CHAPTER 16

Self-disclosure, privacy and the Internet

Adam N. Joinson and Carina B. Paine

In this chapter, we examine the extant research literature on self-disclosure and the Internet, in particular by focusing on disclosure in computer-mediated communication (CMC) and web-based forms – both to surveys and in e-commerce applications. We also consider the links between privacy and self-disclosure, and the unique challenges (and opportunities) that the Internet poses for the protection of privacy. Finally, we propose three critical issues that unite the ways in which we can best understand the links between privacy, self-disclosure and new technology: trust and vulnerability, costs and benefits and control over personal information.

Central to the chapter is the notion that self-disclosure is not simply the outcome of a communication encounter: rather, it is both a product and process of interaction, as well as a way for regulating interaction dynamically. We propose that by adopting a privacy approach to understanding disclosure online, it becomes possible to consider not only media effects that encourage disclosure, but also the wider context and implications of such communicative behaviours.

What is self-disclosure?

Self-disclosure is the telling of the previously unknown so that it becomes shared knowledge, the 'process of making the self known to others' (Jourard and Lasakow 1958: 91). This shared knowledge might exist between pairs of people, within groups, or between an individual and an organization. It has a variety of purposes, in part dependent on the context in which disclosure occurs. For instance, within dyads, particularly romantic relationships, it serves to increase mutual understanding (Laurenceau *et al.* 1998), and builds trust by making the discloser increasingly vulnerable (emotionally or otherwise) to the other person (Rubin 1975). Since self-disclosure is often reciprocated it frequently serves to strengthen the ties that bind people in romantic or friendship-based relationships (Jourard 1971).

Disclosure within groups can serve to enhance the bonds of trust between group members, but it can also serve to legitimize group membership and strengthen group identity. For instance, the admission of a negative identity (e.g. 'I am an alcoholic') within a shared identity group serves both to increase trust by revealing a stigmatized identity and act as a membership card for a particular group (Galegher *et al.* 1998). Personal growth may be an outcome of honest self-disclosure (Jourard 1971). In a study reported by Pennebaker *et al.* (1988), participants assigned to a trauma-writing condition (where they wrote about a traumatic and upsetting experience for four days) showed immune system benefits, compared to a non-trauma

writing group. Disclosure in this form has also been associated with reduced visits to medical centres and psychological benefits in the form of improved affective states (Smyth 1998). For people using the Internet to talk about their problems (or to publish weblogs), their activities may well have unforeseen, positive, health and psychological benefits.

Finally, disclosure between an individual and an organization can serve authentication purposes – for instance, to establish identity, allow authentication of a claim to identity and to enable an organization to recognize you in the future in order to personalize its offerings to you. Organizations might also ask for personal information for marketing purposes – for instance, when registering to access a website or joining an online community. Of course, organizations, in the form of researchers, might also ask for personal information in the name of academic research.

New technology, and in particular the Internet, might well change the demands upon people to disclose personal information, as well as the possible implications of such disclosure. For instance, disclosing personal information to another person online might not involve the increased vulnerability that usually follows self-disclosure of personal information offline (Ben-Ze′ev 2003). Organizations might also demand more information in the name of authentication (although this need not always be personal information). Furthermore, new technology changes the scope of personal information that can be disclosed or collected. For instance, the development of ambient and ubiquitous devices, such as smart mobile phones and RFID tags, makes it likely that information about location, movements and social interactions are likely to be collected in the future in some form. How we negotiate the disclosure of such information is a critical issue, equally as important as how systems are designed to minimize privacy violations while also providing adequate levels of functionality.

