’s dilemma-like situation, similar to the one outlined in Table 4.

If, then, men use the above rules to update their beliefs, following a tremble they will lower their expectation about the proportion of women willing to sanction non-conformity. Furthermore, since punishment never occurs, this expectation will never recover. As mistakes accrue, men’s expectation about the proportion of women willing to sanction will drop further and further until it drops below the Jerks’ threshold value, d_{sanc} . Once this happens, Jerks no longer want to conform to the social norm.

Suppose this happens at the end of round τ . This means that in round $\tau + 1$, all Jerks will switch to a macho jargon. At the end of round $\tau + 1$, all men lower their expectations about how many other men will adopt the professional jargon in the following round. During the course of the next few rounds, this belief will continue to drop until it approaches the proportion of Normies among the men. If this value is above the threshold value for empirical expectations, d_{emp} , then the population will come to a rest where all Normies adopt a professional jargon and all Jerks adopt a macho jargon. This constitutes an equilibrium of the repeated game, since at this point neither Normies nor Jerks become better off by unilaterally switching strategy.

However, if the proportion of Jerks is large enough to bring the proportion of men who adopt the professional jargon below the threshold value d_{emp} , then men who consider the social norm as legitimate will also be unwilling to conform. This shows that an introduction of Jerks may destabilize a population of Normies who have universally adopted the professional jargon.

Note also that if the Normies have a strong preference against becoming ‘suckers’—that is, d_{emp} is high—then a small number of Jerks may be enough to cause the population to abandon the professional jargon. Moreover, if we relax the assumption that all men have identical threshold points, then chain reactions can occur. Suppose that a population of men consists of one Jerk and many Normies who have different threshold values d_{emp} . One of the Normies prefers to follow the norm only if he believes that everyone else follows the norm, another if he believes that at least everyone else but one follows the norm, and so on. Then, when the Jerk switches to a macho jargon, he will cause the most sensitive Normie to abandon the social norm, which in turn causes the next-most sensitive Normie to abandon the social norm. This process will continue until the macho jargon is universally adopted.

Our analysis thus outlines the limits for the effectiveness of consciousness-raising campaigns, like #metoo and previous movements, in introducing and maintaining a norm of professionalism, thereby reducing the occurrence of workplace sexual harassment. For populations consisting only of Normies, who view this norm as legitimate, such campaigns will help establish a non-harassment equilibrium. However, once Jerks, who consider this illegitimate, are introduced into the population, such an equilibrium will become unstable in the absence of credible sanctions. The remedy is to make the Jerks expect that sufficiently many women are willing to sanction transgressions of the norm.

This concludes our answer to our initial guiding questions concerning why sexual harassment remains a problem, despite consciousness-raising campaigns since at least the 1970s, and concerning the mechanisms behind sexual harassment. The final section of this article catalogues different kinds of sanctions and explores their potential to get Jerks to expect that sufficiently many women are willing to sanction. This will answer our guiding question concerning the practical implications of an improved understanding of sexual harassment and its mechanisms.

IV. CONCLUDING REMARKS

There are different types of sanctions, all compatible with our model. *Formal* sanctions use the rules of due process, both in identifying perpetrators and in deciding the appropriate sanction. The formal rules may, in turn, be legal or institutional. Legal rules may mandate prison sentences or fines, whereas institutional rules, in a workplace setting, may mandate warnings, demotions, or sackings.

One problem with formal sanctions, experience tells us, is that only a fraction of all incidents is reported.³⁵ Our model offers an explanation: the dilemma of the harassed, facing victims of sexual harassment as well as bystanders, resurfaces here. Just as for the strategy of speaking up, reporting harassers to the authorities or to one's employer is connected with significant individual costs, while potential benefits are a public good.

A second problem, especially concerning the use of legal sanctions, is that, for incidents that are reported to the authorities, the percentage of legal trials is rather low, and that of subsequent convictions much lower still.³⁶ This is in part due to the relatively high bar of proof that is required by our legal systems—and reasonably so. There are compelling arguments for maintaining a high evidential threshold to avoid false positives, that is, convicting innocent citizens by state authority. However, where low rates of convictions—in addition to low rates of reporting—are well-known facts, rational Jerks will not be deterred from committing sexual harassment.

