Addressing Online Violence Against Women and Girls in the Commonwealth Pacific Region

The Role of Bystanders

The Commonwealth

Addressing Online Violence Against Women and Girls in the Commonwealth Pacific Region

THE ROLE OF BYSTANDERS

The Commonwealth

Foreign, Commonwealth & Development Office

Commonwealth Secretariat
Marlborough House
Pall Mall
London SW1Y 5HX
United Kingdom

www.thecommonwealth.org

Published by the Commonwealth Secretariat.

Contents

List of Tables

Acronyms and abbreviations

APC	Association for Progressive Communications
LGBTQ	Lesbian, gay, bisexual, transgender and queer
UN Women	United Nations Entity for Gender Equality and the Empowerment of Women
VAWG	Violence against women and girls

Acknowledgement

The Commonwealth Secretariat acknowledges with gratitude the financial support of the United Kingdom Foreign, Commonwealth & Development Office to the Commonwealth Cyber Capability Programme.

The report on *Addressing Online Violence Against Women and Girls in the Commonwealth Pacific* Region: The *Role of Bystanders*, is part of a series which investigates the culpability of bystanders in violent act committed against women and girls on the cyberspace.

The report was authored by Professor Rob McCusker, Consultant in Transnational Crime.

The series was prepared under the general guidance of Dr Tawanda Hondora, Adviser and Head of Rule of Law Section, Governance and Peace Directorate (GPD). Dr Nkechi Amobi, Senior Research Officer, Cyber Capability Programme GPD, led and co-ordinated the review and editorial process of the report. Ms Emma Beckles, Programme Officer, GPD and Mr Shakirudeen Ade Alade, Programme Coordinator GPD provided valuable input while Ms Helene Massaka, Programme Assistant GPD, provided logistical and administrative support.

The team is grateful to Mrs Elizabeth Bakibinga-Gaswaga, former Legal Adviser Rule of Law Section, GPD, for conceptualising this research project.

The team is also grateful for the constructive feedback received from internal reviewers- Ms Jennifer Namgyal, Adviser Gender Mainstreaming, Economic, Youth, and Sustainable Development Directorate (EYSD) and Ms Monika Pinder, Programme Officer, Gender, and Development, EYSD, and Clive Lawson, Publications Assistant, Communication Division.

Executive summary

This report focuses on the role of online bystanders in reducing online violence against women and girls (VAWG), including consideration of any related policies or laws. Given the relatively recent emergence of bystanders as potential facilitators and exacerbators of online crime, policy responses have tended to focus on the virtual versions of existing physical crimes, such as 'cyber' bullying and 'cyber' stalking, and/or the dissemination of intimate images online (when previously they might have been posted on a physical bulletin board). Moreover, these policy responses have tended to focus on capturing perpetrators, but not bystanders, within a legislative framework.

In order to assess the degree to which such bystanders can provide assistance in this regard, it is necessary initially to establish precisely (a) how the notion of 'online violence' should be defined and calibrated, (b) how online bystanders contribute, directly or indirectly, to the perpetration and perpetuation of online violence and (c) why and how legislation regarding the responsibility of bystanders to act should be drafted and applied. In that context, this report explores the notion, types, range, scale and impact of online violence only as a backdrop to the examination and analysis of the role and potential legislative requirements of online bystanders.

As already noted, VAWG in the online space has until recently tended to be examined in terms of offences that are housed in existing legislation on cybercrime. A key driver for a new legislative response is that online violence is now too widespread, broad (in terms of potential offence range) and extra-territorial to be dealt with through existing, sometimes generic, laws rather than through a dedicated and more sophisticated legislative framework. Moreover, whether in a physical or virtual setting, current legislation focuses entirely on the perpetrator and not on bystanders as enablers to the perpetrators' offending. There is a need, therefore, to consider whether – and in what way – bystanders might be brought within the provisions of existing, amended or new legislation. Crucially, in terms of the latter, within the Pacific region (that is, Australia, New Zealand and the Pacific countries), only Australia has created 'good Samaritan' legislation and – save for one Act in one part of Australia (Criminal Code Act 1983, s.155, Northern Territory) – no jurisdiction has created 'bad Samaritan' legislation, which aims principally to penalise bystanders who fail to intervene. In terms of any future creation of such legislation, a full understanding is required of its rationale, the difficulties involved in its creation and its complexities in terms of implementation and enforcement.

1. Introduction

1.1 The nature of online violence against women and girls

Central to any assessment and analysis of online violence, and of the nature and role of online bystanders to that violence, is understanding what 'violence' means in the virtual environment. It differs in both form and context from that exhibited in the physical world, although there are some points of similarity. For example, violence in the physical world may not actually encompass physical injury, and such is primarily the case in the virtual world, where expressions of hatred, attempts at ridicule and similar behaviour may lead to the same visceral impact without its physical manifestation. Equally, violence perpetrated in a virtual context may lead to others responding by physically assaulting the target, such as instances where physical bullying occurs against a victim of cyberbullying. In that broad context, therefore, the first issue is to determine what constitutes online violence against women and girls (VAWG).

VAWG in the physical space is related to violence perpetrated in the online space since both rely on ill-conceived perceptions of women and girls as justification of, or rationalisation for, the violent behaviour. It follows, therefore, that addressing those perceptions in the physical space may assist in actions in the virtual space being likewise reflected and acted on. At the same time, however, the characteristics of cyberspace can provide offenders with a safe environment from which to launch their violent behaviour.

The United Nations has defined VAWG as '[a]ny act of gender-based violence that results in, or is likely to result in, physical, sexual or psychological harm or suffering to women, including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or in private life' (UN General Assembly 1993). Such harm would include forced intimate partner violence and sexual assault, dowry-related violence, marital rape, sexual harassment, intimidation at work and in educational institutions, forced pregnancy, forced abortion, forced sterilization, trafficking and forced prostitution and gender-related killings.

The 1994 Inter-American Convention on the Prevention and Eradication of Violence against Women defined violence against women as '...any act or conduct, based on gender, which causes death or physical, sexual or psychological harm or suffering to women, whether in the public or private sphere' (OAS 1994).

These definitions pre-date both the arrival and concomitant impact of communication within, and abuse and exploitation of, the online space. When the term 'cyber' is applied to instances of VAWG, it should be recognised that the internet can exacerbate such violence and that the typologies of sexual violence are likely to continue to evolve as the digital and virtual platforms, on which the violence is perpetrated and encouraged, continue to develop.

The first challenge in any discussion pertaining to online VAWG thus lies in the huge variety of actions that might be deemed to fit within the confines of 'online violence'.

A lack of universality in approach has led to the creation of different categories and sub-categories in a number of jurisdictions. Aware of this difficulty, the Council of Europe chose 'cyberviolence' as the overarching and encompassing term around which to build a typology of offending. It defined this as '...the use of computer systems to cause, facilitate, or threaten violence against individuals that results in, or is likely to result in, physical, sexual, psychological or economic harm or suffering and may include the exploitation of the individual's circumstances, characteristics or vulnerabilities' (Council of Europe 2018). It recognised that, in practice, cyberviolence might entail different forms of activity that, although equally severe in terms of impact and effect on the victim, might not all require the intervention or application of the criminal law.

Its broad typology (see Table 1.1) clearly demonstrates the scope of cyberviolence and indicates that not all the types of crime identified lend themselves naturally, or at all, to the presence of bystanders, let alone the facility for enabling or encouraging bystanders to respond in defence of the victim.

Table 1.1. Examples of different types of cyberviolence

ICT-related violations of privacy	ICT-related hate crime	ICT-related direct threats of or physical violence	Cyber-harassment	Cybercrime	Online sexual exploitation and sexual abuse of children
• Computer intrusions • Taking, sharing, manipulation of data or images, including intimate data • Sextortion • Stalking • Doxing • Identity theft • Impersonation	Against groups based on • Race • Ethnicity • Religion • Sex • Sexual orientation • Disability	• Murder • Kidnapping • Sexual violence • Rape • Torture • Extortion • Blackmail • Swatting • Incitement to violence • Transmissions that themselves cause injuries • Attacks on critical infrastructure, cars or medical devices	• Defamation and other damage to reputation • Cyberbullying • Threats of violence, including sexual violence • Coercion • Insults or threats • Incitement to violence • 'Revenge porn' • Incitement to suicide or self-harm	• Illegal access • Illegal interception • Data interference • System interference • Computer-related forgery • Computer-related fraud • Child pornography	• Sexual abuse • Child prostitution • Child pornography • Corruption of children • Solicitation of children for sexual purposes • Sexual abuse via livestreaming

Source: Council of Europe 2018.

The Learning Network has posited six broad categories in which areas of abuse are located (Baker et al. 2013):

1. Hacking: using technology to obtain illegal or unauthorised access to systems for the purpose of obtaining personal information, altering information or slandering individuals and/or their organisations
2. Impersonation: using technology to assume the identity of an individual to access private information or send offensive emails purporting to emanate from the victim
3. Surveillance/tracking: using technology to stalk and monitor a victim's activities and behaviour
4. Harassment/spamming: using technology to continuously contact, annoy, threaten and/or scare the victim
5. Recruitment: using technology to lure potential victims into violent situations, including trafficking victims
6. Malicious distribution: using technology to manipulate and distribute defamatory and illegal materials relating to the victim and/or VAWG organisations

Beyond that broad typology, the Association for Progressive Communications (APC) has identified five unifying characteristics that distinguish cyber VAWG (APC 2017):

1. Anonymity – the abusive person can remain hidden from the victim
2. Action at a distance – the abuse can be launched virtually from any point
3. Automation – abuse delivered via technology is easier to create and disseminate
4. Accessibility – perpetrators have an array of technology available to them
5. Propagation and perpetuity – texts and images may exist virtually for extended periods of time

There have also been other categorisations of cyber VAWG. One, for example, has given rise to the term 'technology-facilitated sexual violence', which posits five instances where technology is used to facilitate or exacerbate sexual and gender-based harm to victims (Powell and Henry 2017):

1. Technology-enabled sexual assault
2. Image-based sexual abuse
3. Cyberstalking and criminal harassment
4. Online sexual harassment
5. Gender-based harassment and hate speech

The Organization of American States (OAS) has suggested that online violence '... may involve threatening or harassing emails, instant messages, or posting information online' and 'targets a specific person either by directly contacting them or by disseminating their personal information, causing them distress, fear or anger' (OAS 2019). The Pew Research Center posited at least six distinct forms of harassment that would fall within the generic category of online violence: (1) offensive name-calling, (2) purposeful embarrassment, (3) physical threats, (4) sustained harassment, (5) sexual harassment and (6) stalking (Duggan 2017).

APC (2017) posits the notion of 'technology-related violence', which it configures as acts '...committed, abetted or aggravated, in part of fully, by the use of information and communication technologies such as mobile phones, the internet, social media platforms and email'. It has been suggested by the Internet Governance Forum (2015) that acts of online violence are often '...an extension of existing gender-based violence, such as domestic violence, stalking and sexual harassment, or target the victim on the basis of her gender or sexuality'. In broad indicative terms, the Forum regards a number of acts as constituting online VAWG (see Table 1.2).

However online VAWG is characterised and defined, it is clear that online technologies, from smart phones and email to social networking and online dating sites, are being used to carry out sexual assault in online spaces. The generic heading of 'revenge pornography' (colloquially 'revenge porn') for the non-consensual creation and distribution of sexual or intimate images has been deemed problematic. This is because that term does not fully capture the range of perpetrator motivations, which extend beyond revenge to, for example, distributing images for profit.

