Policing diversity in the digital age: Maintaining order in virtual communities

David S. Wall and Matthew Williams

Criminology and Criminal Justice 2007 7: 391
DOI: 10.1177/1748895807082064

The online version of this article can be found at:
http://crj.sagepub.com/content/7/4/391
Policing diversity in the digital age: 
Maintaining order in virtual communities

DAVID S. WALL AND MATTHEW WILLIAMS
University of Leeds, UK and Cardiff University, UK

Abstract

Members of ‘terrestrial’ communities are migrating in ever-increasing numbers to a new ‘Third Space’ that manifests outside traditional geographical physical boundaries. This online space consists of purely social relations where interaction and community are performed at-a-distance. The diversifying populations of these virtual villages, towns and cities now constitute very real communities. Online non-gaming spaces such as Ebay, Active Worlds and Secondlife, for example, deliberately utilize the discourse of community in an attempt to instil a sense of communal space and shared responsibility among their members. While the majority subscribe to the rhetoric of ‘netizenship’ others find alternative means to participate online. The avocations of these few have resulted in the endemic deviance/crime problem that exists online. As a result, online communities have developed their own distinct history of control and regulation.

This article explores the ways that online social spaces maintain orderly ‘communities’. It contrasts ‘proximal’ (online) forms of governing online behaviour, such as online reputation management systems, ‘virtual’ police services and vigilante groups that employ ‘online shaming’, with ‘distal’ (offline) forms such as offline policing and criminal justice processes. The central theme of the article is a critical account of how these, often contradicting, nodes of governance interact.

Key Words

cybercrime • governing online behaviour • online community • policing cyberspace • shaming • virtual community
Introduction

Within academic and policy domains the thinking around the policing of contemporary ‘communities’ has been largely preoccupied by common-sense notions of diversity. Communities are seen as internally fragmented in terms of ethnicity, geography, religion, sexuality, class and age. Some groups are identified as ‘hard-to-reach’ (Jones and Newburn, 2001) signalling their abstraction from the ‘homogeneous’ elements of the community within which they live. Reassurance and neighbourhood policing initiatives have targeted these groups in an attempt to fill the ‘reassurance gap’ (ACPO, 2001; Innes, 2004: 151). In tandem, conceptual and empirical work on the plurality of policing has shown how the ‘mixed economy’ of patrol in relation to these initiatives is both beneficial and problematic (Crawford and Lister, 2006; Hughes and Rowe, this issue). Attempts at making communities and the isolated groups within them safer while also reducing fear of crime have shown varying degrees of success in the United States and Britain (Skogan and Steiner, 2004; Tuffin et al., 2006). Innes (2006a) argues that part of the reason for the ‘pockets of success’ is the ‘discursive’ shift from ‘community policing’ to notions of ‘reassurance’ and ‘neighbourhood’ policing. Innes also notes that this is more than simple rhetoric as it reflects an ‘increasing awareness among police officers about the complexity of the policing environment’ (2006b: 96). This complexity is reflected in the recognition that in late modern society ‘traditional’ definitions of community are subject to challenge and, in response, ‘neighbourhood’ becomes a site and focus for policing with specific bounded geodemographies. This article questions the extent to which this awareness and recent trends in policing, extends beyond notions of community, neighbourhood and diversity into the domain of ‘non-terrestrial’ cyberspaces. These ‘cyber-communities’ and ‘cyber-neighbourhoods’ can be conceptualized as deeply embedded within a mixed economy of crime control and in need of reassurance in the face of a burgeoning cybercrime problem.

This article seeks to explore the different ways in which online social spaces maintain orderly ‘communities’. It begins by reviewing our understanding of online communities and the various textual and graphical formats that give rise to them, it then looks at cybercrime and the types of disorder that are specific to virtual environments. Next, the discussion outlines the various methods of ‘policing’ online behaviour that are exercised by proximal (online) nodes of governance such as online reputation management systems, virtual ‘police’ services and vigilante groups that employ ‘online shaming’. These methods of online governance are then compared with the more ‘traditional’ forms of distal (offline) governance found in offline policing and criminal justice processes. Towards the latter part of the article qualitative data from research conducted within an online community provides an insight into community members’ perceptions of various forms of control. A critical account of how these contrasting forms of governance interact in often conflicting ways forms the central theme of the article.
A third space for a second life

Members of ‘terrestrial’ communities are increasingly migrating to a new ‘Third Space’ (Oldenburg, 1999) that manifests outside of traditional geographical physical boundaries. A space in which people can live out, what Presdee (2000: 62) has described as, their ‘second life’, which intersects with the virtual representations of the State and also corporate interests. This online space consists of purely social relations where interaction and community are performed at-a-distance (Lash, 2001). Not only do the virtual villages, towns and cities constitute very real communities for their diverse populations, but this characteristic is now being utilized. Online non-gaming spaces such as Ebay, Cyberworlds and Secondlife, for example, deliberately utilize the discourse of community in an attempt to instil a sense of communal space and shared responsibility among their members. While the majority subscribe to the rhetoric of ‘netizen- ship’, others find alternative and more destructive means to participate online. The avocations of these few have resulted in the endemic deviance/crime problem that exists online. As a result, online communities have developed their own distinctive histories of control and regulation.