The classical remedy for resolving a prisoner's dilemma is changing the payoff structure. One could *decrease women's (expected) costs* of reporting: for example, by making reporting procedures more easily accessible, training police officers and HR personnel to handle harassment cases efficiently and respectfully, and so on. Still, as long as

³⁵A recent report of the US Equal Employment Opportunity Commission (EEOC) states that that 'gender-harassing conduct [in the workplace] was almost never reported; unwanted physical touching was formally reported only 8% of the time; and sexually coercive behavior was reported by only 30% of the women who experienced it'; <https://www.eeoc.gov/eeoc/task_force/harassment/upload/report.pdf>. Moreover, of all allegations of sexual harassment made to the EEOC between 2010 and 2018, less than 1 in 4 has resulted in a 'favourable' outcome (with 54% of the reports dismissed as having 'no reasonable cause', and another 22% dismissed on administrative grounds; <https://www.eeoc.gov/eeoc/statistics/enforcement/sexual_harassment_new.cfm>. To compare with statistics for one of the most severe forms of sexual harassment—rape—according to RAINN, a US anti-sexual-violence organization, out of 1000 cases of rape, 230 are reported to the police, 46 of those lead to arrest, and just 5 of those lead to a felony conviction; see <<https://www.rainn.org/statistics/criminal-justice-system>>.

³⁶See e.g. n. 35 above, regarding rape conviction rates.

there are costs to reporting, the model predicts that willingness to sanction will remain low, and hence Jerks will remain undeterred.

Or one could *increase the (expected) benefits* of reporting: for example, by increasing the rate of conviction, giving women better prospects of punishing their harassers. This might be harder to justify within a legal context, given the above rationale for maintaining a high bar of proof. Yet thresholds of evidence may reasonably be set lower within systems of non-legal institutional rules, for example, in the workplace. An employer may, in the interest of maintaining a well-functioning organization, choose to sanction suspected harassers, even knowing that they would not have been convicted in court. Moreover, in workplace settings there are further measures available for reducing the cost of reporting: for example, technical solutions for anonymous reporting, which detect repeat perpetrators and allow victims to ‘activate’ their reports conditional on others reporting the same individual, thereby reducing their risk from ‘going it alone’.³⁷

One upshot of the model, and the role it assigns to sanctions for combating sexual harassment, is thus that we should pay more attention to institutional settings: workplaces, schools, and other organizations. They are important venues for upholding norms of non-harassment—not just on a consciousness-raising policy level, but also regarding the availability of grievance procedures and use of sanctioning mechanisms. Only the last provides Jerks with reasons to update their relevant expectations and adjust their behaviour.

Still, there are plenty of contexts where institutional grievance and sanctioning procedures do not apply. Some employers simply do not acknowledge their responsibility to provide those procedures; the emergence of ‘gig economy’ jobs undermines employer responsibility on a larger scale; and, moreover, plenty of sexual harassment is committed outside of institutional settings such as the workplace. Where institutional sanctions are unavailable and/or very costly, while legal sanctions have both high (expected) costs and low (expected) benefits, we are thus left to consider *informal* sanctions. Lacking rules of due process, informal sanctions imply ‘taking matters into one’s own hands’—for example, naming and shaming harassers on social media.

Do informal sanctions have the potential to deter Jerks from sexually harassing women? Again, there are considerable costs to naming and shaming, as has been vividly shown by the backlash against women who outed their harassers on social media in the wake of #metoo. Many of them have themselves been publicly outed and condemned; some have also been legally prosecuted under libel laws. As long as there are costs connected to naming and shaming, the dilemma of the harassed resurfaces once again.

Again, we may try to solve the dilemma by changing the payoff structure: we may, for example, try to offset the backlash against a namer-and-shamer by supporting her against her critics, emphasizing her contribution to upholding a non-harassment norm. We may even attempt to limit the applicability or harshness of libel law under certain circumstances.

³⁷For such a solution, implemented at college level, see <<https://www.projectcallisto.org>>.

One final upshot of our model is that the list of possible remedies can be extended beyond the classic strategy of overcoming a prisoner's dilemma by changing the payoff structure, through reducing costs and/or increasing benefits. Our analysis highlights a further, less discussed strategy: the introduction of a social norm, which transforms the prisoner's dilemma game with respect to sanctioning into a coordination game. Jerks can be held in check if a *social norm to sanction* transgressors is introduced: for example, 'whenever you observe sexual harassment, sanction the harasser (by available means)'. In contexts where legal reporting holds little promise of administering sanctions and where institutional sanctions do not apply, this norm effectively calls for informal sanctions.