Moreover, APC (2017) has also argued that 'revenge porn' is something of a misnomer as it attributes some degree of blame to the victim. The term 'revenge' implies that the poster of the content had been provoked by the inappropriate behaviour

Table 1.2. Acts constituting online violence against women and girls

Infringement of privacy	Surveillance and monitoring	Damaging reputation and/or credibility	Harassment	Direct threats and/or violence	Targeted attacks to communities
• Accessing, using, manipulating and/or disseminating private data without consent (by cracking personal accounts, stealing passwords, using/stealing identities, using another person's computer to access a user's accounts while it is logged in, etc.) • Taking, accessing, using, manipulating, and/or disseminating photographs and/or videos without consent (including 'revenge pornography') • Sharing and/or disseminating private information and/or content, including (sexualised) images, audio clips and/or video clips, without knowledge or consent • Doxing (researching and broadcasting personally identifiable information about an individual without consent, sometimes with the intention of providing access to the woman in the 'real' world for harassment and/or other purposes) • Contacting and/or harassing a user's children, extended family, colleagues (etc.) to gain access to her	• Monitoring, tracking and/or surveillance of online and offline activities • Using spyware or keyboard loggers without a user's consent • Using GPS or other geolocator software to track a woman's movements without consent • Stalking	• Deleting, sending and/or manipulating emails and/or content without consent • Creating and sharing false personal data (like online accounts, advertisements, or social media accounts) with the intention of damaging a user's reputation • Manipulating and/or creating fake photographs and/or videos • Identity theft (e.g., pretending to be the person who created an image and posting or sharing it publicly) • Disseminating private (and/or culturally sensitive/controversial) information for the purpose of damaging someone's reputation • Making offensive, disparaging and/or false online comments and/or postings that are intended to tarnish a person's reputation (including libel/defamation)	• 'Cyber bullying' and/or repeated harassment through unwanted messages, attention and/or contact • Direct threats of violence, including threats of sexual and/or physical violence (e.g., threats like 'I am going to rape you') • Abusive comments • Unsolicited sending and/or receiving of sexually explicit materials • Incitement to physical violence • Hate speech, social media posts and/or mail; often targeted at gender and/or sexuality • Online content that portrays women as sexual objects • Use of sexist and/or gendered comments or name-calling (e.g., use of terms like 'bitch'/'slut') • Use of indecent or violent images to demean women • Abusing and/or shaming a woman for expressing views that are not normative, for disagreeing with people (often men) and also for refusing sexual advances • Counselling suicide or advocating femicide • Mobbing, including the selection of a target for bullying or harassment • Mobbing by a group of people rather than an individual and as a practice specifically facilitated by technology	• Trafficking of women using technology, including use of technology for victim selection and preparation (planned sexual assault and/or femicide) • Sexualised blackmail and/or extortion • Theft of identity, money and/or property • Impersonation resulting in physical attack	• Cracking websites, social media and/or email accounts of organisations and communities with malicious intent • Surveillance and monitoring of activities by members of the community • Direct threats of violence to community members • Mobbing, specifically when selecting a target for bullying or harassment by a group of people, rather than an individual, and as a practice specifically facilitated by technology • Disclosure of anonymised information like addresses of shelters, etc.

Source: Internet Governance Forum 2015.

of the victim. 'Revenge porn' also suggests, in a perhaps moralistic tone, that the material distributed is definitionally 'pornographic' in nature, to the point that the victim may receive a degree of societal judgment in addition to the harm caused by the posting itself. In that sense, 'image-based sexual abuse' might be a more apposite term.

In that broader context, sub-categories such as 'sextortion' occur, where perpetrators obtain images and then threaten to distribute them if the victim fails to pay the perpetrator not to do so.

Cyberstalking has been deemed to be '... an extension of conventional stalking using electronic means' (Powell and Henry 2017). However, difficulties have occurred with the application of existing legislation on physical stalking to the cyber variety. In broad terms, legislation governing stalking requires a repetition of behaviour rather than a one-off occurrence. In the case of cyberstalking, however, there may be a range of behaviour associated with what is ostensibly one event. Thus, for example, the one-off action of posting a sexually explicit image online without consent may be accompanied by identifying information and an invitation for others to contact, harass or injure the victim. Equally, the actions of the perpetrator in terms of, for example, posting offensive, malicious or personal information, might be humiliating for the victim but not elicit the visceral fear or apprehension that legislation often specifies, and that arguably features, in physical stalking cases.

Powell and Henry (2017) have suggested that any assessment of behaviour leading to technology-facilitated sexual violence must recognise factors such as gender inequality and unequal power dynamics as underlying issues. Indeed, the Organization of American States recognised the '...structural factors that affect violence against women and socio-cultural and symbolic standards as well as social and cultural stereotypes that perpetuate it' (OAS 2015). UN Women (2020) similarly argues that the key underlying causes of violence against women include '...gender inequality and power imbalances between women and men, reinforced by discriminatory and gender-biased attitudes, norms and practice'. It suggests that pertinent risk factors in this regard include '... inequitable cultural and social norms that support male authority over women, condone or trivialize [violence against women and girls], and stigmatize victims/survivors'. Indeed, the situation is exacerbated by social norms placing a lower value on women and girls, which may then be manifested by a high tolerance and/or acceptance of violence against them.

While women and girls can be at risk solely due to their gender, their experience may be exacerbated by a range of other factors including race, ethnicity, language, sexual orientation or gender identity, age, disability, nationality, migrant status, religion and whether they live in an urban or rural location. In addition, when women have intersecting identities, they are arguably easier to target online and that in turn may exacerbate their experiences of abuse. For example, women who are human rights advocates, active in politics, journalists or from the lesbian, gay, bisexual, transgender and queer (LGBTQ) community are frequent targets of online violence. A regional survey by the Inter-Parliamentary Union (2016) noted that social media had become the primary space in which online violence (in the form of '... sexist and misogynistic remarks, humiliating images, mobbing, threats and intimidation ...') was perpetrated against women parliamentarians. A global survey by the International Women's Media Foundation indicated that more than 25 per cent of verbal, written and/or physical intimidation (including threats to family or friends) occurred online (Barton and Storm 2014).

APC (2017) has argued that the same forms of gender discrimination in social, economic, cultural and political structures that lead to gender-based violence are reproduced and perhaps exacerbated in the online space. Women and girls face specific threats online, including harassment, cyberstalking, attacks on their sexuality, exposure of personal information, manipulation of images and non-consensual distribution of intimate images or videos. APC (ibid.) argues that '...the technology dimension adds elements of searchability, persistence, replicability and scalability which facilitate aggressors' access to women they are targeting and can escalate and exacerbate harm'.

UN Women (2020) has noted, however, that women and girls tend to be internet users with limited digital skills and that, subsequently, they may be more at risk of online violence than their more skilled/experienced counterparts. The Sustainable Development Goals therefore include targets not only for ending discrimination against women and girls but also for enhancing the use of technology to increase their empowerment.

It is also clear that online and offline violence can coincide or coexist. Thus, online violence can be supplemented or exacerbated by offline violence, including harassment, vandalism, telephone calls and physical assault. The viral nature of online violence also means that the scope of impact is widened considerably. The European Parliament (2018) has observed a phenomenon of 'continuity of digital spaces', where the victim is targeted simultaneously across several social media platforms and via messaging apps and email in what can be a coordinated pattern of attacks.

A concomitant effect of the advances in technology, for example, the 'Internet of Things', also threatens to expand the scope of online offending. The Internet of Things strives to connect every machine, residence and vehicle to an intelligent communications infrastructure with responding chips being placed in everyday devices. Already, technology including smart phones, smart cameras, watches, toys and tablets has been exploited in online violence against ex-partners. This has been characterised by Yee Man (2021) as technology-facilitated domestic abuse, which is defined as '... the use of digital technologies to control, coerce, intimidate, humiliate, stalk, or harass an intimate partner (usually a female) both during a relationship and after separation'. She has argued that women from culturally and linguistically diverse backgrounds are particularly vulnerable (ibid.), an assertion supported, inter alia, by the Office of the eSafety Commissioner in Australia (eSafety Commissioner undated).

1.2 Responding to online VAWG

The first point of note in relation to legislation is that jurisdictions have tended, as a result of the historical evolution of cybercrime, to configure their laws around discrete themes such as 'cyber security', 'cyber fraud' and more recently 'cyber harassment'. Other legislation, not necessarily with a readily apparent cybercrime component, has tended to evolve in response to changes in the threat environment. Those changes have been created, facilitated or exacerbated by developments in technology, cyberspace more broadly and the inevitable negative consequences of the wholesale embracing (by some societal demographics more than others) of social media in all its various forms, each of which can be used for exploitation and abuse.

In large part, online violence has traditionally been countered indirectly, in the sense that legislation may, for example, have prohibited certain use or exploitation of computers or computer violence against women and girls. Thus, there are measures in place in several jurisdictions that deal with certain offences committed in relation to social media postings. For example, posting non-consensual images on the internet may be captured by legislation, which renders illegal 'revenge porn', voyeurism, harassment, extortion or defamatory libel actions.

Equally, if a video depicts a person under the age of 18 engaged in sexual activity, or if the main purpose of the video is the depiction of that person's sexual organs, this could be covered by child pornography legislation.

However, the degree to which such provisions encapsulate the behaviour of online bystanders who view and/or share and/or comment on such images is debatable. Thus, in broad terms, 'revenge porn' generally requires the images to be intimate and the person featured in them to have had a reasonable expectation of privacy in relation to the image circulated. Criminal harassment requires that the victim feared for their safety, or the safety of someone known to them. Voyeurism requires the image to be taken surreptitiously or the offence cannot usually be made out.

These constructions do not facilitate the inclusion of online bystanders who view and/or share the images taken by the original offender. In other words, laws prohibiting online posting reflect the fact that undertaking proceedings against the main offender was seemingly the most effective way of assisting the victim. However, the very nature of the original offence, allied with the behaviour of online prurient observers, means that the damage continues long after the original offence has occurred. In that broad sense, the offending continues unabated.

Equally, from a largely pragmatic viewpoint, it might be possible that an offence of, for example, 'interfering with a telecommunications service' (Telecommunications Act 2019, (No 7) Cook Islands, ss.85-87) could, if argued appropriately in a court setting, capture an assailant's entry into a victim's smart phone for the purposes of abstracting images or data for subsequent dissemination online as an act of violence. This is

certainly the situation in the Pacific Islands where, except for direct references to the abuse of devices for the purposes of child pornography or child exploitation, cyber bullying, cyber harassment and cyber extortion, it would be necessary, and certainly possible, for online violence offences to be prosecuted under the more widely defined cybercrime legislation in each jurisdiction (see Table 1.3). These could include 'sending messages of an offensive, indecent, obscene or menacing character' (Telecommunications Act 2002, Nauru, s.45.) and '... modification, interception and disclosure of messages' (Telecommunications Act 2004, Kiribati, Part VI.).

The position has advanced by degrees. For example, in New Zealand, the Harmful Digital Communications Act 2015 created a criminal offence of posting harmful digital communications (s.22). Under Part 1 of this Act (Approved Agency and enforcement, Subpart 1—Purpose, interpretation, the Crown, and communication), 'harm' is defined as meaning 'serious emotional distress' (4 Interpretation) and the digital communications referred to should not disclose 'personal facts about an individual' (6 Communication, Principle 1), should not be 'grossly offensive to a reasonable person in the position of the affected individual' (6 Communication, Principle 3) and should not 'contain a matter that is published in breach of confidence' (6 Communication, Principle 7).

However, the legislation has been found wanting because, for example, the perpetrator must have intended to cause serious emotional distress to the complainant and there must have been serious harm resulting from the disclosure of the images complained of. Reaching a determination as to the causation of harm, consideration is given to factors such as the age and characteristics of the victim and the extent to which the images were circulated. The Act also does not prohibit the taking of nude or sexual images without consent (such as 'upskirting') or making threats to share images ('sextortion'). However, such activities can be proceeded against in New Zealand under alternative legislation such as the Crimes Act 1961, which (under the provisions of the Crimes (Intimate Covert Filming) Amendment Act 2006) makes it a criminal offence to take an 'intimate visual recording' (s.216G) of another person using any device without their knowledge or consent, where there is a reasonable expectation of privacy.

Further, in New Zealand, under the Films, Videos and Publications Classification Act 1993, it is a criminal offence to make an 'objectionable publication', that is, one that depicts 'sex, horror, crime, cruelty or violence in such a manner that the availability of the publication is likely to be injurious to the public good' [Part 1 s. 3(1)].

Since the aforementioned Act, while many pieces of legislation contain broadly defined overarching technology protection provisions (which could arguably be configured to online violence offences), there has been a rise in provisions that specifically speak to the typology of offences indicated above.

In Australia, for example, under Federal law, the Criminal Code Act 1995 (Commonwealth) includes offences in its telecommunications section on the use of a carriage service for child pornography material (s.474.19) and for making a threat to kill [474.15 (1) (a) and (b)] or cause serious harm (ibid.). Also, under s.474.17 of the Crimes Legislation Amendment (Telecommunications Offences and Other Measures) Act (No.2) 2004, it is an offence to use a carriage service in such a way that is '... menacing, harassing or offensive', and this has been utilised in so-called 'revenge porn' cases (more appropriately referred to, for reasons noted earlier, as 'non-consensual creation and distribution of sexual or intimate images').