Measuring self-disclosure

Within person-to-person and person to group interactions, self-disclosure has tended to be studied using either content analysis or self-report. In the case of content analysis, the issue of what constitutes self-disclosure, and how it is scored, is particularly important. One option would be to count the number of instances within a conversation in which a person discloses information about themselves. However, there are a number of problems with this approach. First, it is not always clear what constitutes an act of self-disclosure – for instance, to express an opinion may well be classified in some contexts, but not in others. Second, self-disclosure can only be properly understood in terms of the ongoing interaction. For instance, does one count answers to a specific question – 'How old are you?' – as self-disclosure, or only spontaneous occurrences of disclosure (see Antaki *et al.* [2005] for a recent discussion of this issue). Moreover, given the dynamics of reciprocity, it may not even be possible to count occurrences of spontaneous disclosure as independent of the conversational dynamic. For these reasons, it is usual to treat discussions between people as a single unit of analysis (Kenny and Judd 1986).

Finally, not all self-disclosure is equal – disclosing your season of birth is not the same as disclosing your age, which is not the same as disclosing your sexual fantasies. One option is to use a three-layer categorization scheme proposed by Altman and Taylor (1973) to guide the content analysis of depth. Altman and Taylor suggest that disclosure can be categorized into either peripheral, intermediate, and core layers. The peripheral layer is concerned with biographic data (e.g. age), the intermediate layer with attitudes, values and opinions and the core layer with personal beliefs, needs, fears, and values. Joinson (2001b) instead used a 7-point Likert scale with which two scorers allocated the degree to which an utterance 'revealed vulnerability'. However, Antaki *et al.* (2005) argue that the act of disclosure needs to take into account the interactional context rather than simply being scored on a checklist. For instance, the phrase 'I'm the world's worst cook' could be disclosure, a plea for help or self-deprecation. Without the context, they argue, it is not possible to be certain.

Alternatively, lists of topics can be used to score intimacy – although again there are a number of problems with their application in practise to communication research (see Tidwell and Walther 2002, footnotes).

Self-report measures of disclosure have been used successfully, for instance to compare levels of disclosure in face-to-face (FtF) and online relationships, or to link marital satisfaction with disclosure within the relationship. For instance, Parks and Floyd (1996) asked their participants to report the level of self-disclosure in their Internet relationships using self-report (e.g. high scores on 'I usually tell this person exactly how I feel' and low scores on 'I would never tell this person anything intimate or personal about myself'). However, the same problems – a lack of context – arise for such self-report measures too.

Measures of dispositional self-disclosure can also be used. For instance, within the International Personality Item Pool (IPIP) the RD3 subscale of items 'similar to the Temperament and Character Inventory (TCI)' has 10 items such as 'Am open about myself to others' (positive coding) and, 'Reveal little about myself' (negative coding) to measure general self-disclosure. However, it is not currently clear how such personality type measures might interact with different media, or indeed with people's behaviour within a specific interaction.

Self-disclosure outside of person-to-person and group interactions can also be measured in a number of different ways. One system is to count the number of words typed into text boxes in response to a personal or sensitive question, and to rate those responses by their intimacy or depth (e.g. Moon 2000; Joinson 2001b). Joinson (2005) also describes the use of non-response as a measure of self-disclosure in studies. There are two main ways in which non-response can be operationalized in survey methodology and e-commerce. The first is non-response – either submitting a default selection, or where there is no default option, submitting no response. A second is to add an option that allows participants to select 'I prefer not to answer' (Buchanan *et al.* 2002; Knapp and Kirk 2003). The use of 'I prefer not to answer' as a response option to a sensitive question is methodologically similar to the provision of a 'no opinion' response in attitudinal surveys. While it has been argued that the provision of 'no opinion' choices may increase satisficing in attitude surveys (Holbrook *et al.* 2003), there is little reason to assume that a similar process would operate in the use of 'I prefer not to answer' responses to sensitive personal questions. Indeed, Joinson *et al.* (in press) report that the provision of 'I prefer not to answer' options in a salary question may improve data quality by reducing the number of non-responses or default selections. In our own research (in preparation) we established that people are more likely to use an 'I prefer not to say' option when faced with a sensitive rather than non-sensitive question, and that priming participants for online privacy (by asking them about their privacy concerns and behaviours) significantly increases the use of 'I prefer not say' as an option to sensitive questions.