If women come to expect other women to conform to such a sanctioning norm, and expect that other women believe that women ought to conform to this norm, they will be willing to sanction men who do not adopt the professional jargon. Thus, the Jerks' expectations of punishment will be matched by the actual behaviour of women. Accordingly, men will not lower their expectations about the proportion of women willing to punish in case some men tremble and adopt a macho jargon by mistake. Thus, if a sanctioning norm is introduced, it is possible to make the universal adoption of a professional jargon a stable equilibrium, even when a significant part of the population consists of Jerks.

Critics of the #metoo movement may have many good reasons to oppose an introduction of such a norm, as well as the suggested strategies of changing the payoff structures underlying the harassed's failure to, formally or informally, sanction harassers. However, in pointing out the role sanctions play, our model highlights the incompatibility of such general opposition with a serious commitment to fight sexual harassment.

Our reply, then, to Saul's challenge to make progress on what to do about sexual harassment is this:³⁸ in the wake of #metoo, if we are still serious about eliminating sexual harassment, we need to keep up the consciousness-raising efforts, but also put sanctions on the table.

REFERENCES

- Anderson, D. E. 2017. Conceptual competence injustice. *Social Epistemology*, 31, 210–23.
- Anderson, E. 2000. Beyond homo economicus: new developments in theories of social norms. *Philosophy and Public Affairs*, 29, 170–200.
- Anderson, E. 2006. Recent thinking about sexual harassment: a review essay. *Philosophy and Public Affairs*, 34, 284–311.
- Bicchieri, C. 2005. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge: Cambridge University Press.
- Bicchieri, C. 2016. *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. Oxford: Oxford University Press.
- Bicchieri, C. and Y. Fukui. 1999. The great illusion: ignorance, informational cascades, and the persistence of unpopular norms. *Business Ethics Quarterly*, 9, 127–55.

³⁸Saul 2014; see sect. I above.

- Bowes-Sperry, L. and A. M. O'Leary-Kelly. 2005. To act or not to act: the dilemma faced by sexual harassment observers. *Academy of Management Review*, 30, 288–306.
- Crouch, M. A. 2001. *Thinking about Sexual Harassment: A Guide for the Perplexed*. Oxford: Oxford University Press.
- Dotson, K. 2012. A cautionary tale: on limiting epistemic oppression. *Frontiers: A Journal of Women Studies*, 33, 24–47.
- Farley, L. 1978. *Sexual Shakedown: The Sexual Harassment of Women on the Job*. New York: McGraw-Hill.
- Fricker, M. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.
- Gibbard, A. and H. R. Varian. 1978. Economic models. *Journal of Philosophy*, 75 (11), 664–677.
- Gilbert, M. P. 2008. Social convention revisited. *Topoi*, 27, 5–16.
- Jackson, D. L. 2018. 'Me too': epistemic injustice and the struggle for recognition. *Feminist Philosophy Quarterly*, 4 (4), (online first).
- Lewis, D. 1969. *Convention: A Philosophical Study*. Oxford: Blackwell.
- MacKinnon, C. A. 1979. *Sexual Harassment of Working Women: A Case of Sex Discrimination*. New Haven: Yale University Press.
- Mason, R. 2011. Two kinds of unknowing. *Hypatia*, 26, 294–307.
- O'Hare, E. A. and W. O'Donohue. 1998. Sexual harassment: identifying risk factors. *Archives of Sexual Behavior*, 27, 561–80.
- Pina, A., T. Gannon, and B. Saunders. 2009. An overview of the literature on sexual harassment: Perpetrator, theory, and treatment issues. *Aggression and Violent Behavior*, 14, 126–38.
- Saul, J. 2014. Stop thinking so much about 'sexual harassment'. *Journal of Applied Philosophy*, 31, 307–21.
- Selten, R. 1975. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4, 25–55.
- Siegel, R. B. 2003. Introduction: a short history of sexual harassment. Pp. 1–40 in C. A. MacKinnon and R. B. Siegel (eds), *Directions in Sexual Harassment Law*. New Haven: Yale University Press.