The cultural, spatial and technological shifts that characterize the ‘information age’ have also been used to explain the increasing use of computer-mediated communications for social interaction. Jones notes that ‘crucial to the rhetoric surrounding the Internet … is the promise of a renewed sense of community and, in many instances, new types and formations of community’ (1998: 3). The increasing numbers of individuals who are experiencing culture-at-a-distance may also be experiencing community-at-a-distance (Lash, 2001). Like Jones (1998), others have written on the growth and development of social formations online (Rheingold, 1993; Turkle, 1995; Baym, 1998; Markham, 1998; Reid, 1999; Wellman and Gulia, 1999). At the core of much of this literature is a series of debates about the existence of ‘virtual community’. Rheingold (1993) was the first author to represent these increasingly complex social relations and formations as a form of community. He attributed the escalation of online social interaction to the demise of contemporary offline communities in the United States. In the same way that Lash (2001) talks of ‘technological forms of life’, Rheingold (1993) sees the proliferation of social groups sustained by non-tangible technological spaces. However, Rheingold’s (1993) utopian perception is quite unique. Lash (2001) writes of the thinning and stretching of social relations in a time of technological dependence, while Rheingold (1993) considers online social formations to be so rich and meaningful that they warrant being called communities. This bifurcation is at the centre of the ‘virtual community’ debate and raises the question over whether or not social formations can be maintained in a non-tangible technological environment where interlocutors experience interaction at-a-distance. If they can, then are these formations meaningful enough to warrant being considered communities?
Critics of online community

The work of Beniger (1987) and Peck (1987) on pseudo-community has been used as a comparison with ‘virtual community’. The demise of gemeinschaft community, characterized by communal relationships, and the subsequent rise of gesellschaft relations, typified by impersonal indirect relations, is taken as the basis of pseudo-community. Beniger (1987) discusses the demise of personal relations in terms of increasing mass communication. The primary criticism here is the lack of authenticity in pseudo-community communication. Clear parallels can be drawn with online social relations. Simply put, non face-to-face communication, with a reduction in social cues and a lack of presence, results in fleeting encounters characterized by ephemerality, non-linearity and a lack of depth (Beniger, 1987). Rheingold’s (1993) central question was whether computer-mediated communication would allow for a community divorced from Beniger’s (1984) and Peck’s (1987) pseudo-relations.

Before this question can be adequately answered, the criticisms of those who oppose the use of the term community to describe online social formations have to be addressed. Fernback draws attention to the central difficulty when dealing with such slippery concepts: ‘Community is a term which seems readily definable to the general public but is infinitely complex and amorphous in academic discourse. It has descriptive, normative, and ideological connotations … [and] encompasses both material and symbolic dimensions’ (1997: 39).

The traditional conceptualization of community emphasizes its physical, temporal and moral dimensions: a common geographic territory; a common history, common value system; a shared religion and language (also see discussion in Hughes, 2007). Using this definition several conflicts begin to emerge. First, while the Internet can be considered to be a space, it is not physical. Individuals cannot physically live in cyberspace. Second, there are questions over the extent to which groups in cyberspace can have a shared history, value system, religion or language. Commentators including Lockard (1997) are reluctant to apply the term community to online social formations. They state that communication alone cannot replace community’s manifold functions. Others such as Healy (1997) argue that with such a small fraction of the world’s population online, social formations do not oblige their participants to deal with diversity. In this sense, online social formations are no more than voluntary associations of like-minded people, without the additional demands of offline community, which helps forge its very nature.

Further criticisms revolve around the technologically deterministic nature of Rheingold’s (1993) claims. Morley and Robins (1995) reject Rheingold’s ideas labelling them conservative and nostalgic. The idea that technological change and innovation can turn around the social and cultural decay in contemporary American society is naive. In line with contemporary thinking around the demise of offline community and the emergence of culture at-a-distance, Wittel (2001) draws attention to the way that much of the literature discusses the imaginary aspects of virtual community (Anderson, 1983).
This in itself is an indication that online social formations are not communities in the traditional sense. Further, the idea of the ‘virtual community’ is arguably flawed in its presumed dichotomy between the ‘real’ and the ‘unreal’. Many writers have taken the presumed dichotomy of the ‘real’ and the ‘virtual’ as the basis for their analyses (see Virilio, 1997; Baudrillard, 1998). However, the empirical studies that do exist in this area have repeatedly shown that experiences online are not considered as ‘virtual’ or apart from ‘real’ life, and that this presumed dichotomy is a false one (Markham, 1998; Miller and Slater, 2000).

Yet, there still remains the vexed question as to the quality of any online community, if indeed it exists. That virtual relationships within online communities contain neither the full panoply of social relationships nor the cohesive or organic expectations of traditional gemeinschaft community relationships speaks for itself because the complex social stimuli which create them are largely absent. Instead, there emerges a line of argument that the Internet has been successful in supporting the subsequent growth of new social ‘networks of networks’ (Licklider and Taylor, 1968/1990: 38). Individuals lock into networks that give them the illusion of satisfying emotional or informational needs, but these experiences are nevertheless restricted, and despite expectations to the contrary, networked individuals cannot encounter a holistic community experience. This illusory community is a form of gesellschaft, which encourages the social deskilling of the individual through the specialization and compartmentalization of interactions.

Following this line of debate, Wittel (2001) argues that online social formations are in fact a form of networked sociality rather than community. He notes that in contrast to community, which involves strong long-lasting ties, proximity and a common history, network sociality is characterized by informational relations, being devoid of history, and based more upon an exchange of data and on ‘catching up’. This concept of network sociality resonates with Sennett’s (1992, 1998) commentary upon the decline of long-term, sustained, deep relationships. In a society where ‘the personality’ transcends politics (Sennett, 1992: 238), individuals are no longer accustomed to doing things together or performing public roles and are therefore unable to find community and the intimacy it gives. Community, in its satisfying and organic form, is no longer to be found. So, when culturally reinforced expectations of community become embedded in policy as a desired outcome, the result is what Sennett calls ‘destructive gemeinschaft’ (1992: 238). In other words, expecting to find and utilize the qualities of community when it does not exist can have adverse consequences.

Sennett’s (1992, 1998) analysis provides a stark contradiction to Rheingold’s (1993) utopian vision. Instead of computer-mediated communications providing a stage for a renewed sense of community, Sennett paints a dystopic picture of relations devoid of trust, commitment and loyalty. However, Wittel’s (2001) network sociality is a less bleak interpretation of contemporary social relations. While it does not depend upon shared history and narratives, replacing them with an exchange of experiences and
biographies, trust and commitment can still be engineered. Instead of basing trust upon the knowledge of someone’s character, knowledge of their resources or social position can be used. Giddens’ (1994) notion of ‘active trust’ is such an example.