Finally, self-disclosure can be measured using statistical techniques, for example the randomized response technique (Musch *et al.* 2001). In the randomized response technique, participants are asked to answer a sensitive question either truthfully or with a prespecified answer, depending on the result of a random event such as a coin toss. So, for instance, the question might be, 'do you lie to your partner about anything important?' Participants are asked to toss a coin, and if it is heads, they tell the truth, if it is tails they say 'yes' regardless of the truthful answer. Using statistical probabilities, a population estimate for a behaviour can be found, without knowing if any one individual told the truth or simply followed the instructions for 'tails'.

As noted earlier, self-disclosure in the age of ubiquitous computing poses novel challenges. For instance, it is likely that people will disclose information without full awareness or control (e.g. their location via a cell phone) – instead they may need to rely on privacy profiles or preferences to negotiate the disclosure on their behalf. In these circumstances, discussion or measurement of a single instance of disclosure is meaningless without full consideration of the context in which disclosure occurred.

Self-disclosure and the Internet

A rapidly increasing body of experimental and anecdotal evidence suggests that CMC and general Internet-based behaviour can be characterized as containing high levels of self-disclosure. For instance, Rheingold (1993) claims that new, meaningful relationships can be formed in

cyberspace because of, not despite, its limitations. He further argues that 'the medium will, by its nature . . . be a place where people often end up revealing themselves far more intimately than they would be inclined to do without the intermediation of screens and pseudonyms'. Similarly, Wallace (1999) argues that 'The tendency to disclose more to a computer . . . is an important ingredient of what seems to be happening on the Internet' (1999: 151). Self-disclosure has been studied in a number of different settings using computers. For instance, Parks and Floyd (1996) studied the relationships formed by Internet users. They found that people report disclosing significantly more in their Internet relationships compared to their real life relationships. Similarly, in their study of 'coming out on the Internet', McKenna and Bargh (1998) argue that participation in online newsgroups gives people the benefit of 'disclosing a long secret part of one's self' (1998: 682). Chesney (2005), in a small-scale study of online diaries, reported high levels of disclosure of sensitive information, with half of his participants claiming to never withhold information from their diaries.

In the series of studies reported by Joinson (2001a), the level of self-disclosure measured using content analysis of transcripts of FtF and synchronous CMC discussions (Study one), and in conditions of visual anonymity and video links during CMC (Study two). In keeping with the predicted effect, self-disclosure was significantly higher when participants discussed using a CMC system as opposed to FtF.

In the second study, incorporating a video link while the participants discussed using the CMC program led to levels of self-disclosure similar to the FtF levels, while the comparison condition (no video link) led to significantly higher levels of self-disclosure.

These two studies together provide empirical confirmation that visually anonymous CMC tends to lead to higher levels of self-disclosure. The results of these studies also suggest that high levels of self-disclosure can effectively be designed out of an Internet interaction (e.g. through the use of a video link or accountability cues (Joinson 2001a, Study 3), as well as encouraged.

Further empirical confirmation of increased self-disclosure during CMC comes from the work of Tidwell and Walther (2002). They proposed that heightened self-disclosure during CMC may be due to people's motivation to reduce uncertainty. According to Uncertainty Reduction Theory (URT) (Berger and Calabrese 1975), people are motivated to reduce uncertainty in an interaction to increase predictability. In FtF interaction, uncertainty can be reduced through both verbal and non-verbal communication and cues. Tidwell and Walther hypothesize that during CMC, uncertainty reducing behaviours are text-based only, including increased levels of self-disclosure and question asking. To test this, Tidwell and Walther recruited 158 students to discuss in opposite sex pairs with an unknown partner using a CMC system or FtF. The subsequent conversations were content-analysed for disclosure using the breadth and depth indices developed by Altman and Taylor (1973; see above for a description).