The Federal Parliament in Australia creates legislation that applies to the whole country while six states (Queensland, Victoria, Western Australia, New South Wales, South Australia and Tasmania) and two territory (Australian Capital Territory and Northern Territory) parliaments create legislation for their respective state or territory. A number of states have elected to deal with image-based sexual abuse more overtly within their legislation (see Table 1.4). South Australia championed the issue with its offence (Summary Offences Act 1953 (SA), s26C: Distribution of invasive image) criminalising the non-consensual distribution of 'invasive images' in a situation where it can be proven that the distributor knew, or should have known, that the victim did not consent.

In Victoria, it is an offence (Summary Offences Act 1966 (Vic), s41DA and s41DB) to intentionally distribute, or threaten to distribute, 'intimate' images of another person without their consent. Victoria also has the Summary Offences Amendment (Upskirting) Act 2007, which renders

Table 1.3. Pacific laws on cybercrime

Country	Statutes	Forms of cybercrime
Cook Islands	Telecommunications Act 2019	Ss85–87: Interfering with telecommunications service Ss135–138: Offences related to indecent material, which may apply to online child pornography
	Copyright Act 2013	Offences for infringement of copyright
Fiji	Crimes Act 2009	Part 17, Division 6: Offences for unauthorized modification of data to cause impairment; unauthorized impairment of electronic communication; unauthorized access to (or modification of) restricted data; unauthorized impairment of data held on a computer disk; and possession, control, production, supply or obtaining of data with intent to commit a computer offence; Part 11, Sub-division D: Forgery and related offences
	Posts and Telecommunications Decree 1989	Part IV: offences for modification, interception and disclosure of messages
Kiribati	Telecommunications Act 2004	Part VI: Offences for modification, interception and disclosure of messages Part VII: Computer misuse offences – unauthorized access, modification, use or interception Part VI: Grossly offensive use of a telecommunications system Part VIII: Distribution and exhibition of obscene matter, including in electronic form, which may apply to child exploitation material
Nauru	Cybercrime Act 2015	Part 2: Offences for illegal access, interception, data interference, data espionage, system interference, distributing/possessing software/a device for committing a crime; computer-related forgery and fraud, including identity theft; producing, offering, distributing, procuring, possessing or knowingly obtaining access to child pornography 'through an electronic system'; online solicitation of children; publishing of indecent material in electronic form
	Telecommunications Act 2002	Ss43–44: Offences for intercepting, using or disclosing communications (applicable only to telecommunications employees) and interfering with communications (by any person) S45: Sending messages of an offensive, indecent, obscene or menacing character (via telecommunications)
	Crimes Act 2016	Ss139–143: Offences for pornography and exposing child to offensive material S189: Offence for making or possessing device for making false document

Papua New Guinea	Cybercrime Code Act 2016	Ss6–31: Extensive offences for unauthorized access or hacking, illegal interception, data interference, system interference, data espionage, illegally remaining, electronic fraud, electronic forgery, electronic gambling or lottery by a child, identity theft, illegal devices, pornography, child pornography, child online grooming, animal pornography, defamatory publication, cyber bullying, cyber harassment, cyber extortion, unlawful disclosure, spam, cyberattack, online copyright infringement, online trade-mark infringement, patent and industrial designs infringement, unlawful advertising
	Protection of Private Communications Act 1973	Part II: Offences for intercepting private communications and divulging such communications
	Criminal Code Act 1974	Division 2B: Offences for producing and distributing child pornography
Samoa	Crimes Act 2013	Part 18: Extensive offences for electronic systems – accessing without authorization, accessing for a dishonest purpose, illegal remaining, illegal interception, damaging or interfering with electronic data, illegal acquisition of electronic data, illegal system interference and illegal devices; identity fraud, forgery of electronic data, spam, solicitation of children and harassment utilizing means of electronic communication
	Telecommunications Act 2005	S74: Offences for telecommunications networks or computer systems - unauthorized access, interception, alteration of data, hindering of a network or system, sale of devices or data to facilitate above offences, use of a telecommunications network to offend or harass
	Copyright Act 1998	Criminal sanctions for copyright infringement
Solomon Islands	Telecommunications Act 2009	Part 19: Telecommunications offences – infringing security to obtain data, intercepting messages, altering/destroying/deleting data, revealing contents of messages, impeding or delaying messages and possessing a device to do any of the above
Tonga	Computer Crimes Act 2003	Ss4–8: Offences for Illegal access, interfering with data, interfering with computer systems, illegal interception of data, illegal devices
	Communications Act 2015, Pornography Control Act 2002, Copyright Act 2002	Other offences for: Identity theft, fraud and forgery, child pornography
Tuvalu	Tuvalu Telecommunications Corporation Act 2008 (Revised Edition)	S33: Telecommunications offences – modifying or interfering with messages, interception and disclosure of intercepted messages; sending of a grossly offensive message
Vanuatu	Penal Code (Consolidated Edition) 2006)	S73C(vii): Computer offences related to terrorism Ss147A&147B: child pornography offences
	Telecommunications Act (Consolidated Edition) 2006	Part 10: Telecommunications offences – intentional damage, interception and disclosure

Source: Le Nguyen and Golman 2021.

Table 1.4. Australian legislation on image-based sexual abuse

Jurisdiction	Legislation	Offences	Maximum Sentence
Federal	Federal Criminal Code Act 1995	s474.17: Using a carriage service to menace, harass or cause offence	3 years
Australian Capital Territory	Crimes Act 1900 (ACT)	s61B: Intimate observations or capturing visual data, etc. s72C: Non-consensual distribution of intimate images s72E: Threaten to capture or distribute intimate images s72H: Court may order rectification [for non-compliance with court order to take reasonable action to remove, retract, recover, delete or destroy an intimate image involved under s72C, s72D or s72E]	2 years 3 years 3 years 2 years
New South Wales	Crimes Act 1900 (NSW)	s91K: Filming a person engaged in private act s91L: Filming a person's private parts s91M: Installing device to facilitate observation or filming s91P: Record intimate image without consent s91Q: Distribute intimate image without consent s91R: Threaten to record or distribute intimate image s91S: Court may order rectification [for non-compliance with court order to take reasonable steps to remove, retract, recover, delete or destroy any intimate image involved under s91P or 91Q] s578C: Publishing indecent articles	2 years 2 years 2 years 3 years 3 years 3 years 2 years 12 months
Northern Territory	N/A	N/A	N/A

Queensland	Criminal Code 1899 (Qld)	s227A: Observations or recordings in breach of privacy s227B: Distributing prohibited visual recordings	2 years 2 years
South Australia	Summary Offences Act 1953 (SA)	s26B: Humiliating and degrading filming s26C: Distribution of invasive image s26D: Indecent filming s26DA: Threat to distribute invasive image or image obtained from indecent filming	2 years 1 year 2 years 1 year
Tasmania	Police Offences Act 1935 (Tas)	13A: Observation or recording in breach of privacy 13B: Publishing or distributing prohibited visual recording	12 months 12 months
Victoria	Summary Offences Act 1966 (Vic)	s41A: Observation of genital or anal region s41B: Visually capturing genital or anal region s41C: Distribution of image of genital or anal region s41DA: Distribution of intimate image s41DB: Threat to distribute intimate image	3 months 2 years 2 years 2 years 1 year
Western Australia	Restraining Orders Act 1997 (WA)	10G(2)(g): Restraints on \ from distributing or publishing, or threatening to distribute or publish intimate personal images of the family member	2 years

Source: Henry et al. 2019.

it an offence (s. 41A) to use a device to intentionally observe, intentionally visually capture or distribute an image of another person's genital or anal region.

In New South Wales, the Crimes Act 1900 (NSW) created three criminal offences: recording intimate images without consent (s91P), distributing intimate images without consent (s91Q:) and threatening to record or distribute intimate messages (s91R).

In the Australian Capital Territory, new criminal offences were created under the Crimes (Intimate Image Abuse) Amendment Act 2017, which concern non-consensual distribution of intimate images (s72C), distribution of intimate images of a person under the age of 16 (s72D) and capturing or distributing intimate images (s61B).

Given the vagaries of cybercrime legislation in terms of its application to the myriad types of online violence (noted earlier), it has long been maintained that a combination of legislative frameworks in lieu of, or as complement or alternative to, cybercrime legislation could be utilised. Thus, for example, information and communication legislation (concerning regulation of internet content and allied services), data protection legislation, human rights legislation (pertaining, inter alia, to relative rights of freedom of speech) and criminal legislation (concerning, inter alia, violence against persons, prostitution and cyber bullying) could be invoked in support of online violence victims (Broadband Commission for Digital Development 2015). The caveat to such applications, however, rests in large part on the lack of extra-territoriality of that kind of legislation in a world where online malfeasance necessitates legislation with a cross-jurisdictional ambit. This is important, not least because of the varying definitions of seemingly straightforward concepts such as 'cyberbullying'.

APC (2017) argued that 'governments tend to prioritise legislative solutions, but they take time and are frequently outpaced by technology and online gender-based violence practices. Adapting existing gender-based violence and cybercrime legislation or opening interpretation to encompass technology related gender-based violence, may be more practical than creating new legislation'.

Crucially, however the legislative response is framed and online violence defined, the net effect is that current laws focus on the initial perpetrator. This raises, for the victim, if not the legislature, a fundamental issue, which is that the harm caused by the act of online violence is exacerbated by the reactions and actions of members of the online community. Those members receive the manifestation of the online violence (whether through information, a comment or image) and then elect to tacitly support or condone that violence by dint of their subsequent action or non-action. A short-term solution to the involvement of the wider social media community would be to reconfigure some of the statutory offences noted above so that individuals other than the perpetrator could be brought into the offence category. For example, under the New South Wales Crimes Act 1900, the offence under s.91Q ('distributing an intimate image without consent') or s.578C ('publishing indecent articles') could theoretically be applied to anybody who received the original image and shared or otherwise distributed it among their peers.

However, to take that step would also involve the authorities in accurately tracing all such parties in an online environment where anonymity and geographical distance are key features. Moreover, the courts would also have to be satisfied in relation to third parties' intention to cause harm or their knowledge of whether or not consent was given, or sought, by the original perpetrator. Given the finite rules pertaining to criminal liability, this proposed approach would require a good deal of analysis and would doubtless be subject to a wide degree of criticism, notwithstanding its potentially positive benefits.

There has been an increased focus on endeavouring to understand why online bystanders seemingly distance or remove themselves from the plight of the online victims they encounter. Also, consideration has been given on how to persuade them to assist those victims by intervening and/or reporting the violence they witness or discern after the event. Moreover, driven in part by the apparent apathy or disengagement of online bystanders (who do not intervene or assist victims where it would be relatively easy to do so), the proposition that they be held legally accountable for their omission or failure to act has been promulgated. Each of these areas of debate requires a degree of analysis because what might appear at first glance to be a logical avenue for assuaging or reducing the impact of online violence is riven with actual, prospective or perceptual difficulties.

2. The role of bystanders in aggravating, perpetuating and mitigating online VAWG

The study of bystanders and their role in perpetuating online violence tends to focus mainly on 'cyberbullying'. This is largely due to the widespread use of, and reliance on, social media by the population demographic in which bullying commonly occurs, that is, young children and adolescents (see, for example, Pacer's National Bullying Prevention Center 2020; Tas'adi et al. 2020). However, cyberbullying constitutes but a fraction of the range of offences (noted above) that make up online violence.

In its physical manifestation, bullying is subject to a complex group dynamic comprising the 'bully', the 'victim' and 'bystanders'. Bystanders are important because bullying is in essence a public display of the bully's power, and that power derives from, and is increased by, the mere presence of and the attention given by the bystander to the bully's actions. In the cyberbullying context, the consequences for the victim are magnified due to a tendency for individuals to speak more harshly online than they would in a physical space. This is because the degree of accountability in the latter is lacking in the former. Where those users also have anonymity, both the degree and level of abuse they release are higher than would be the case in the physical sphere. As Shariff and Hoff (2011) note, '[i]n the absence of authority to set and enforce clear boundaries and structure, teens lose their inhibitions and engage in negative behaviour'. In a bizarre twist, it is also possible for those who would never become a bully in the physical world to become empowered to become one in the virtual world, possibly as a result of their own victimisation by bullies.