With the lack of meaningful encounters Lockard (1997) speaks of, and the banality of interaction Healy (1997) refers to, network sociality may seem a far more accurate analysis of online social relations. However, most of the analysis in the early to mid-1990s of computer-mediated communications seems parochial and lacks empirical insight. More and more is the Internet now being considered a site for empirical enquiry and a phenomenon intrinsically linked to cultural and social processes that occur offline. This empirical work has highlighted how the Internet is a collection of very different social spaces, all with unique characteristics that shape the way in which individuals interact. To talk of the Internet or computer-mediated communications as one form of communication is short-sighted (Miller and Slater, 2000). Further, empirical work has highlighted how individuals use computer-mediated communications very differently (Markham, 1998). Aggregated experiences of Internet social interaction are no longer an accurate reflection of social relations online.

Of the studies that have been conducted within online communities, the majority have focused upon text-based environments, limiting the ability to form a general analysis of computer-mediated communications as a whole—for examples see Baym (1998); Danet (1998); Markham (1998); Burkhalter (1999). The analyses from these studies have been controversial in that they challenge the majority of the criticisms of online community. The idea that networked society has ‘thinned’ out interaction, and that notions of history and narrative are no longer integral to sociality are being questioned. The analysis presented in this article supports both Markham’s (1998) and Miller and Slater’s (2000) analyses, and rejects the blanket application of the term ‘network sociality’ to explain all social relations online. Rather, it is argued here that ‘history’ and ‘meaning’ are an integral part of some online social formations, and that while some exchanges are characterized as ephemeral, they are not simply reducible to mere informational exchanges—they have social meaning.

**Online graphical communities**

The data used in this article were gathered from an online graphical community called Cyberworlds, one of a handful of graphically represented three-dimensional virtual reality online social spaces that form part of the Internet (see further Williams, 2006). An online focus group comprised of 70 community members ran over three months to generate data on perceptions and experiences of online community, online deviance and its control (see further Stewart and Williams, 2005; Williams, 2007a).

Online communities like Cyberworlds are second-generation Multi-User Domains (MUDs) that are much more technologically sophisticated than
their predecessors. The traditional MUDs developed in the late 1970s used text to depict action, description and communication. Users therefore had to rely upon their imagination to create the environment being described to them. Cyberworlds, on the other hand, take advantage of broadband network communications, which allow much more information to be transmitted in shorter periods of time. This means that information portraying action, description and communication can be represented by more than just simple text. A graphical window allows a user to see themselves represented as an avatar—a three-dimensional persona. Via this interface users can locate each other and navigate around their three-dimensional environment. Text still functions as the means of communication and is essential to the maintenance of the ‘community’, but it has become relegated where visual components now represent action and description. Cyberworlds and other second-generation MUDs have grown in popularity over the past 10 years due to the level of social and ‘physical’ immersion they provide to the user.

As increasing numbers of people became connected to computer networks, social spaces emerged to provide an escape from the ‘real’ world. MUDs and the social networks they facilitate have been the subject of much research in media and communications disciplines (see Jones, 1995, 1997, 1998 for an overview). Many studies have focused upon the substantive issues of race, gender and sexuality within MUDs and other associated technologies (such as Newsgroups and Internet Relay Chat). The effects of anonymity and disembodiment, experienced by each MUD user, have forged much of what is now understood of online social experience (see Dietrich, 1997; Shaw, 1997; Danet, 1998; Kramarae, 1998; Poster, 1998). However, graphical MUDs, like Cyberworlds, have received little academic attention. Furthermore, few sociological studies have attempted to examine empirically the phenomena of cybercrime, disorder and the governance (policing) of online behaviour in communities within these social settings.

### Cybercrime and disorder online

Increasing media attention has elevated cybercrime in the public consciousness (Wall, 2007a: 14). Online paedophilia, cyberterrorism, identity theft, online fraud, malware infections, spams, denial of service attacks, hacking and online hate crime (to name a few) have transformed the public’s perception of the Internet from a new social space associated with unprecedented freedoms into a ‘dangerous place’ riddled with escalating, often misunderstood, risks. While there is fairly widespread consensus that cybercrimes exist, there remains much confusion as to what they actually are and what risks they pose (Wall, 2005a: 77; see Brenner, 2001; Walden, 2003). The confusion over what constitutes a cybercrime creates a ‘reassurance gap’ between crimes experienced and those felt (see Innes, 2004: 151), and leads to public concern about ‘cybercrime’, which subsequently shapes the demands made of the police (for reassurance).
An effective way of identifying if an act is a cybercrime is to apply what Wall (2007a: 34) calls the ‘transformation test’. Simply put, what happens to the ‘crime’ if the Internet is removed? By applying this logic, three different groups of cyber-criminal opportunity can be identified as points on a spectrum. At the one extreme are ‘traditional’ crimes that masquerade as cybercrimes. In such cases the Internet has typically been utilized for communication or information gathering to facilitate an ‘offline’ crime. If the Internet is removed from the activity then the criminal behaviour persists because the offenders will revert to using other information sources or types of communication. A good example is the use of locksmiths’ newsgroups by burglars as documented by Mann and Sutton (1998). ‘Hybrid’ cybercrimes occupy the middle ground. These are ‘traditional’ crimes for which entirely new global opportunities have emerged (e.g. globalized frauds and deceptions, also the global trade in pornographic materials including child pornography). Take away the Internet and the behaviour will continue by other means, but not by the same volume or across such a wide span. At the far end of the spectrum are the ‘true’ cybercrimes which are solely the product of opportunities created by the Internet and which can only be perpetrated within cyberspace. These are the spawn of the Internet and therefore embody all of its transformative characteristics. Spamming is a good example of a true cybercrime. It is an illegal behaviour in its own right in the laws of the USA, EU and many other jurisdictions, but it also facilitates secondary offending by enabling engagement with potential victims (Wall, 2005b). Many of the offences that result are small-impact bulk victimizations—de minimis offences. Take away the Internet and spamming and true cybercrimes vanish (Wall, 2007: 52–129).