Tidwell and Walther found that those in the CMC condition displayed higher levels of both question asking and self-disclosure compared the FtF condition. The questions asked by CMC discussants were also more probing and intimate than those asked by those talking FtF, while both the questions and disclosure by FtF interactants tended to be more peripheral than those in the CMC condition. Tidwell and Walther conclude that the limitations of CMC encourage people to adapt their uncertainty-reducing behaviours – they skip the usual asking of peripheral questions and minor disclosure, and instead opt for more direct, intimate questioning and self-disclosure.

Surveys and research administered via the Internet, rather than using paper methodologies, have also been associated with reductions in socially desirable responding (Joinson 1999; Frick *et al.* 2001), higher levels of self-disclosure (Weisband and Kiesler 1996) and an increased willingness to answer sensitive questions (see Tourangeau 2004).

In a similar vein, survey methodology techniques that tend to reduce human involvement in question administration also increase responses to sensitive personal questions. For instance, compared to other research methods, when data collection is conducted via computer-aided self-interviews (where participants type their answers on to a laptop) people report more health-related problems (Epstein *et al.* 2001), more HIV risk

behaviours (Des Jarlais *et al.* 1999), more drug use (Lessler *et al.* 2000), and men report less sexual partners, and women more (Tourangeau and Smith 1996). Medical patients tend to report more symptoms and undesirable behaviours when interviewed by computer rather than FtF (Greist *et al.* 1973). Clients at a STD clinic report more sexual partners, more previous visits and more symptoms to a computer than to a doctor (Robinson and West 1992). Ferriter (1993) found that pre-clinical psychiatric interviews conducted using CMC compared to FtF yielded more honest, candid answers. Similarly, automated or computerized telephone interviews, compared to other forms of telephone interviewing, lead to higher levels of reporting of sensitive information (see Lau *et al.* 2003; Tourangeau 2004).

Conversely, methods that increase the social presence of the surveyor (e.g. by using photographs of the researcher) have been predicted to lead to a reduced willingness to answer sensitive questions (Tourangeau *et al.* 2003), although the findings of Tourangeau *et al.* were equivocal. However, Sproull *et al.* (1996) found that participants 'present themselves in a more positive light to the talking-face displays' (1996: 116) than to text-only interfaces. Joinson *et al.* (in press) report that although personalizing the research experience leads to higher response rates to a self-administered survey, it also reduces self-disclosure.

Within the Human–Computer Interaction (HCI) literature, the assumption seems to be that people will avoid disclosing information to commercial web services (Metzger 2004) due to their privacy concerns (Jupiter Research 2002). An online survey stated that the three biggest consumer concerns in the area of online personal information security were: companies trading personal data without permission, the consequences of insecure transactions, and theft of personal data (Harris Interactive 2002). For example, Hoffman *et al.* (1999) report that almost 95% of Internet users declined to provide personal information when requested to do so by a website, and over 40% provided false demographic information when requested. Quittner (1997) reports that 41% of survey respondents would rather exit a web page than reveal personal information. Clearly then, open self-disclosure is not a universal experience on the Internet: for commercial organizations, consumers are often less than forthcoming, usually because of a combination of privacy concerns, lack of trust and concern about how personal information will be used (Hoffman *et al.* 1999; Metzger 2004). For instance, Olivero (2001) studied the willingness to disclose information about the self to a commercial organization, and manipulated the level of trustworthiness of the organization, whether a financial reward was offered for disclosure and the level of intrusiveness of the questions. She found that the level of trust was associated with participants' willingness to disclose to highly intrusive questions, but that an awareness of data mining/privacy concerns moderated this effect of trust. Andrade *et al.* (2002) conducted a similar study by examining three approaches to encourage self-disclosure of personal information online – the completeness of a privacy policy, the reputation of a company and the offer of a reward. They found that the completeness of privacy policy and reputation of the company reduce the level of concern over self-disclosure while the offer of a reward heightens concern.