Cyberbullying more broadly may also involve a rise in scale and severity because the bully cannot see, as they would in the physical world, the impact on their victims. This may lead them to overcompensate by increasing the nature, scale and timeframe of the bullying episodes. Equally, cyberbullying facilitates participation by an infinite audience of bystanders and a never-ending onslaught against the victim. As Kowalski et al. (2012) note, 'the ubiquity of online social interaction prevents a victim from ever completely eluding the reach of the bully, unless [s/he] also chooses to cut himself off from his entire social network'.

The cyberbullying phenomenon is worsened by dint of the 'code of silence', in relation to reporting bullying to authority figures, to which victims and bystanders seem to adhere. Kowalski et al. (2012) note that '[b]ystanders are especially prone to inaction without the support of other observers to bolster their confidence to intervene on behalf of the victim [and] might also fear that reporting bullying to adults will make them the next target of the bully's ire'. Moreover, there is a disinclination for victims or bystanders to report cyberbullying for fear that their parents' reaction might simply be to remove them from social media. This would undermine a focal point of their lives since, as 'digital natives', they may '... communicate more through electronic devices than through face-time interaction with family and peers' (Coyl 2009).

In this context, bystanders are witnesses to acts of violence that they do not perpetrate, and they can play either a positive or negative role in relation to those acts. Thus, they can be deemed to be 'reinforcers', that is, people who engage in negative behaviour by clicking the 'like' button, leaving comments supportive of the perpetrator or degrading to the victim and/or forwarding the message or image, or similar, on the social media platform (Wong et al. 2021). Any, or all, of those actions serve to reinforce the behaviour of the perpetrator and extend and exacerbate the harm caused to the victim. Conversely, bystanders can be deemed to be 'upstanders', that is, people who elect to engage in a positive manner by reporting the violence to social media platform administrators or mediators or defending, consoling or overtly supporting the victim in the hope that either the

violence will stop or others will be persuaded not to share the post elsewhere.

A potential issue that can arise when considering the role of online bystanders is that the phenomenon has tended to be analysed in relation to cyberbullying (more from the point of school-aged children rather than employees) and not in relation to the myriad other types of online violence (Council of Europe 2018). In addition, in terms of utilising online bystanders as a tool in the fight against online violence, there is a prerequisite that the bystander has access to the online violent acts, but not all online violence is perpetrated in spaces inhabited by other online users. However, those broad points notwithstanding, it is arguable that the actions or inactions of bystanders in the cyberbullying context could be applied to other forms of online violence. This is because the same three parties will invariably be involved: the perpetrator, the victim and the bystander.

Valdés-Cuervo et al. (2021) define cyberbullying as a '...repetitive and wilful electronic communication to bully a person, typically by sending messages or posting information of an intimidating or threatening nature'. Those who witness that communication, the bystanders, may stimulate the cyberbullying (by adding equally violent commentary), reinforce the cyberbullying (by encouraging others to add their own commentary) or remain passive (by ostensibly ignoring it). Yet, even a degree of passivity can contribute to the ongoing nature of the online violence because a decision to remain neutral can convey silent assent to, or condoning of, the perpetrator's behaviour.

According to Valdes-Cuervo et al. (2021), defenders or upstanders may fall into one of two categories: 'constructive' or 'aggressive'. Constructive interveners focus on advising the perpetrator that their actions were wrong and/or by advising the victim that they were blameless and should not have been targeted. Aggressive interveners react vehemently but focus their attention on threatening the perpetrator.

It has been demonstrated that the degree to which, if at all, an online bystander will defend a victim can be predicted by her or his social group, attitudes and situational influences (Lytle et al. 2021). Perhaps unsurprisingly, given the demographics of the victims of online violence, intervention is more likely when the bystander is female. Furthermore, the propensity to intervene increases with the age of the bystander and if they have strong social support or popularity. Factors legislating against intervention include situations where the bystander holds negative views of the victim and where the nature of the bystander's social environment, including lessons learned in a family setting, renders them disinclined to intervene.

It is arguable that perpetrators of online bullying may engage in such behaviour in part because of the way in which they gain and maintain status in the online social media environment they inhabit (Moretti and Herkovits 2021). Thus, '... by making fun, they believe they can entertain their peers, build bonds, and obtain recognition that gives them social status' (ibid., p. 9). The role of bystanders in this context is to rate that status by indicating how much they 'like' the perpetrator's latest post. Where the victim lacks positive peer assessment generally (and that will often be a causal factor in their victimisation), this strategy of rating the perpetrator is rendered easier. In this way also, however, the harm caused to the victim is exacerbated for, as Moretti and Herkovits note, '... in a space where image represents a medium for experiencing one's identity and values, practices aimed at disseminating photos and videos are perceived as more harmful than written or physical attacks' (ibid., p. 9). By extension, even a passive observer of that material contributes, wittingly or otherwise, to the exaggeration of the harm caused by the original bullying incident.

One might imagine that the proclivities of adolescent users in the online space, in terms of engaging in online violence largely to gain social media popularity and/or becoming swept up in the frenzy that often surrounds posts of difficult or dangerous activity, would dissipate within the adult online population. Moreover, aside from perhaps not engaging with or directly or indirectly supporting the perpetrator's victimisation of the victim, one might posit that there would be a preponderance of bystanders who would take more positive action. This might include reporting the violence to the social media platform administrators or to the authorities, or indeed engaging with the perpetrator in an overtly critical manner and/or with the victim in a supportive manner. Unfortunately, this supposition may be premature given the existence and perseverance of the 'bystander effect'.

3. The bystander effect

The bystander effect provides that the greater the number of passive individuals who witness an emergency, the less likely it is that any one of them will assist the victim (Darley and Latane 1968). The concept stemmed from an event in New York in 1964 when 38 people witnessed the attack, and ultimately murder, of Kitty Genovese. This led to a desire to explain the behaviour of the bystanders (which had been characterised variously as 'indifferent', 'morally callous' and 'dehumanising'), not least of all in terms of why none of them acted in response to such a violent event.

Rather than focus on facets of the individuals' personalities – in terms, for example, of apathy or indifference – Darley and Latane (1968) instead honed in on situational factors in terms of the relationship among the bystanders. They determined that the presence of a crowd of bystanders could inhibit an individual's prosocial impulse to intervene in an emergency. The impact of being seen by others, and the effect of seeing others, was termed the 'bystander effect'. Under this effect, factors such as 'diffusion of responsibility' (whereby the responsibility to intervene is determined by the number of bystanders present), 'evaluation apprehension' (a fear and/or embarrassment of being judged in a negative light by the other bystanders) and 'pluralistic ignorance' (the belief that inaction by other bystanders means that no action is required) were deemed to hinder a series of decisions the bystanders deemed necessary to make before electing to intervene. These decisions involved:

1. Noticing that something was happening
2. Interpreting that the event being viewed was an emergency
3. Taking personal responsibility for acting
4. Determining how to act with the belief that one had the skills to succeed
5. Implementing the action chosen

In the online rather than physical context, some basic points of distinction can be seen. First, the number of potential witnesses to an online incident may be dramatically higher than those witnessing a physical event. Second, fellow online witnesses may have little real sense of the actual numbers of their contemporaries. Third, all online witnesses will have a greater sense of disassociation from the victim, and the point in time at which witnesses observe the online incident will vary, with some watching an event live and others via video and/or images posted after the event.

Other points of distinction include the fact that, in the online space, any non-explicit response to an event by bystanders will not be readily identifiable to other bystanders (Domínguez-Hernández et al. 2018). This will render the diffusion of responsibility more likely to apply on the assumption that there will be many other parties witnessing the event on social media. An added complication, which tends not to apply at the scene of a physical accident, is that bystanders' ability to discern the existence, nature and extent of an online event is hampered, including whether or not action has already been taken in relation to it.

Studies have been conducted into the prevalence of the bystander effect in the online environment. Markey (2000) created repeated requests for assistance in some pre-existing internet chatrooms and discovered that individuals were slower to respond when other bystanders were present, although making the request of one named individual produced the fastest response, irrespective of how many others were present. Barron and Yechiam (2002) noted that, in their study, email requests for assistance sent to one recipient were more likely to be responded to than if they were sent to five recipients. Similarly, Blair et al. (2005) discovered that the probability of receiving a response to an email request for assistance declined in direct proportion to an increase in the number of recipients.

In relation to cyberbullying, where the bystander effect has been studied more prominently than in relation to other online incidents, Obermaier et al. (2014) utilised Facebook to present participants in their study with a screenshot depicting a post made on the 'wall' of a university Facebook group. The post was a request made by a fictitious student for lecture notes, to which another 'member' of the group added insults and an invitation to others to post similar derogatory comments. Different

bystanders were informed that the post had been seen by 2, 24, 224 or 5,025 members of the group, respectively. The study found that the number of bystanders did not directly affect the willingness of the participants to intervene, but that there was an indirect impact on that willingness when there were fewer bystanders present, which led to a need to demonstrate some personal responsibility.

3.1 Mitigating the online bystander effect

Perceptually, responding to online violence as an online bystander would be a relatively simple and straightforward affair since, if nothing else were to occur, a simple telephone call to the authorities would reduce the occurrence and/or its impact dramatically. In practise, however, this is not the case. Kaufman (2021) has noted a series of incidents where online bystanders did nothing to assist the victims of violence. In one case in Steubenville, United States, two 16-year-old boys raped a girl of the same age while the culprits and others discussed the assault, shared photos and videos of it and ridiculed the girl through text messages and through posts on YouTube, Twitter and Instagram. Nobody who witnessed the rape physically or virtually called the police.

On the other hand, there have been incidents witnessed by many that have seen an intervention by one or more individuals and demonstrate, in principle, the capacity for online bystander intervention. For example, in relation to the Steubenville case, Alexandria Goddard, a former resident of the town where the crime was perpetrated and who, at the time, was in another state, took screenshots of messages, photographs and videos on social media and sent them to the police before they could be deleted. In 2017, in Sweden where the gang rape of a woman was livestreamed via a closed Facebook page, a number of online witnesses notified the police, who managed to intervene, stop the assault and arrest the perpetrators (BBC, 2017).

However, there have also been cases where there has been no online bystander intervention and the issue had to be resolved by relatives of the victim. Thus, in 2020 in Providence, Rhode Island, a mother discovered a video on Facebook of her unconscious daughter being sexually assaulted by some men. She sent a copy of the video to the police.

Online users measure their worth in terms of 'friends' and 'followers', 'likes' and 'comments', with the number of each being determined proportionately to the level of activity one demonstrates in the online space. When a user posts or shares an image or video, the response they receive from those 'friends' or 'followers' in the form of external approval '... triggers a powerful psychological response ...' (Love 2018). It is a tragic tacit corollary that shocking footage invariably elicits the greatest response. By extension, this creates a powerful inducement on the part of the poster to put up yet more disturbing content. At a common-sense level of course, it is certainly arguable that a reduction in original posted content would minimise in turn the reposting by online bystanders.

Discussion on potential responses to the wilful apathy, or lack of empathy, on the part of online bystanders who omit to act in regard to the online violence they have witnessed has proceeded down particular routes:

1. Enabling the online bystander to be able to report directly to the social media platform on which they witnessed the illegal activity
2. Usurping the moral duty of the online bystander by requiring social media platforms to act more aggressively or comprehensively – this could involve identifying illicit content, reporting the perpetrator to the authorities and/or removing their ability to engage on that platform as well as identifying those who demonstrate their presence through their positive affirmations and/or negative commentary regarding the online violence, and either report or remove them along with the perpetrators
3. Raising the ethical expectation society has for online bystanders (as it has progressively in relation to physical bystanders) by creating and enforcing a legal duty to respond to the online violence they witness, with punishment imposed for failing to do so

3.2 The role for and expectations of platform providers

It seems logical to bring internet intermediaries into the debate regarding the activity and inactivity of online bystanders. The internet is a relatively freely accessible forum, with access accommodated by

a number of intermediaries that provide hosting facilities and control the sites that house the social media platforms (such as Facebook and Twitter) and video-sharing applications (such as TikTok) that facilitate the exchange of information, including that related to online VAWG. The simplest route to encouraging the intermediaries to engage in this debate lies in the fact that the online violence noted in this report is transmitted via the vast web platforms that they own and which they operate extraterritorially.