Along this spectrum exist myriad crimes and misdemeanours. ‘Computer-integrity’ crimes are offending behaviours that assault the integrity of network access mechanisms. They include hacking and cracking, vandalism, spying, denial of service and the planting and use of viruses and Trojans. Many jurisdictions now have legislation such as the Computer Misuse Act 1990 (UK), the Computer Fraud and Abuse Act 1986 (USA) (18 USC 1030), internationally harmonized conventions such as the Council of Europe’s Convention on Cybercrime (ETS No. 185), to protect against unauthorized access to computer material; unauthorized access with intent to commit further offences; and unauthorized modification of computer material. ‘Computer-related’ crimes are committed using networked computers to engage with victims in order to acquire dishonestly cash, goods or services. In addition are socially engineered variants such as ‘phishing’, advanced fee frauds and the manipulation of new online sales environments, particularly auction sites. Again, most jurisdictions have legislation concerning thefts and provide legal measures for the recovery of lost assets, as well as intellectual property laws to protect against the unauthorized exploitation of intellectual property. ‘Computer-content’ crimes relate to the content of computers—materials held on networked computer systems. They include the trade and distribution of pornographic materials, the dissemination of hate crime materials and, more recently, the publication of video nasties of the murders of kidnapped foreign
nationals. Once more, most jurisdictions have variants of the obscenity laws and laws that prohibit incitement, although their legislative strength can vary where Internet content is also protected by laws of free speech. In common with the other two crime groups, legislation does nevertheless vary across jurisdictions in terms of judicial seriousness. Particularly challenging is where the cybercrimes are both specific to virtual environments and actually take place within them.

Disorder in online communities

Online communities are plagued by variants of the cybercrimes and misdemeanours mentioned previously. We see, for example, in computer gaming environments the increased criminal exploitation of gaming artefacts that have strategic importance in online role play gaming. Players in Project Entropia need to obtain artefacts that sustain their place in their games and help them progress through it. The artefacts are therefore highly desired because they represent not only high levels of ability and power, but also the hours of labour put into their construction. Because of this, players are willing to pay large amounts of real money for them. In 2004, for example, a virtual island was sold on eBay for $26,500 (£13,700) and in 2005 a virtual space station went for $100,000 (£56,200) (BBC, 2005a). Consequently, the high values of these artefacts have generated a string of new criminal opportunities. Already there have been examples of buyers being defrauded through e-auction sales, artefacts being stolen from players’ accounts by hackers and even an ‘online mugging’ where a Japanese student was subsequently arrested for using automated bots in a ‘first person shooter’ game by making his avatar move faster than other players and shoot with pinpoint accuracy, thus attacking fellow players and stealing items (BBC, 2005b). The challenges that these forms of offending pose for criminal justice systems are considerable, not least because the victims can point to real economic harms done to them through the illegal use, or sale, of their ‘virtual currency’. A key question is how best to represent legally the loss in the victims’ interests.

Non-gaming-oriented communities, such as Cyberworlds, also experience variants of cybercrime. Both citizens and tourists fall victim to harassment and textual assaults (see Williams, 2001), vandalism of private virtual property (see Williams, 2004), identity theft, intellectual property theft, obscenity and profanity (see Williams, 2006). The meaning and status of these harmful and disruptive acts was initially shaped by community-based understandings of what was, and was not, acceptable behaviour within online environments. The tacit rules that emerged were initially enforced ad hoc by vigilante-type actions, however, they subsequently became incorporated by community members and the corporate arm of the Cyberworlds venture into a formal community charter enforced by ‘Peacekeepers’, the local online private policing service. It is hard to identify whether the ‘true’ impetus behind the labelling of these acts was the virtual public good or,
given the growing market for online communities, whether the corporate element of the community encouraged the condemnation of disruptive acts in an attempt to protect its assets. The answer is probably a combination of both. What is clear, however, is that cybercrimes within these communities have ‘real’ consequences for ‘real’ people. They are shaped by the both the distal offline factors that can influence the architecture of the online environment and also the proximal behaviour of participants. Any remedies therefore have to reflect these characteristics.

**Distal (offline) forms of governance**

**Law and policing**

The use of existing laws to deal with cybercrime can be seen as a distal (offline) node of governance. While law is generally slow to respond to escalating online threats it has been utilized in several cases to prosecute cyber-offenders. The distinctions made between different types of cybercrime previously (‘fake’, ‘hybrid’ and ‘true’ cybercrimes) are important because the first two types tend to be subject to existing laws and existing professional experience can be applied to law enforcement regarding these offences. Any legal problems arising tend to relate more to the application of legal procedures than the substantive law itself. The final group, however, are solely the product of the Internet and pose the greater regulatory challenges. It is also important, of course, to look for any common features in the substantive behaviours. In this way they can be linked to existing bodies of substantive law and associated experience within the criminal justice processes.

Yet law only serves the purpose up to a point. A key challenge is the *de minimis* trap (Wall, 2007a: 161). A common characteristic of many cybercrimes is that they lead to low-impact, bulk victimizations that cause large aggregated losses, which are spread globally, potentially across all known jurisdictions. Consequently, they fall outside the traditional Peelian paradigm of policing dangerousness (risk populations), which frames the police–public mandate. Since local policing strategies often depend upon decisions made at a local level about the most efficient expenditure of finite resources (Goodman, 1997: 486), it is often hard to justify the ‘public interest’ criteria that would release police resources for the investigation of individual cybercrime victimizations.

A second and further challenge is the problem of *nullum crimen* legal disparities in inter-jurisdictional cases. Protocols, including the COE Cybercrime Convention and the establishment of multi-agency partnerships and forums, assist in facilitating inter-force co-operation, but they rely upon the offence in question being given similar priority in each jurisdiction. If, for example, a case is clearly a criminal offence for which the investigation carries a strong mandate from the public, such as the investigation of online child pornography, then resourcing its investigation is usually fairly unproblematic from a
police point of view. However, where there is not such an implied mandate, such as with offences other than child pornography, then resourcing becomes more problematic, especially if the deviant behaviour in question is an offence in one jurisdiction but not another. Of course, the other inter-jurisdictional problem is that there may be cultural differences in defining the seriousness of specific forms of offending, or some offences may fall under civil law in one jurisdiction and criminal law in another, as with the theft of trade secrets, which is a criminal offence in the USA, but a civil offence in the UK. In the UK, only the manner by which trade secret theft takes place falls within the criminal law (see Law Commission, 1997).