However, there are a number of 'counter surveys' and empirical evidence suggesting that there is a significant discrepancy between privacy principles and privacy practices. Very few individuals actually take any action to protect their personal information, even when doing so involves limited costs (Berendt *et al.* 2005; Jenson *et al.* 2005) i.e. there is a dichotomy between stated attitudes and actual behaviours of people in terms of their protection of personal information.

Models of self-disclosure online

Explanations for high levels of self-disclosure in person-to-person CMC have tended to focus on the psychological effects of anonymity: 'This anonymity allows the persecuted, the controversial, and the simply embarrassed to seek information – and disseminate it – while maintaining their privacy and reputations in both cyberspace and the material world' (Sobel 2000: 1522).

Theoretically, it has been argued that anonymity in CMC works by replicating a 'strangers on the train' experience (Bargh *et al.* 2002), promoting

private self-awareness and reducing accountability concerns (Joinson 2001a), creating a need for uncertainty reduction (Tidwell and Walther 2002) or a combination of the media and the process of interaction itself (Walther 1996).

Similarly, explanations for increased self-disclosure to online surveys and web forms have also tended to stress anonymity (Joinson 1999), alongside the reduced social presence (and judgement) of the researcher (Tourangeau 2004), reduced vulnerability (Moon 1998) and increased privacy of the research environment (Tourangeau 2004). Once privacy is reduced, or social presence increased, self-disclosure also tends to be reduced (Joinson *et al.* in press).

However, explanations for people's unwillingness to disclose personal information to e-commerce services invariably stress people's privacy concerns (e.g. Hoffman *et al.* 1999), in particular, issues surrounding the level and type of information collected, and people's lack of knowledge about how it may be used in the future, or control over that use (Metzger 2004).

These differing approaches to understanding disclosure and non-disclosure of personal information illustrate the paradox of self-disclosure on the Internet. On the one hand, the Internet provides an environment in which people can express themselves with relative immunity via pseudonyms, but in order to access these services and sites they often need to disclose high levels of personal information during the registration process.

Within the privacy literature, this paradox is relatively easy to solve – the provision of information about the self is treated quite separately from the use of privacy or pseudonymity to express one's inner desires. However, it is rare for CMC self-disclosure research to explicitly consider privacy, in particular the multifactor approaches to privacy discussed in the socio-legal literature.

Within e-commerce, there are further paradoxes which may be solved by looking at both the literature on interactional person-to-person disclosure and the privacy literature concurrently. For instance, there are occasions when you need to disclose a lot of personal information (e.g. purchasing online), but other factors (e.g. lack of social presence) make such privacy concerns less pressing. The answer to this paradox is that it is the author to whom one is disclosing that is critical – if one trusts the recipient of the personal information, then one can act with relative freedom in the pseudonymous world such disclosure purchases. Only by considering the wider context can such seemingly paradoxical impacts of new technology on personal disclosure be fully understood.

This interpretation also strongly suggests that any explanation of self-disclosure online that relies solely on media effects (i.e. visual anonymity) is mistaken. Disclosure, while often 'given away' is also something that is carefully considered within the context of an ongoing interaction and wider context – regardless of whether that interaction is interpersonal or human–computer. We would suggest that a wider theoretical scope is needed – not only is it important to consider the particular context of an interaction, but also how the person accessed that environment in the first place. For instance, while the use of pseudonyms may enable expressive freedom on a discussion board, we would also ask how access was gained to the board, what registration process was in place, what records of postings are kept remotely and locally and so on? Without this knowledge, one is forced to assume that people somehow dropped into an online environment out of the sky, rather than as a motivated act (see Joinson 2003).

While concern about the privacy implications of new technology are nothing new (Home Office 1972), the development and linking of databases with biometrics, and the tension between the need for identification, protection of privacy and full participation in the e-society (Raab *et al.* 1996) makes an understanding of the relations between privacy and the disclosure and use of personal information critical. In the next section of this chapter, we consider what privacy is, how the Internet and new technologies threaten privacy, and the implications of privacy for understanding self-disclosure within an interaction.

What is privacy?