UN Women (2020) has argued that '... intermediaries have a responsibility to balance their business imperative to encourage traffic on to their platforms while protecting freedom of speech, and removing violent, inappropriate and harmful content'. Wong et al. (2021) note that online platforms are the places where '... people form communities in which they create, exchange, comment, recreate and cocreate content'.

In that sense, it has been suggested that users should be responding to '... the shaping of moral obligations that are set by platforms' (Haber 2020). As Haber contends, '[s]uch self-regulation could aid in reframing the ways in which human rights and liberties should be protected in this era, which might better accommodate the technological and social changes ...' (ibid., p. 1627). It has been suggested that this transition might be achieved were platforms to provide, for example, a way in which users could state a trigger phrase. This would automatically lead to law enforcement or to a trusted person being contacted, who would be able to see all relevant data, view the event in question as it unfolded and react accordingly. In this vein, some platforms have already begun to respond to livestreaming activity on their sites. Facebook, for example, does not allow users who proclaim a 'violent mission' or are 'engaged in violence' to have an account, which would include, not unnaturally, those involved in organized violence or criminal activity.

However, the requirement for a proclamation may allow much violent material to be posted if an overt statement of violent intent by the user is not present. Facebook has also introduced a 'one strike' rule in response to the livestreaming of crimes, under which those who do so are restricted (but not permanently banned) from using the live feature for a finite period. However, a mere restriction arguably undermines the seriousness of posting criminal activity. This is because the obvious reaction to such livestreaming (if Facebook were cognisant of the impact of the material's subsequent dissemination by bystanders) would be to remove indefinitely the user's opportunity to stream (and, by logical extension, the ability of bystanders to access such sites). Moreover, Haber (2020) has noted that '... the most dominant factor in self-regulation is the economic incentive of platforms to encourage users' engagement in the platform'. On that logic, it would take a brave platform to impinge on what is clearly for many users a key draw for its use, namely hosting and sharing content (with lurid and criminal content retaining a high degree of popularity). As Haber has put it, '[p]latforms' potential role as digital Samaritans ... will be influenced not just by their users' attitudes but also by their own perspective of whether their actions constitute an optimal form of business and management' (ibid., p. 1633).

Many platforms now provide reporting mechanisms within their virtual environment. Haber (2020) has argued that such mechanisms should remove, or at least mitigate, the reluctance of online bystanders to overtly intervene in terms of reporting an online offender. This is particularly so if that offender is known to them or to the wider social networking group to which they belong.

Thus, a concerned online bystander can report anonymously to the provider, but that willingness is affected by four determining factors:

1. The bystanders' assessment of the emergency of the social media online violence
2. The bystanders' sense that it is their personal responsibility to report the incident
3. The bystanders' belief that they have the capacity to intervene
4. The bystanders' sense that others are present and that they too will join in utilising the self-reporting mechanism.

Wong et al. (2021) have attempted to capture the degree to which online bystanders are likely to intervene via reporting mechanisms (see Table 3.1).

Table 3.1. Bystanders' willingness to use built-in reporting functions of social media platforms

Decision of intervention	Bystanders' willingness to report social media harassment incidents to platform owners by using built-in reporting functions
Assessment of the event	— Perceived emergency – the extent to which bystanders believe that the social media harassment incident needs to be addressed urgently
Assessment of personal responsibility	— Perceived responsibility to report – bystanders' subjective assessment of their sense of personal obligation to deal with social media harassment incidents
Assessment of capability to intervene (personal and situational factors)	— Perceived self-efficacy to report – bystanders' subjective assessment of their ability to successfully report the harassment using built-in reporting functions — Perceived outcome effectiveness of reporting – extent to which bystanders believe that using built-in reporting functions will effectively tackle social media harassment
Presence of others	— Pluralistic ignorance – defined as the extent to which bystanders believe that other bystanders who have also witnessed the incident will remain unconcerned with the social media harassment incident — Diffusion of responsibility – extent to which bystanders believe that reporting responsibility should be transferred to other bystanders who have also witnessed the incident — Evaluation apprehension – bystanders' fear of being judged or negatively evaluated when using built-in reporting functions to report social media harassment incidents

Source: Adapted from Wong et al. 2021.

4. Good Samaritan legislation

In most jurisdictions, there is no legal duty on ordinary bystanders to rescue a person in danger. As Ames (1908) once observed '[t]he law does not compel active benevolence between [person and person]. It is left to one's conscience whether he shall be the good Samaritan or not'. However, a number of those same jurisdictions have endeavoured to persuade members of society to intervene through the creation of so-called 'good Samaritan' legislation. The key driver of such legislation has been to provide a statutory limitation for liability in relation to the good faith efforts of any prospective rescuer.

In Australia (but not in New Zealand or the Pacific Islands), for example, under New South Wales legislation (Civil Liability Act 2002 (NSW), s56), a 'good Samaritan' is defined as someone who 'in good faith and without expectation of payment or other reward, comes to the assistance of a person who is apparently injured or at risk of being injured'. An indirect aim of such legislation is to encourage bystanders to intervene when another person is in obvious need. Although the legislation was created in relation to physical emergencies, there is no real difficulty in applying the principle of it to the online context. This is because virtual witnesses may be deemed to have a similar degree of proximity to a physical witness (in terms of an ability to witness an event). Accordingly, apart from an inability to physically intervene in a rescue of, for example, a sexual assault victim, such witnesses can nevertheless have a positive impact through their intervention in alerting the emergency services.

In Australia, each state and territory provides statutory protection from liability for people who assist strangers in need to encourage people to provide assistance (see Table 4.1). However, as Al-Alosi et al. (2016) argued that, '... good Samaritan laws currently do not impose a duty to intervene and, combined with the natural disengagement characteristic of the online environment, may require a more substantial form of motivation for bystanders to intervene ....'

Overall, Benzmiller (2013) has posited that a cyber good Samaritan has a duty to report when either one of two situations presents itself: (1) the online violence includes threats of violent criminal behaviour; and (2) the witness knows, or reasonably believes, that the online violence will cause physical harm or the fear of physical harm.

Table 4.1. Australian legal protections for good Samaritans

	Legislation	Protection	Exclusion from protection
Australian Capital Territory	Civil Law (Wrongs) Act 2002	Honestly and without recklessness in assisting, or giving advice about, an injured person, or person in need of emergency medical assistance.	✓ Liability falls within ambit of a scheme of compulsory third-party motor vehicle insurance ✓ Capacity to exercise appropriate care and skill was significantly impaired by a recreational drug
New South Wales	Civil Liability Act 2002	In good faith in an emergency when assisting a person who is apparently injured or at risk of being injured.	✓ If the good Samaritan's intentional or negligent act or omission caused the injury or risk of injury ✓ Ability to exercise reasonable care and skill was significantly impaired by being under the influence of alcohol or a drug voluntarily consumed (whether or not it was consumed for medication); and failed to exercise reasonable care and skill
Northern Territory	Personal Injuries (Liabilities and Damages) Act 2003	In good faith and without recklessness while giving emergency assistance/advice about the treatment of a person being given emergency medical assistance.	✓ Intoxicated while giving the assistance or advice
Queensland	Law Reform Act 1995	In good faith and without gross negligence in rendering medical care, aid or assistance to an injured person at the scene of emergency or during transfer to hospital	
South Australia	Civil Liability Act 1936	In good faith and without recklessness in assisting a person in need of emergency assistance/giving advice about assistance, e.g., by telephone.	✓ Liability falls within ambit of a scheme of compulsory third party motor vehicle insurance ✓ Capacity to exercise due care and skill was significantly impaired by alcohol or another recreational drug
Tasmania	Civil Liability Act 2002	In good faith and without recklessness in providing assistance, advice or care at the scene of emergency/providing advice, e.g., by telephone to a person at scene of emergency	✓ Ability of the good Samaritan to exercise reasonable care and skill was significantly impaired by reason of the good Samaritan being under the influence of alcohol or a drug voluntarily consumed (whether or not it was consumed for medication); and ✓ The good Samaritan failed to exercise reasonable care and skill in connection with the act or omission.
Victoria	Wrongs Act 1958	In good faith in providing assistance, advice or care at the scene/providing advice, e.g., by telephone to a person at scene of emergency or accident.	
Western Australia	Civil Liability Act 2002	In good faith and without recklessness in assisting person in need of emergency assistance/ providing advice about the assistance to be given.	✓ Ability to exercise reasonable care and skill was significantly impaired by being intoxicated by alcohol or a drug or other substance and intoxication was self-induced

5. Bad Samaritan legislation

As the contrasting name suggests, 'bad Samaritan' legislation seeks to penalise those who witness and have an opportunity to intervene, directly or indirectly, in support of those in need of rescue, or who are otherwise in an apparent state of distress, but who fail to do so.

This presupposes that all witnesses to violence may have the propensity and/or opportunity to intervene. However, it is important to note that not all online violence against women and girls occurs in a manner capable of being observed by people other than those specifically invited to witness that violence and who, by dint of that invitation, will not intervene against, nor report, the incident. Thus, in cases of online sexual exploitation of children, the witnesses to the abuse are bystanders in the common sense notion of the word, that is, they are watching without intervening, but the nature of their invitation to witness the abuse will result in them never reporting it to the authorities. Consequently, the role of the online bystander, in terms of reporting offences in this context is redundant. In broad terms, therefore, the range of activity that might be contained within bystander legislation will not, even if effective, reduce violence against women and girls in situations such as these.

Benzmiller (2013) has posited that a cyber-Samaritan has a duty to report when either one of two situations presents itself: first, the online violence includes threats of violent criminal behaviour; and second, the witness knows, or reasonably believes, that the online violence will cause physical harm, or the fear of physical harm. However, the enforcement of such a statute would face a number of problems. For example, there might be several potential offenders and difficulty interpreting any given situation for potential threats. Moreover, it would be problematic to determine culpability among a group of witnesses and perceptually unfair to punish the entire group when some among them might have just cause for not reporting.

As is common in many jurisdictions, there are no legislative provisions in the Pacific region that overtly make it an offence to fail to intervene as an online bystander, and that broadly reflects the norm present in the physical world. Thus, for example, it is not an offence to omit to act in response to a person drowning in a river (unless a recognised duty exists, such as that of a parent in relation to a child). The only potential exception to this situation lies in the Northern Territory in Australia, albeit in the physical space only, where the Criminal Code Act 1983 provides that '... [a]ny person who, being able to provide rescue, resuscitation, medical treatment, first aid or succour of any kind to a person urgently in need of it and whose life may be endangered if it is not provided, callously fails to do so is guilty of an offence ...' (s.155).

The term 'innocent', which is often applied to those who observe but do not participate in online offending, is something of a misnomer in a virtual context and care should be taken before applying it. The other issue is one of scale and scope. Technically, anyone who views, or witnesses, posts (whether in textual, pictorial or video format) online is a bystander. The very fact of clicking on a post to view it increases its currency and impact on the victim. Without that action, there would be no bystanders and the impact of the post would be mitigated. It is possible to call someone who 'simply' opens a post an innocent bystander, unless the nature of its content is apparent before it is opened from, for example, a heading. If, however, that bystander then elects to share or respond (whether in the form of words or emoticons or similar) they arguably become 'non-innocent' bystanders.

Whether innocent or otherwise, it is then necessary to configure whether such people should be required to assist the victim. Furthermore, it should be determined whether that requirement should be routed through education and information programmes pertaining to good citizenship or through legislative provisions. The former are in abundance largely in the context of cyberbullying but, while an important issue, this is but one of many variants of online violence perpetrated against women and girls. The legal route has many actual and/or perceived concerns and difficulties. More promising and operationally more efficient are pieces of legislation that overtly capture those who assist in the impact of online violence, or at least are open to the interpretation/argument that they do so.

In 1883, Sir James Fitzjames Stephen (1883) noted that '[a] number of people who stand around a shallow pool in which a child is drowning and let it drown without taking the trouble to ascertain the depth of the water, are no doubt shameful cowards, but they can hardly be said to have killed the child'. In a modern reflection of that sentiment, in 2017, five teenagers video recorded Jamel Dunn drowning in a pond but did nothing to save him. Although their behaviour was construed as wholly immoral, the Florida Penal Code contained no provision that could render them culpable for Dunn's death.