While the first two types of cybercrime can be found in the police crime diet, they are not a particularly large part of it and tend to fall within the scope of specialist, rather than everyday, police work while other behaviours, those referred to as true cybercrimes, are entirely alien to the police. This raises questions not only about whether it is the police who should deal with these offences, but also, in the light of the contrast between the high levels of incidence reported by some statistical sources and very low levels of computer misuse prosecutions, who should be policing cyberspace if the police are not? In the UK, for example, during the first decade following the introduction of the Computer Misuse Act 1990 there were only about 100 or so prosecutions against hackers and even fewer convictions (Hansard, 26/3/02, Col. WA35) (Wall, 2007a: 158), and this trend is also found outside the UK (see Smith et al., 2004).

Proximal (online) forms of governance

Technology

Technology is ‘native’ to the environment within which cybercrimes take place and it can be considered a mode of governance. The idea that technology is a more effective regulator of cyberspace than laws has been advanced by Lessig (1999) and others (see discussion in Wall, 2007a: 190–2). It was Lessig’s (1999) aim to counter the technological deterministic view that the Internet could not be regulated. Instead he subscribed to a ‘digital realism’ that recognized the disruptive capacity of technology within cyberspace. Rejecting Boyle’s (1997) notion of an Internet Holy Trinity—that regulation was impossible due to the technology of the medium, the geographical distribution of its users and the nature of its content—Lessig (1999) found that the thread that links all of the Internet’s characteristics together—code or architecture—could be used to control behaviour.

The effectiveness of technology as regulator can be accounted for in several ways. First, technology can disrupt human action, forcing individuals to renegotiate paths and goals (Latour, 2000). Second, technology, code or architecture is malleable; it is easily shaped by actors that have access to its control. Third, the way in which technology imposes constraints on how people can
behave is more pervasive and immediate than other modes of regulation. Fourth, technology is more readily and rapidly adaptive than laws, norms or markets to cyber-criminal threats, allowing it to control both criminal and sub-criminal behaviour. Fifth, changes to system architecture incorporate a preventative approach. It is far more effective to prevent an online offence as opposed to reactively identifying and apprehending an offender. Finally, it is a native form of regulation making it less contentious. Often the origins of the technology are concealed, and hence its regulatory practice is perceived as less coercive than a state-sponsored regime. Technology is then perceived to be more benign, merely shaping—or even facilitating—individual choices (Boyle, 1997). The effectiveness of technology as a regulator then lies in its ability to alter behaviours, its ability to be shaped, its rapid adaptability, its ex ante approach, its wide-reaching scope, its sensitivity towards criminal and sub-criminal activity and its less visible approach to social control.

However, some writers argue that this rather uncritical approach to technomediated regulation is short-sighted. Instead of being a self-executing benign regulator, Hosein et al. (2003) talk of technology as a biased cultural artefact, which is embedded with subjectivity. For this reason alone, there can be no certainty that technology will produce a particular behaviour. Hosein et al. (2003) continue to complicate this relationship. Instead of claiming it is the technology that determines freedom and rights, they take a non-technologically deterministic approach, arguing instead that individuals (codewriters) become the alternative sovereign. Concerns are raised about the accountability of these new masked regulators, and the basis or root of their authority is questioned. If we are to accept that not all cybercrimes can be controlled via technology and law then other reactive modalities, such as community-led and proximal (online) modes of policing, are required to make the remainder of the ‘matrix’ of cybercrime governance.

**Social control within online communities**

A range of mechanisms built upon the dynamics of online communities are currently being employed to govern online behaviour in order to maintain order. Within Cyberworlds, regulation of deviant behaviour rapidly matured over several years from an oligarchic and vigilante-based system to a formal policing model (see further Williams, 2007b). These changes in regulatory practice within Cyberworlds can be neatly rationalized by using Gill’s (1994) model of societal policing. In the early years of the community, social control was organized around a Community Involvement model, in that the community took responsibility for the policing and introduction of crime prevention measures on a non-structured basis (Gill, 1994). World creators would monitor behaviour when they were online, forming regulations that were specific to their environment. If and when an ‘incident’ occurred community members could either log a complaint with the world creator who would then decide how to deal with the offender, or resort to vigilantism. Implementation of these forms of regulation was ad hoc, and there was no formalized method for dealing with troublesome individuals.
In a response to the increasing sophistication and proliferation of deviant acts, it was recognized that an equivalent organized response had to be developed if the community were to maintain cohesion. As a result the Peacekeepers, the community voluntary police, were established. This was seen by community members as an important landmark in the process of civilizing Cyberworlds, and served to increase the commitment, involvement and attachment of citizens to the community. These developments can be considered a progression to Gill’s (1994) Volunteer Community Policing model where the community provides some limited structure to policing. The Peacekeepers are essentially an organization of volunteers who are formally trained to deal with rule breakers within Cyberworlds. They are a structured organization, which consists of a Peacekeeper Academy for training new recruits, and teams of Peacekeepers who work on a shift basis. A Special Response Team is dedicated to the investigation of complex and serious cases of vandalism and harassment.

However, the effectiveness of these proximal virtual rule enforcers is limited due to the failure of the technology or code of the Cyberworlds architecture and a lack of citizen confidence. The punishment options available to Peacekeepers include ejection from Cyberworlds for a short period, a lockdown of the offender’s account, account closure and contacting the offender’s Internet Service Provider (ISP). While each of these sanctions has a disabling effect upon communication, a troublemaker can easily subvert each one if persistent. Even the most drastic sanction—the closure of account and contacting an ISP—is flawed due to the autonomy granted to all Internet users to reinvent themselves, allowing them to choose a different ISP and open a new account in Cyberworlds under a different identity. These technological shortfalls leave Peacekeepers with a limited arsenal to combat deviant activity. Not only are they ill equipped, they also face high levels of opposition from community members who believe the Peacekeepers are unaccountable and impose justice in an overt crime-control fashion. As a result many community members resort to the shaming tactics utilized in the vigilante style of regulation existent prior to the establishment of the Peacekeepers. This pattern of governance is also common in other online communities, including Secondlife and Ebay, which utilize reputation management systems to shame those who break the rules. It is this topic the article now turns to by drawing upon qualitative data gathered from within Cyberworlds on experiences and perceptions of shame as a regulator.

**The effectiveness of shaming as social control in online communities**

Williams’ (2006) finding that community members prefer the use of vigilante justice, peer pressure and ostracism to maintain order in Cyberworlds finds sympathy in the findings of research conducted in other online communities. Reid (1999) notes how the medieval practice of ‘charivari’ in 13th and 14th-century France has similar characteristics to modes of informal regulation within an online community named JennyMUSH (see later).
Whereas ‘charivari’ involved the public ridicule and physical taunting of an individual who had broken community rules, the mediation of deviance within online environments often involves the use of shaming via textual performances. This process finds resonance in Braithwaite’s (1989) work on shaming. Essentially the virtual offender’s moral conscience is considered to be a far more effective deterrent against further deviance than any formal (proximal or distal) form of governance. Braithwaite’s theory relies on the assumption that to be shamed is undesirable, so undesirable as to prevent someone from committing a crime. Shaming can also work on two levels, it can be applied directly by an audience, be it verbal or tacit communication, such as a frown, or it can be internally applied by the wrongdoer. What is important for both forms of shaming to work is the existence of a conscience. Drawing from learning theories, Braithwaite (1989) contends that we all have a learnt conscience, or classical conditioning, inculcated from early childhood, which tells us the difference between wrong and right, and it is this conscience that allows us to feel the pains of shaming. To go further, Braithwaite (1989) has argued that internal shaming is far more powerful than shaming by onlookers for two reasons. First, shaming oneself has the benefit of immediacy because wrongdoers know instantly that they have done something wrong. For this reason the pains of shaming are amplified. Second, because shaming can be immediate, it has the power of educating the wrongdoer and reinforcing his or her classical conditioning. Here we can see clear similarities with both control and learning theories, and from this we could say Braithwaite’s (1989) theory takes for granted that people are motivated to offend, and have to be pressured socially not to do so. The effectiveness of shaming is then reliant on an individual being bonded to family and wider social structures such as the community.

Braithwaite (1989) devotes considerable attention to distinguishing between shaming that is reintegrative and shaming that is not. Drawing from labelling theory, he states that shaming that is disintegrative can result in stigmatization. Braithwaite continues to note that if shame is applied without forgiveness at a later stage then this may lead to secondary deviance, and possibly abjection in a criminal subculture. It is Braithwaite’s (1989) contention that if the current ‘disintegrative’ shaming process in western societies were to be replaced by ‘reintegrative’ practice, the result would be a reduction in recidivism. The next stage in this process is the reintegration of the wrongdoer back into the community. For reacceptance to work there can be no stigmatization, so the disapproval must be placed on the wrongful act or the behaviour itself and not the individual as a person. This prevents any criminal label being attached to the person holistically. Once the wrongdoer has been shamed, they are forgiven and it is hoped that the severity of the shame would have re-educated them or built upon their classical conditioning which would hopefully prevent recidivism.

**Reintegrative shaming in the online context**

Applying such thinking to an online context, there is little doubt that complex social relationships are played out within many online communities such as
Ebay, Cyberworlds and Secondlife. Markham (1998) and Mnookin’s (1996) account of social interaction online show how individuals developed complex interdependencies, characterized by social and emotional bonds. These online interdependent social networks provide a fertile ground for the use of shaming in Internet deviance reduction. While the constraining effects of the family are absent within these social spaces (unless family members are also present online), the bonds to other ‘netizens’ and online community are as significant as offline ties. For example, within commercial online spaces, such as Ebay and Amazon, good relationships with buyers are essential to forming a solid reputation as a seller. Similarly strong relations within purely social spaces, such as Cyberworlds and Secondlife, are important to sustaining positive reputation and status. Acts of shaming online invariably draw on these ties, attempting to either humiliate the wrongdoer or draw upon their classical conditioning in an attempt to induce feelings of guilt. The following narratives taken from citizens within Cyberworlds demonstrate how the shaming process is effective within online social spaces.

Bola: Were punishments effective? Mostly, the offending party usually left or modified his behaviour. Why? Because the offender was often shamed into modifying his behaviour as he was often publicly ridiculed for his abusive behaviour. As soon as the behaviour stopped, so did the ridicule.

Teamdoyobi: I have been to worlds [outside of Cyberworlds] where harassers have been turned into toads because of their behaviour. Banning them from the community doesn’t work as they can get around the technical blocks. The only way to really make sure they stop is to make them feel small and ashamed.

The process of shaming draws upon bonds to community in attempting to maintain social order within Cyberworlds. As Braithwaite (1989) contends, it is deterrence that is the key to crime reduction—the fear of being humiliated by your peers or being made to feel the pains of guilt. Offenders are made to feel discomfort as a result of internal guilt, triggered by a classical conditioning that has expanded to encompass elements of the online social arena. Community members with attachments and commitments, and who are involved and believe in their online community become subject to the potential pains of shaming, and as such become subjects of social control. Within Cyberworlds many citizens rejected overt crime-control methods espoused by the Peacekeepers, preferring these methods of reintegrative shaming and mediation:

Remy: When I have tried to mediate problems such at this, I first talk to the people involved separately. Then, if I feel that it will help, I get them together and try to help them work through the problem.

Rocketscientist: Now, there is another way that this does work if you exclude the word ‘punishment’ and replace it with ‘rehabilitation’ though that word is too clumsy for what I mean. In other words, treat the offender as a human, determine why they are acting the way they do (often it is merely young teens who crave attention but do not know positive ways of getting attention), and meet them on level ground as an equal. This paragraph does NOT apply
to older persons though. With older persons, it’s often socially acceptable
behavior in their country and not in ours, or, a few are just plain head-cases.

Buxton: Nothing except communicate with them. Get through to them
that first time, before they get mad. It works. I’ve done it, and I have seen it
done far more times than I have seen ejection work. The sad thing is that
Cyberworlds and its volunteers are no longer about communication. Some
of the Peacekeepers don’t begin to have the communication skills needed to
do this even if they wanted to.