As Ames (1908) once opined, '[w]e should all be better satisfied if the man who refuses to throw a rope to a drowning man or to save a helpless child on the railroad track could be punished'. Some commentators have indeed argued for the criminalisation of a failure to act, with Ashworth (1989), for example, observing that '[t]he general principle in criminal law should be that omissions liability should be possible if a duty is established, because in those circumstances there is no fundamental moral distinction between failing to perform an act with foreseen bad consequences and performing an act with identical bad consequences'.

It has indeed been argued that the failure to act constitutes 'misfeasance' not 'nonfeasance' (Grande Montana 2018). As Grande Montana observed,

> 'A bystander who witnesses a crime upon a victim has the power to affect the situation and the crime in progress by notifying the authorities or directly assisting the victim. If the bystander does nothing, the resulting harm to the victim is a consequence of the bystander's decision not to use this power to intervene. Not only is the harm that results to the victim connected to the bystander's failure to intervene, but the bystander's failure to intervene is also a "causally relevant factor" in the resulting harm to the victim. Failing to summon the authorities on behalf of the victim is not the sole cause for the resulting harm ... [h]owever, [it] plays a relevant role in the harm that results' (ibid., p. 533).

As Haber (2020) has suggested, '[t]he developments of digital technology and the markets that drive them have led to a potential new form of digital bystanders'. In recent years, social media platforms have facilitated the ability to livestream on the internet, which has segued into a proclivity of online users to broadcast and/or share criminal activities, not least sexual violence. In that context, support for the notion that an affirmative decision not to intervene constitutes misfeasance in the online space is provided by the fact that spectators who film and then disseminate video footage, images or messages might be deemed to be 'engaged spectators'.

As Andersson and Sundin (2021) have observed, for example, 'Swedish news media reported how witnesses at accident sites used their smartphones to film the situation, instead of offering help ... The reports on these "mobile bystanders" were soon followed by a debate. While this debate included various positions and perspectives, there was one perspective shared by most commentators: that the behaviour was morally deplorable'.

5.1 Purpose of and rationale for bad Samaritan legislation

Benzmiller (2013) maintains that the primary motivation for bad Samaritan legislation is to punish the bystander who fails to help and to encourage future bystanders to act on behalf of a person in danger. In most jurisdictions, there is an obligation imposed on parties for whom a duty of care is deemed to exist (parent of a child, lifeguard at a beach, etc.), but bad Samaritan legislation creates a duty between parties for whom no formal connection exists. It has been suggested that good Samaritan legislation protects the bystander who endeavours to assist a victim, but it does not drive those who do not assist to do so (Guiora and Dyer 2020). In terms of the suggestion often made to rely on individuals' moral compass, or on a moral foundation that draws on humanity, respect and decency in society more broadly rather than to legislate, Guiora and Dyer have contended that '[w]hile, doubtlessly, there are numerous examples of individuals acting on behalf of another in distress, to suggest that is a cultural norm reflecting consistent normative values and behavior is inaccurate' (ibid., p. 297).

Love (2018) has argued that '[e]very day, people post footage of other people dying or being assaulted to online platforms for public consumption and entertainment. While it is morally offensive, in the digital age, choosing to document

and repost instead of actually preventing a tragedy is shockingly commonplace'.

In this vein, social media has redefined the notion of 'bystander', from the traditional construction of one who is physically present at an event or incident but does not participate to one who actively records and posts events that appear on their devices onto their digital platforms. It is this changing role of the online bystander from a passive observer to one who participates through the act of distribution of material that has led to calls for criminal liability.

Haber (2020) notes that there are different justifications for bad Samaritan legislation. These include some utilitarian arguments, including that assisting will minimise needless death and injury, create an outlet for moral outrage and reinforce the notion of social solidarity that underpins democratic societies, as well as provide a potential deterrent effect for non-compliance. He posits that '[t]he underlying assumption is that criminal law's deterrence and stigma will regulate the behavior of those who otherwise would not have aided and will potentially achieve other goals of criminal law like retribution for the wrongdoer' (ibid., p. 1572).

Support for the rationale for seeking to punish online bystander has been drawn from the intrinsic purpose of criminal law which, at its heart, is driven by the prevention or punishment of harmful acts. There are arguably two acts of harm in failing to intervene on behalf of, and then posting images in relation to, a person in distress: (1) the failure to provide emergency assistance, when in a position to reasonably do so; and (2) the dissemination of the footage of distressing events on the internet for public consumption. As Love has put it, '[w]atching a sexual assault or a man drown and then laughing at the person in need of emergency aid is morally reprehensible by any standard' (2018, p. 9).

In terms of the justification for addressing this harm in relation to the failure of online bystanders to intervene and/or their action in disseminating the evidence of harm there are two broad aspects in evidence. First, as Love observes, 'social media incentivizes people to record an emergency event instead of intervene in it' and 'social media creates a new context for the victimization of those individuals who are captured on bystanders' recordings' (2018, p. 9). In that sense, there is an aggravation of the harm originally caused to the victim that would not exist but for the existence and ubiquity of social media outlets.

Second and conversely, Haber observes that some commentators have opined that '… forced altruism is inherently wrong – even when reporting or rescuing is rather easily performed and even if rendering aid will not place the Samaritan in danger or peril – thus they have a right to refrain from reporting or rescuing' (2020, p. 9). Moreover, he has maintained that, while it might indeed be important to find ways to regulate human conduct to the point that members of society will become more willing to assist others in need, '… it is not the task of the criminal law to do so' (ibid.).

Thus a central argument against the creation of criminal liability for a failure to act is the oft-cited notion that the role of the criminal law has always been to punish action rather than inaction. As Love has suggested, '… opposition to punishing the failure to act with the criminal law starts with the proposition that society's greater concern should be punishing wrongdoing as opposed to bringing its lethargic or ignorant members up to scratch' (2018, p. 13). Indeed, it might be argued that there is a moral superiority of those who fail to act on behalf of the victim over those who perpetrate the crime.

On that logic, it is arguably unfair to expect observers to act under a duty of beneficence. To do so would require a determination as to the point at which the duty to assist others would begin and end. This very indeterminacy necessitates that legislation would need to carefully define the ambit of responsibility and provide exemptions for those who, for demonstrable reasons, could not or did not respond.

There is a further distinction to be made in terms of, for example, sharing a video or image as opposed to witnessing an attack, where it is conceivable that the nature of that attack is not entirely clear to the observer. In addition, it has been argued that imposing a requirement to act will inevitably dilute individual accountability for harm. This is because one person may assume that other observers will take the necessary steps to intervene, which thereby removes the necessity for him/her to do so. Allied to this notion is the issue of 'autonomy', whereby it could be argued that imposing a responsibility on an observer to act impacts on their liberty more than does the imposition of a sentence on the person who commits the crime observed. Support for this notion stems in part from an argument holding that a failure to act in relation to an event does not, in of itself, stop

the act from happening. Therefore, one should not attach liability to a person who simply fails to intervene. This of course cannot be held to apply during an ongoing assault since an intervention might very easily reduce or stop the attack. Similarly, even if the observer is viewing the aftermath of an attack, posting images or video of the attack impacts directly on the consequences for the victim of that act.

More importantly, perhaps, is that there is arguably a considerable difference between the perceived or required responsibilities of a bystander in the physical world compared to the online space. In the former, bystanders will invariably be observing an accident or an assault, and their willingness or ability to assist may be compromised, not least of all because of the bystander effect. In the latter, all that is effectively required of a bystander is to notify the police or similar authority, including, where facilitated, the platform provider. Ideally, it would be hoped, if not expected, that the bystander would in addition refuse to disseminate the content, post messages in support of the victim and/or challenge the poster of the online violence material and urge other bystanders to do likewise. In the virtual situation it is arguably less viable for a bystander to cite the fear or apprehension that might occur in the physical environment. In that sense, bad Samaritan legislation might be deemed to be acting against unjustified apathy as much as against an omission to act when to do so would place the bystander in little or no danger of reprisal.

However, Lord Devlin noted that 'the morals which underly the law must be derived from the sense of right and wrong which resides in the community as a whole; it does not matter whence the community of thought comes, whether from one body of doctrine or another or from the knowledge of good and evil which no [person] is without' (1959, p. 149).

Beyond the moral imperative, there are also practical aspects of the rationale for bad Samaritan legislation in terms of, for example, the prospective evidence gathered by users of social media while witnessing and/or disseminating the content created during the commission of a crime.

Ultimately, as Guiora and Dyer (2020) have summarised the position, '[b]ystander legislation bridges the gap between instances where individuals act in concert together to commit a crime and those where an individual simply chooses not to summon assistance for someone they know is suffering a serious bodily injury. It is the difference between active criminal intent and callous indifference for the life of another human being'.

5.2 Challenges of bad Samaritan legislation

It has been suggested that bystander liability does not materially affect behaviour (Hoffman in Guiora and Dyer 2020, p. 301). Lord Hoffman observed that

> 'One can put the matter in political, moral or economic terms. In political terms it is less of an invasion of an individual's freedom for the law to require him to consider the safety of others in his actions than to impose upon him a duty to rescue or protect. A moral version of this point may be called the "Why pick on me?" argument' (Stovin v. Wise [1996] 3 All ER 801 (HL) 819).'

The irony of that assertion, when applied to bystanders who refuse to engage in the mitigation of crimes they witness online, is that in other contexts, social media users have engaged in a number of popular movements. These include #MeToo (in connection with sexual harassment – ironically a key constituent of the online violence discussed earlier) and #BlackLivesMatter (in connection with racially prejudicial killings). The former hashtag was used more than 19 million times on Twitter between October 2017 and September 2018 (Anderson and Toor 2018).

Guiora and Dyer (2020) have suggested that fear of prosecution does not tend to compel those who are subject to a legislative requirement to act if they would otherwise feel no compunction to do so. The disconnect between people's belief that they should not have to intervene in an emergency, and legislation requiring that they do so, is often too pressing.

This disconnect is driven in part by a phenomenon of online activity where, for example, '... [t]echnology may ... have created a desensitization to violence and a decrease in empathy, which make it more likely for bystanders to pull out their phones to film an incident than dial 911' (Guiora and Dyer 2020, p. 299).

Moreover, as Lytle et al. observed, '[c]ybervictimization, along with the online social environment they occur within, is relatively new

[and] the norms and values for observing and responding to status updates in social media continue to evolve, which may create uncertainty for witnesses regarding what is and is not acceptable behavior online' (2021, p. 714).

More broadly, there arguably remain technical, ethical and/or logistical issues in relation to the application and operation of bad Samaritan legislation.

In terms of the application of such legislation, several factors need to be considered.

First, at what point should a duty arise for the online bystander? In the physical environment, this has usually been configured in terms of a threat to life or of serious injury. However, in the online environment, that level of danger might only arise in, for example, livestreaming of a sexual assault. In relation to the plethora of offences that fall under the 'online violence' category, it may be more difficult for a bystander to know when to intervene.

Second, who should be legally bound to intervene from the infinite number of other bystanders known or perceived by the bystander to be present? Moreover, how would a determination be made by the bystander that another had, or had not, already intervened?

Third, what should a bystander be obliged to do under any statutory duty to respond? Would a condemnatory comment posted online suffice, or the raising of a report to the intermediary hosting the platform or alerting the authorities?

Fourth, what in practice would be the criminal features of an offence of failing to respond? In the physical environment in relation to an attack or accident, it is relatively easy to configure what would constitute an omission. Thus, for example, in the Northern Territory Criminal Code (Australia), the offence (under s.155) is made out if the bystander fails '...to provide rescue, resuscitation, medical treatment, first aid or succour ...'.

In terms of the practical operation of bad Samaritan legislation, a number of other factors pertain.

First, there may be issues regarding the transmission and recording of the digital evidence with which online bystanders are concerned in terms of the applicability of territorial legislation in an extra-territorial environment.

Second, there may be hundreds of potential bystanders liable for prosecution in relation to countless online events. It seems probable that only a smattering of bystanders and/or a small collection of events will be considered for prosecution. That raises a double issue of (a) the majority of bystanders escaping liability and (b) the creation of an anti-deterrent atmosphere online where bystanders calculate that the likelihood of them being apprehended or prosecuted is, for them, acceptably low.