Rocketscientist adopts an amateur therapeutic approach to punishment.
Instead of ejecting or negatively shaming the wrongdoer, a more liberal neo-
rehabilitative approach is taken. Challenging the deviant’s behaviour and
addressing the ‘root’ of the problem forms part of this online therapeutic
discourse. However, a distinction is drawn in relation to eligibility.
Rocketscientist and others only consider this process suitable for ‘young
teens’ who ‘don’t know better’. The rationalization of deviant behaviour
exhibited by ‘adults’ is less sympathetic. Deviant adults are either con-
sidered culturally ‘alien’ or suffering from a level of sociopathy. This distinc-
tion seems to work well in practice as it is believed the majority of those
expressing deviant behaviour within Cyberworlds are indeed adolescent.
Continuing the therapeutic theme Buxton emphasizes the value and effect-
iveness of communication. Condemning the Peacekeepers, Buxton has rec-
ognized how a crime-control model is being implemented at the cost of
more therapeutic and due process methods. Both Remy’s and Billgatessux’s
narratives of how they would deal with an online harasser illustrate how
the zero tolerance policing employed by the Peacekeepers is considered a
last resort, maintaining the therapeutic rhetoric (Williams, 2006). There is
a clear desire among members of online communities for the maintenance
of vigilante styles of regulation. The relative effectiveness of shaming in
comparison to techno-mediated punishments is more than evident to com-


Disintegrative shaming in the online context

While the practice of shaming within Cyberworlds was based upon a ther-
apeutic instead of punitive discourse, evidence from other online commu-
nities indicates this may be an isolated phenomenon. A common example
of ritual shaming can be seen in the practice of ‘toading’, often found in Multi-
User Domains. The process usually involves the system administrator,
or the person at the peak of the community hierarchical scale, altering the appearance and/or description of the offender’s persona into something shameful (commonly a toad, which can be traced back to fantasy gaming and the Dungeons and Dragons role-playing genre). A process of public ridicule then begins with the victims and other sympathetic community members venting their anger upon the offender via derisory speech. This way of controlling deviant behaviour is well established in many online communities and discussion groups (Reid, 1999). Examples of shaming have been seen in the mediation of ‘virtual rape’. After the failed efforts of technological mediation (e.g. exclusion from the community) in the case of Mr Bungle, textual methods were employed to disempower his actions (Dibbell, 1999). These textual performances involved the community ‘shunning’ the assailant and ultimately ‘shaming’ him. These acts, which were textually and socially performative, secured the social ostracism of the offender. Similar instances of vigilante justice have emerged in other cases of sexual harassment online. Reid (1999) recounts an instance of ‘virtual rape’ in the Multi-User Domain named JennyMUSH, an environment created for the purpose of counselling women who had suffered offline sexual abuse. The harasser taunted and verbally abused several of the online visitors while performing acts of sexual violence through text. Given the nature of the environment the assailant’s actions were considered exceptionally harmful. Members of the community were encouraged to harass textually the assailant in an attempt to deliver punishment.

The outcome of the shaming process in these examples was to ostracize the wrongdoer permanently. While some online communities have adopted external disintegrative tactics in order to shame wrongdoers (Reid, 1999), the earlier extracts show how much of the shaming within Cyberworlds draws upon internal classical conditioning, in an attempt to induce guilt and to reintegrate the offender. Part of the reason for this disparity in approach may be technological. In both the JennyMUSH and Mr Bungle examples the environments were textually mediated, allowing individuals to alter the offender’s appearance via their textual performances. The graphical component of Cyberworlds does not allow for such flexibility. Vigilante groups cannot change the appearance of a wrongdoer, making external disintegrative shaming more difficult to achieve.

A further explanation for variance in approaches to shaming relates to levels of interdependency. Not all online environments are characterized by complex social networks. As noted earlier, Braithwaite (1989) contends that internal shaming is more effective at reducing deviance than is shame applied by onlookers. The extent to which this can be said to be an accurate portrayal of the shaming process in all online environments is dependent upon the social embeddedness of its members. The level and complexity of interdependencies within online environments vary and are more likely to be less multi-farious than those existent offline. For example, while complex social and emotional interdependencies may exist online, the more significant relationships evident in family networks are absent. The ephemeral nature
of some online relationships has undoubtedly resulted in the use of externally imposed disintegrative shaming. Humiliation is a more effective punishment if the individual has no complex set of online interdependencies. In these circumstances any appeal to the individual’s classical conditioning would fail to bring about the desired shame due to a lack of local (online) socialization. The more draconian rituals of public ridicule then become the only effective way of shaming these less bonded individuals. This leaves the question of public ridicule as punishment. In the offline world the mode of punishment as spectacle (Foucault, 1979) was successful in the 18th and 19th century due to the offender’s constant residence among their community peers, allowing the shaming to have a sustained effect. Yet shaming in a virtual environment is short lived. The purpose of public ridicule must then lie in a form of justice based on the alleviation of the feelings of those harmed. The chief concern is to protect the community’s integrity and to expel anyone that threatens its solidarity, while simultaneously repairing the harm done via a process of retribution. However, this process of justice may result in the opposite effect; by encouraging community members to taunt and abuse other members, hatred and derisory performances are given a free and legitimate reign, encouraging a lack of trust and interdependence. Without trust and interdependency a community becomes fragmented when members grow ever more anxious over the possibilities of victimization and the eventual demise of their online environment.

The application of shaming within online communities is complicated by the differing levels of social bond to the immediate online context. Where individuals lack complex online interdependencies it is unlikely that reintegrative shaming would be effective in reducing recidivism. Alternative disintegrative shaming tactics may be more effective in humiliating the offender, but the punishment is temporary given the ephemeral and anonymous nature of the online context. In relation to Cyberworlds, vigilante modes of regulation that rely on reintegrative shaming are marginally effective when applied to bonded individuals. However, such marginal effectiveness cannot justify the use of vigilante methods alone. It seems most beneficial to combine both proximal (online) methods of vigilantism and organized policing, relying on techno-mediated forms of punishment where social forms fall short. The uncertain effectiveness of proximal nodes of governance within online communities to deal with cybercrime drives us to evaluate the potential role of the public police in virtual order maintenance.

A role for the public police?

Although the public police and their constitutional position has changed considerably since their formation, many of the original Peelian police principles survive, though adapted to modernity: a bureaucratically organized responsive local police that maintains order and enforces law; officers who
are identifiable from the rest of the public, professional in conduct, accountable to law and the community for their actions. However, the increasing pervasiveness of the Internet and the cyberspace it creates, along with its global, transformative impacts create a range of entirely new challenges for public police which question their traditional local dominance over the security domain and could in fact marginalize them completely. Not only does the concept of cybercrime produce problems for the police because Internet-related offending takes place within a global context whereas crime tends to be nationally defined with policing decisions made locally, but policing the Internet is also a complex affair by the very nature of policing and security being networked and nodal (Johnston and Shearing, 2003). While the application of concepts of networked and nodal security may be challenged in the ‘terrestrial world’ (Crawford and Lister, 2004: 426), nowhere is it more networked and nodal than in cyberspace (see further Wall, 2007b: 184).

The public police role has to be understood within the broader and largely informal architecture of Internet policing, which not only enforces norms and laws but also maintains order in very different ways. Understanding this position enables more realistic expectations and understanding of the public police role. It also helps to identify a broader range of cross-jurisdictional and cross-sectoral issues that the police have to attend to in order to participate fully in policing the Internet, by fully embracing both the concept of networking and the subsequent network technology. This growing networking of sources of security (which includes police as a node) during recent decades (Johnston and Shearing, 2003; Dupont, 2004; Shearing, 2004) has arisen as one part of the shift towards the networked society (Castells, 2000). To further this trend a range of new transformations need to take place in order to enhance the effectiveness and legitimacy of the nodal architecture—a flattening of policing structures, parity of legal definitions across boundaries, broadly accepted frameworks of accountability to the public, shared values, multi-agency and cross-sectoral dialogues, and more.

The current limited role of the police, especially in relation to low-level reassurance policing within online communities, means a continuum of regulation is required (Bonnici and Cannataci, 2003) from order maintenance to law enforcement. As Grabosky and Smith note, ‘much computer-related illegality lies beyond the capacity of contemporary law enforcement ... security in cyberspace will depend on the efforts of a wide range of institutions’ (2001: 8). While it is not realistic to ask for immunity from state regulation, there are clear regions where a sovereign need not become involved. A realistic expectation is that state intervention is likely to remain where economic issues or public concerns are strong motivators. In all other cases, where the cost of intervening is likely to be non-commensurate to the outcome, a sovereign will more than likely devolve regulation to the online community members themselves (Bonnici and Cannataci, 2003). This general line of thinking is shared by Grabosky and Smith (2001) who perceive that the resource constraints
upon governments will force them to enlist the assistance of the private sector and the ‘community’.

Conclusions

Online behaviour has become an increasingly complex matter. As we enter the information age proper, online communities are no longer simply an escape from the responsibilities of the first life, they are becoming meaningful extensions of it across a broader span of time and space. They have become the third space for the ‘second life’, but a space that is now very connected to their physical life. The evidence of this transformation is all around us: millions participate in online communities of interest; online auctions are shaping consumer behaviour; emotional relationships are being forged online before the partners actually meet in person; work at home is becoming more frequent; online relationships with government with regard to tax; advice; various forms of registration and information sharing are now commonplace. It therefore follows that if online behaviour has become a complex matter, then so has its governance.

In developing models for governing online behaviour it is useful to work with that which already exists. By comparison with early dystopic prophecies from both left and right, the Internet has become neither anarchic nor an Orwellian nightmare, rather, it has become remarkably ordered. This sense of order results from a complex ‘assemblage’ of networked nodes of security that continually shape virtual behaviour (see Wall, 2007a: 167–77). The term ‘assemblage’ is particularly useful here because, without attributing causality, it describes the relationship between heterogeneous contributors to governance that work together as a ‘networked’ and functional entity, but do not necessarily have any other unity (see Miller and Rose, 1990; Haggerty and Ericson, 2000: 605).

When seeking to understand the relationships between those who provide governance and how they do so (their auspices—Shearing, 2004), it is useful to do so in terms of the proximal and distal qualities outlined earlier. Especially as we find their reflection in the online replication of the bifurcation of the broader functions of terrestrial policing—the need to maintain order online and the need to enforce the law (in the latter case the formal rules of the online community). By separating out the order maintenance and law enforcement functions, some sense can also be made of the rather conflicting messages that are emerging in debates over policing the Internet where the ‘cybercrime’ problem is perceived either as a problem of weak law or poor technological control. In fact, online communities employ a form of online community policing through norm maintenance which tends to exploit the ‘natural surveillance’ of networked technologies to facilitate both primary and secondary social control functions, while also mediating, where they arise, any disparities arising from national or jurisdictional legal differences in definition (Wall, 2007b).
What this article has demonstrated is that online communities are very real places inhabited by very real people who want, as they do in their terrestrial world, to be reassured of their safety while online. As the list of online activities continues to grow longer, deeper and more diverse, then the need to maintain order will continue to be a priority as will the debates over how we police the Internet. This is especially the case if its ability to embrace diversity is to be maintained meaningfully—after all this was one of the main qualities that made the Internet popular in the first place.

Notes
1 Authors listed in alphabetical order.
2 Name of community changed to preserve anonymity.
3 http://www.entropiauniverse.com

References


DAVID S. WALL is Professor of Criminal Justice and Head of the School of Law at the University of Leeds. As a member of the Centre for Criminal Justice Studies, of which he was previously Director, he conducts research and teaches in the fields of criminal justice and information technology, policing and cyberlaw. His latest publication is Cybercrime: The Transformation of Crime in the Information Age (Polity, 2007).

MATTHEW WILLIAMS is a lecturer in Criminology and Criminal Justice at the School of Social Sciences, Cardiff University. His research main areas are cybercrime, digital & online qualitative methodology and policing, diversity & equal opportunities. He has recently published Virtually Criminal: Crime, Deviance and Regulation Online (Routledge, 2006).