Third, there is a very real possibility that the sheer number of potential bystanders who might fall within the ambit of bad Samaritan legislation may negatively impact the already finite resources of police and prosecutorial offices to investigate, gather evidence and then proceed against those bystanders. This is particularly likely because the bystanders may be located in a number of jurisdictions where mutual legal assistance and cooperation between law enforcement and prosecutorial offices may be variable.

Fourth, even if the necessary resources are available to pursue an active investigation, it will be difficult in practical terms to locate and substantiate evidence of bystanders' involvement, given the online framework in which the failure to act occurs. Moreover, bad Samaritan legislation will require the same level of proof as would be expected in a physical environment, and there would be a very real difficulty in proving that bystanders knew the scenes they were witnessing were real and accurate. It would certainly be arguable, from the perspective of the bystander, that digital alteration of audio and visual data gave them the impression that the footage they witnessed was fake, or that they were too unsure to justify taking action in relation to it. Prosecutors would also have to show causation in relation to any assertion of inaction by a bystander. That is, they would have to establish that it was the bystander's failure to act that caused the injury to the victim and/or that the bystander was aware of the danger to the victim. It is arguable, from the bystanders' perspective, that they are in fact inconsequential given that the online violence would have occurred whether they witnessed it or not and whether they reported it or not.

Conversely, the very reason for a drive towards bad Samaritan legislation lies in the notion that it is the presence of bystanders that drives the online violence and/or that those bystanders exacerbate the impact of the original offence by either failing to report it or disseminating the evidence of the offence and/or commentary about it across various

social media platforms. As Love has observed, '...individuals who are unlucky enough to be not only unassisted, but recorded, by an ineffectual bystander are forced to re-live their trauma as it spreads, unchecked, on online channels' (2018, p. 11).

Fifth, the very nature of the online world obviates against the successful adherence to, and therefore enforcement of, bad Samaritan legislation. Individuals who post footage of online violence do so with the very real expectation that there will be an active audience viewing and reacting to it. Until the relative thrill experienced by online users who receive positive affirmation of their social media posts dissipates, any prospective bad Samaritan legislation will, in terms of its impact, be somewhat hampered.

5.3 Bad Samaritan legislation: Additional considerations

Given the unique environment in which online bystanders reside and the potentially endless stream of bystanders who might fall within the purview of bad Samaritan legislation, thought also needs to be given to a categorisation of respective liability.

One potential avenue for this lies in distinguishing between different types of bystanders and their respective motivations and actions/inactions.

A broad distinction can firstly be drawn between 'transmitters' and 'receivers'. Transmitters are witnesses who are physically present at a scene and share evidence of it electronically through various online platforms. It is possible to differentiate further within this category: (1) 'upstander' transmitters who intervene by, for example, attempting to assist the person in distress directly or by notifying the authorities; (2) 'enablers' who intervene harmfully by, for example, facilitating or exacerbating the crime; and (3) 'bystanders' who do not intervene at all. It is also possible to distinguish between 'contemporaneous' transmitters, who share evidence of a crime as it is occurring, and 'delayed' transmitters, who share evidence of a crime after it has occurred. Across those types and sub-types it may also be possible to discern motives for bystanders' conduct, which in turn may assist in determining a culpability ranking. Thus, transmitters may variously be acting in a benevolent, malevolent or neutral fashion or be either endeavouring to prevent a crime or to preserve evidence of it for the authorities.

Equally, transmitters may act in pursuance of the attention and positive affirmation that have become symbolic of the social networking environment and shame or humiliate the victim in that pursuit.

Receivers tend to be remote witnesses who observe live or recorded material from transmitters. From their physically disconnected position, receivers may engage in a range of actions or omissions in relation to the online violence. This might include (1) passive observance, (2) active engagement (through the provision of positive or negative feedback) with the transmitter and/or perpetrator via social media platforms, (3) further dissemination of the provided content across the same platforms or (4) reporting the online violence to the platform administrators or authorities.

6. Non-legislative responses to online VAWG

It is a logical argument to suggest that strategies and policy drivers aimed at preventing or mitigating physical violence against women and children, and/or at supporting the victims of it could simply be applied within the online space. However, the first issue is that, save for instances where online violence or the threat of violence manifests itself subsequently in the physical environment, most online violence does not result in physical harm to the victim. Equally, though, if the broader interpretation of 'online violence' is adopted and utilised, it would be possible to apply policy responses that are used in terms of, for example, 'face-to-face' harassment to similar incidents in the online space. However, the consequence of this policy application would, as in the physical environment, focus primarily on the victims and/or the perpetrators of the offences. Unless those policy determinations also seek to deal with online bystanders, it is likely that, in terms of response, the online environment may simply mirror, yet not improve on, the physical environment. Moreover, engaging perpetrators and bystanders, through policy-driven and/or strategically based campaigns, in relation to their respective contributions to online violence rests on a presupposition: that there is, on their part, simply a misunderstanding of the impact of their behaviour rather than a clear understanding that what they do or, in the case of bystanders, fail to do, creates, drives and sustains online VAWG.

A key issue is the preferred focus of policy drivers and strategies in targeting online bystanders. Thus, is the desire to (1) have them recognise the impact of, and desist from participating in, their passive and/or active engagement as online bystanders, or (2) have them engage directly and overtly with the online perpetrator and other bystanders in an attempt to suppress their output and negative bystander response rate (in the form of comments or 'likes', etc.) and/or to encourage others to do so.

In relation to one common form of online violence, cyberbullying, there are a number of websites that provide support for the victim and/or routes through which the bystander might report the incidents they witness. However, the relatively comprehensive dual approach (that is, serving victim and bystander) is missing in relation to most online violence offences.

In broad terms, it has been suggested that '[d]igital abuse, harassment, and violence are ... simultaneously both extensions of conventional forms of violence, and at the same time, they produce new types of harm that must be addressed using a variety of legal and non-legal strategies' (Powell and Henry 2017, p. 214).

To that end, consideration should be given to a combination of micro (individual), meso (organisational) and macro (societal) responses.

At the micro level, as with physical violence, specialised support services for victims of online violence need to be provided and perfected. This is because long-term psychological and emotional harm is suffered as a result of the original online violence coupled with its prolongation by the behaviour of bystanders. Moreover, the injunction for women and children to 'switch off' their devices or desist from engaging in the social media environment following online violence would have debilitating effects beyond the trauma suffered. Rather, technology provides women and children with significant opportunities to report ill-treatment, as well as to gain a degree of freedom via educational opportunities, health advice etc.

The plethora of approaches created to reduce the motivation for offending in the physical space could be adapted to focus on the online dynamic. Beyond that, at the macro level, there should be education-based initiatives such as endeavouring to create, nurture and apply a notion of 'digital citizenship', which refers to an individual and organisation. In this context, a commitment by intermediaries' (such as Facebook) to protect users' '...capability to partake freely in the internet's diverse political, social, economic and cultural opportunities, which informs and facilitates their civic engagement', could be introduced (Citron and Norton 2011). The realisation of this citizenship lies in the conscious efforts of users not only to be ethical citizens in terms of their own engagement but, moreover, to

intervene (by way of complaining, documenting and reporting online violence and associated activity) when they witness unethical behaviour in others.

At the meso level, reliance in relation to this citizenship also requires intermediaries to proactively patrol and police their platforms and to effectively react to reports received from individuals. This seemingly simple stratagem, if resolutely employed, could significantly impact the prevalence and/or effects of online violence. However, it must be situated within jurisdictions where, for example, the right to freedom of speech might be sacrosanct, and, more broadly, it must consider the exponential growth in the volume of online messaging and posting across many platforms.

The corollary of these premises, however, is that, unlike school-aged cyberbullying activity (where the relative immaturity of both bully and bystander and the fixation with social media are influential factors in the online space), both the perpetrators and bystanders in relation to other forms of online violence are arguably of greater maturity and possessed of a greater propensity to recognise both the nature and impact of their respective behaviour, acts and omissions.

The fact that, despite those points of maturity-related advantage, online violence and the bystander response and/or failure to respond continue to occur – and to occur in relation to objectively disturbing witnessed events of life and death – is problematic. In that context, to assume, as a policy driver or strategic premise, that both perpetrator and bystander simply do not appreciate their respective impacts on the victim is arguably naïve. To that end, therefore, at the meso level, the direction of policy and strategic responses should arguably be focused on the way regulatory and legislative architecture is configured. This will effectively force both perpetrator and bystander to a point of recognition (as a result of criminal or civil action brought against them and/or permanent removal from the various networks that facilitate their respective contribution) as to the existence and/or exacerbation and/or perpetuation of online violence resulting from their respective behaviour.

As discussed above, there is a good deal of existing legislation that either speaks directly to certain (if limited) forms of online violence or is currently phrased with sufficient ambiguity to bring perpetrators and bystanders within its ambit. Where existing legislation does not provide either sufficient coverage of the myriad of online offences or does not allow for bystanders to be brought within its ambit, 'good Samaritan' legislation might be amended to include (alongside the limitation of liability for bystanders) a requirement to assist unless too dangerous to do so. 'Bad Samaritan' legislation might also be introduced that would create for bystanders a statutory duty to intervene (even if, in terms of the actions possible in the online space, that simply involved alerting the authorities), subject to the caveats noted earlier.

7. Conclusion

The definition of online violence is complex and encompasses a range of activities that might not fit within the notion of 'violence' as it is understood in the physical environment. In this sense, given the ubiquity of cyberspace, its extra-territorial reach, its centrality in the social, economic and political spheres and its propensity to facilitate rapid deposition of anonymous multifarious content with immediate often debilitating impact, the definitional parameters of 'violence' have rightly been extended in the online context. The demonstrable level and impact of online violence suggests that the point of government, law enforcement or civil society intervention may already have passed, or at least placed those wishing to reduce its occurrence in a disadvantageous position. In that sense, the logical and pragmatic response might be simply to engage with the situation that exists now rather than to devote a disproportionate amount of time to improving the situation in the future, although that must remain, in terms of prevention, a key underpinning driver.

As Weisel (1986) has observed, '[w]e must take sides. Neutrality helps the oppressor, never the victim. Silence encourages the tormentor, never the tormented. Sometimes we must interfere'. To that end, the situation as it stands now is that online violence against women and girls is growing exponentially. There are two reasons for that growth: first, the ease with which perpetrators can utilise the online space to undertake their attacks; and second, bystanders who, through a combination of blunt action and/or reaction and/or omission and/or apathy, create, drive and sustain the perpetrators' behaviour and thereby exacerbate the harm sustained by victims.

In consequence, therefore, the logical and most direct point of intersection lies in a combination of approaches. These include more vigorous and proactive application of current, and the creation of new, legislation; and more systematic engagement by intermediaries in terms of (1) policing their sites, (2) engaging with their users and (3) encouraging them (including through provided platform conduits) to report online violence. Moreover, there needs to be a more effective reaction to any malfeasance identified by those intermediaries directly and/or or through their user engagement. As a corollary to that approach, it is essential that governments systematically engage with intermediaries and, where necessary, regulate to ensure adherence to the rooting out of online malfeasance on the part of perpetrators and bystanders.

Bibliography

Al-Alosi, H, A Vidahata and A Maurushat (2016), 'The role of bystanders in cyberbullying', *Precedent*, 132, 20–24.

Ames, JB (1908), 'Law and morals', *Harvard Law Review*, Vol. 22, No. 2, 97–113.

Anderson, M and S Toor (2018), 'How social media users have discussed sexual harassment since #MeToo went viral', Pew Research Center [sic], available at: www.pewresearch.org/fact-tank/2018/10/11/how-social-media-users-have-discussed-sexual-harassment-since-metoo-went-viral/ (accessed 17 November 2022).

Andersson, L and E Sundin (2021), 'Mobile bystanders and rubbernecks, disaster tourists, and helpers: Towards a theoretical framework for critically studying action possibilities at accident sites', *Mobile Media & Communication*, Vol. 9, No. 3, 531–545.

APC (Association for Progressive Communications) (2017), 'Online gender-based violence: A submission from the Association for Progressive Communications to the United Nations Special Rapporteur on violence against women, its causes and consequences', APC, November, available at: www.apc.org/sites/default/files/APCSubmission_UNSR_VAW_GBV_0_0.pdf (accessed 17 November 2022).

Ashworth, A (1989), 'The scope of criminal liability for omissions', *Law Quarterly Review*, Vol. 105, 424.

Baker, L, M Campbell and E Barreto (2013), 'Understanding technology-related violence against women: Types of violence and women's experiences', Learning Network Brief 6, Centre for Research and Education on Violence Against Women and Children, London, ON.

Barron, G and E Yechiam (2002), 'Private e-mail requests and the diffusion of responsibility', *Computers in Human Behavior*, Vol. 18, No. 5, 507–520.

Barton, A and H Storm (2014), *Violence and Harassment against Women in the News Media: A Global Picture*, International Women's Media Foundation and International News Safety Institute, Washington, DC and London.

BBC (2017), 'Facebook Live 'broadcasts gang rape' of woman in Sweden', available at: https://www.bbc.co.uk/news/world-europe-38717186 (accessed 24 November 2022).

Benzmiller, H (2013), 'The cyber-Samaritans: Exploring criminal liability for the "innocent" bystanders of cyberbullying', *Northwestern University Law Review*, Vol. 107, No. 2, 927–962.

Blair, C.A, L Foster Thompson and K.L. Wuensch (2005), 'Electronic Helping Behavior [sic]: The Virtual Presence of Others Makes a Difference', *Basic and Applied Social Psychology*, 27(2), 171–178

Broadband Commission for Digital Development (2015), 'Cyber violence against women and girls: A world-wide wake-Up call', Working Group on Broadband and Gender Discussion Paper, available at: www.broadbandcommission.org/publication/cyber-violence-against-women/ (accessed 17 November 2022).

Citron, DK and H Norton (2011), 'Intermediaries and hate speech: Fostering digital citizenship for our information age', *Boston University Law Review*, Vol. 91, 1435–1483.

Council of Europe Working Group on Cyberbullying (2018), *Mapping Study on Cyberviolence*, Council of Europe, Strasbourg.

Coyl, DD (2009), 'Kids really are different these days', Phi Delta Kappan, Vol. 90, No. 6, 404–407, cited in Benzmiller, H (2013), 'The cyber-Samaritans: Exploring criminal liability for the "innocent" bystanders of cyberbullying', *Northwestern University Law Review*, Vol. 107, No. 2, p927–962.

Darley, JM and B Latane (1968), 'Bystander intervention in emergencies: Diffusion of responsibility', *Journal of Personality and Social Psychology*, Vol. 8, No. 4 (Pt.1), 377–383.

Devlin, P Hon Sir (1959), 'Maccabaean lecture in jurisprudence: The enforcement of morals', In *Proceedings of the British Academy*, 130–151, available at https://psi329.cankaya.edu.tr/uploads/files/Devlin,%20The%20Enforcement%20of%20Morals%20(1959)(1).pdf (accessed 17 November 2022).

Domínguez-Hernández, F, L Bonell and A Martínez-González (2018), 'A systematic literature review

of factors that moderate bystanders' actions in cyberbullying', *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, Vol. 12, No. 4, article 1.

Duggan, M. (2017), 'Online harassment 2017', Pew Research Center, 11 July, available at: www.pewresearch.org/internet/2017/07/11/online-harassment-2017/ (accessed 17 November 2022).

eSafety Commissioner (n.d.) , 'Online abuse targeting women', Australian Government, available at: www.esafety.gov.au/women/online-abuse-targeting-women (accessed 17 November 2022).

Parliament, European (2018), 'Cyber violence and hate speech online against women', study for the FEMM Committee, available at: www.europarl.europa.eu/RaegData/etudes/STUD/2018/604979/IPOL_STU(2018)604979_EN.pdf (accessed 17 November 2022).

Grande Montana, P (2018), 'Watch or report? Livestream or help? Good Samaritan laws revisited: The need to create a duty to report', *Cleveland State Law Review*, Vol. 66, No. 533, available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3204470 (accessed 17 November 2022).

Guiora, AN and JE Dyer (2020), 'Bystander legislation: He ain't heavy, he's my brother', *Kansas Journal of Law & Public Policy*, Vol. 49, No. 2, 291–328.

Haber, E (2020), 'The digital Samaritans', *Washington and Lee Law Review*, Vol. 77, 1559–1645.

Henry, N, A Flynn and A Powell (2019), 'Responding to "revenge pornography": Prevalence, nature and impacts', Report to the Criminology Research Advisory Council, available at: www.aic.gov.au/sites/default/files/2020-05/CRG_08_15-16-FinalReport.pdf (accessed 17 November 2022).

Internet Governance Forum (2015), 'Best Practice Forum (BPF) on online abuse and gender-based violence against women', November, available at: https://schd.ws/hosted_files/igf2015/1d/Draft%20JP_Online%20Abuse%20and%20Gender%20Based%20Violence%20Against%20Women.pdf (accessed 21 November 2022).

Inter-Parliamentary Union and Parliamentary Assembly of the Council of Europe (2016), *Sexism, harassment and violence against women parliamentarians*, Issues Brief, October 2016, pp.3, available at: https://www.ipu.org/resources/publications/issue-briefs/2016-10/sexism-harassment-and-violence-against-women-parliamentarians (accessed 24 November 2022)

Kaufman, ZD (2021), 'Digital age Samaritans', *Boston College Law Review*, Vol. 62, No. 4, 1117–1192.

Kowalski, RM, SP Limber and PW Agatson (2012), *Cyberbullying: Bullying in the Digital Age*, John Wiley & Sons, Inc., Hoboken, NJ, cited in Benzmiller, H (2013), 'The cyber-Samaritans: Exploring criminal liability for the "innocent" bystanders of cyberbullying', *Northwestern University Law Review*, Vol. 107, No. 2, 927–962.

Le Nguyen, C and W Golman (2021), 'Diffusion of the Budapest Convention on Cybercrime and the development of cybercrime legislation in Pacific Island countries: "Law on the books' vs 'law in action"', *Computer Law & Security Review*, Vol. 40, April, 105521.

Love, H (2018), 'Not so innocent bystanders: A case for criminal liability for social media posts', paper presented at Conference: Canadian Law and Society Mid-Winter Meeting, available at: www.researchgate.net/publication/322901420 (accessed 17 November 2022).

Lytle, RD, TM Bratton and HK Hudson (2021), 'Bystander apathy and intervention in the era of social media', in Bailey, J, A Flynn and N Henry (Eds.), *The Emerald International Handbook of Technology-Facilitated Violence and Abuse*, Emerald Publishing, Bingley, UK, 711–728.

Markey, PM (2000), 'Bystander intervention in computer-mediated communication', *Computers in Human Behavior*, Vol. 16, No. 2, 183–188.

Moretti, C and D Herkovits (2021), 'Victims, perpetrators, and bystanders: A meta-ethnography of roles in cyberbullying', *Cadernos de Saúde Pública*, Vol. 37, no. 4, 1–18, available at: www.scielo.br/j/csp/a/5tqGgm7fGVs8xDsKqmW9v7r/?lang=en&format=pdf (accessed 17 November 2022).

Obermaier, M, N Fawzi and T Koch (2014), 'Bystanding or standing by? How the number of bystanders affects the intention to intervene in cyberbullying', *New Media & Society*, Vol. 18, No. 8, 1491–1507.

OAS (Organization of American States) (1994), *Inter-American Convention on the Prevention, Punishment and Eradication of Violence against Women ('Convention of Belém do Pará')*, 9 June.

OAS (Organization of American States) (2015), 'Declaration on Political Harassment and Violence against Women', Follow-up Mechanism to the Belém do Pará Convention, Lima, Peru.

OAS (Organization of American States (2019), 'Combating online violence against women: A call for protection', White Paper Series, Issue 7, available at: www.oas.org/en/sms/cicte/docs/20191125-ENG-White-Paper-7-VIOLENCE-AGAINST-WOMEN.pdf (accessed 17 November 2022).

Pacer's National Bullying Prevention Center (2020), 'Bullying statistics: By the numbers', available at: www.pacer.org/bullying/info/stats.asp (accessed 17 November 2022).

Powell, A and N Henry (2017), 'Sexual violence and harassment in the digital era', in Deckett, A and E Sarre (Eds.) (2017), *The Palgrave Handbook of Australian and New Zealand Criminology, Crime and Justice*, Springer Nature, Cham, Switzerland.

Shariff, S and DL Hoff (2011), 'Cyber bullying: Legal obligations and educational policy vacuum', In Jaishankar, K (Ed.), *Cyber Criminology: Exploring Internet Crimes and Criminal Behavior*, CRC Press, Boca Raton, FL, 359–392, cited in Benzmiller, H (2013), 'The cyber-Samaritans: Exploring criminal liability for the "innocent" bystanders of cyberbullying', *Northwestern University Law Review*, Vol. 107, No. 2, 927–962.

Stephen, FJ(1883) , *A History of the Criminal Law of England*, Vol. III, Cambridge University Press, Cambridge.

Tas'adi, R, N Gistituati Mudjiran and A Ananda (2020), 'Cyberbullying in the digital age: A common social phenomenon', *Advances in Social Science, Education and Humanities Research*, Volume 504, Proceedings of the 2nd International Conference Innovation in Education, 196–200.

UN Women (United Nations Entity for Gender Equality, the Empowerment of Women) (2020), *Online Violence against Women in Asia: A Multi-Country Study*, UN Women Regional Office for Asia and the Pacific, Bangkok.

United Nations General Assembly (1993), 'Declaration on the Elimination of Violence against Women', A/RES/48/104.

Valdés-Cuervo, AA, C Alcantar-Nieblas, LG Parra-Pérez, GM Torres-Acuna, FJ Alvarez-Montero and H Reyes-Sosa (2021), 'Unique and interactive effects of guilt and sympathy on bystander aggressive defender intervention in cyberbullying: The mediation of self-regulation', *Computers in Human Behavior*, Vol. 122,106843.

Wiesel, E (1986), 'Nobel Peace Prize acceptance speech', available at: https://www.nobelprize.org/prizes/peace/1986/wiesel/acceptance-speech/(accessed 17 November 2022).

Wong, RYM, CMK Cheung, B Xiao and JB Thatcher (2021), 'Standing up or standing by: Understanding bystanders' proactive reporting responses to social media harassment', *Information Systems Research*, Vol. 32, No. 2, 561–581.

Yee Man, L (2021), 'Technology-facilitated domestic abuse and culturally and linguistically diverse women in Victoria, Australia', in Bailey, J, A Flynn and N Henry (Eds.), *The Emerald International Handbook of Technology-Facilitated Violence and Abuse*, Emerald Publishing, Bingley, UK, 447–467.

Statutes

Australia

Federal
Federal Criminal Code Act 1995

Australian Capital Territory
Crimes Act 1990
Civil Law (Wrongs) Act 2002

New South Wales
Crimes Act 1900
Civil Liability Act 2002

Queensland
Criminal Code 1899
Law Reform Act 1995

South Australia
Summary Offences Act 1953
Civil Liability Act 1936

Tasmania
Police Offences Act 1935
Civil Liability Act 2002

Victoria
Summary Offences Act 1966
Wrongs Act 1958

Western Australia
Restraining Orders Act 1997
Civil Liability Act 2002

Northern Territory
Personal Injuries (Liabilities and Damages) Act 2003

New Zealand
Harmful Digital Communications Act 2015
Crimes Act 1961

Pacific Islands

Cook Islands
Telecommunications Act 2019
Copyright Act 2013

Fiji
Crimes Act 2009
Posts and Communications Decree 1989

Kiribati
Telecommunications Act 2004

Nauru
Cybercrime Act 2015
Telecommunications Act 2002
Crimes Act 2016

Papua New Guinea
Cybercrime Code Act 2016
Protection of Private Communications Act 1973
Criminal Code Act 1974

Samoa
Crimes Act 2013
Telecommunications Act 2005
Copyright Act 1998

Solomon Islands
Telecommunications Act 2009

Tonga
Computer Crimes Act 2003
Communications Act 2015
Pornography Control Act 2002
Copyright Act 2002

Tuvalu
Tuvalu Telecommunications Corporation Act 2008 (Revised Edition)

Vanuatu
Penal Code (Consolidation Edition) 2006
Telecommunications Act (Consolidated Edition) 2006

Caselaw

Stovin v. Wise [1996] 3 All ER 801 (HL) 819

Commonwealth Secretariat
Marlborough House, Pall Mall
London SW1Y 5HX
United Kingdom

thecommonwealth.org

The Commonwealth

D19099