There have been many attempts at definitions of privacy. In a legal context, privacy is largely synonymous with a 'right to be let alone' (Warren and Brandeis 1890). However, others have argued

that privacy is only the right to prevent the disclosure of personal information. Many researchers have referred to the difficulties involved in trying to produce a definition (e.g. Burgoon *et al.* 1989) and despite various attempts to create a synthesis of existing literature (e.g. Parent 1983; Schoeman 1984) a unified and simple account of privacy has yet to emerge. Despite there being no unitary concept of privacy it is clear that both individuals, and society, attach a level of importance to privacy. For example, Ingham states that 'man, we are repeatedly told is a social animal, and yet he constantly seeks to achieve a state of privacy' (1978: 45).

Within psychological literature both Westin's and Altman's theories figure prominently in the major reviews of privacy in the 1970s. Westin (1967: 7) provides a link between secrecy and privacy and defines privacy as 'the claim of individuals, groups, or institutions to determine for themselves when, how and to what extent information about them is communicated to others'. At the psychological level, Westin states that privacy provides opportunities for self-assessment and experimentation and therefore the development of individuality. Specifically, Westin (1967) proposes four main functions of privacy:

1. *personal autonomy* applies to the need for the development of individuality and the avoidance of manipulation by others;
2. *emotional release* refers to the need for opportunities to relax and escape from the tensions of everyday life in order to support healthy functioning;
3. *self-evaluation* is the application of individuality onto events and the integration of experience into meaningful patterns, and
4. *limited and protected communication* refers to both the sharing of personal information with trusted others and the setting of interpersonal boundaries.

Altman (1975) incorporates both social and environmental psychology in understanding the nature of privacy. He defines privacy as 'the selective control of access to the self' (p. 24) and believes privacy is achieved through the regulation of social interaction, which can in turn provide us with feedback on our ability to deal with the world, and ultimately affect our definition of self. Although Westin and Altman emphasize different aspects of privacy, there are many similarities between the theories. For example, both theories are examples of 'limited-access' approaches to privacy in that they both emphasize controlling or regulating access to the self. Such commonalties between these theories suggest that 'their ideas provide a reasonable foundation for understanding the fundamentals of privacy as a psychological concept' (Margulis 2003: 424). A large amount of work has supported and extended both Westin's (e.g. Marshall 1974; Pederson 1979) and Altman's work (e.g. Kupritz 2000) and as such both of their theories have stimulated much of the research and theory development of privacy.

Since these earlier definitions, the highly complex nature of privacy has resulted in an alternative way of defining it – through its various dimensions. Burgoon *et al.* (1989) distinguish four dimensions of privacy and define it using these dimensions as 'the ability to control and limit physical, interactional, psychological and informational access to the self or one's group' (Burgoon *et al.* 1989: 132). Each of the dimensions they distinguish is briefly described below with some examples.

1. *The physical dimension* Physical privacy is the degree to which a person is physically accessible to others. This dimension is grounded within the human biological need for personal space. Examples of violations to physical privacy include: surveillance, entry into personal space and physical contact.
2. *The interactional dimension* Interactional (or social/communicational) privacy is an individual's ability and effort to control social contacts (Altman 1975). Burgoon *et al.* (1989) summarizes the elements of this dimension as control of the participants of, the frequency of, the length of and the content of an interaction. Non-verbal examples of violations to social privacy include close conversational distance and public displays of affection. Verbal examples include violations of conversational norms (e.g. commenting on mood or appearance) and initiating unwanted conversation.
3. *The psychological dimension* Psychological privacy concerns the ability of human beings to control cognitive and affective inputs and outputs, to form values, and the right to

determine with whom and under what circumstances thoughts will be shared or intimate information revealed. As such, psychological privacy can either develop or limit human growth. Examples of violations to psychological privacy include psychological assaults through name-calling and persuasion.

4. *The informational dimension* Informational privacy relates to an individual's right to determine how, when, and to what extent information about the self will be released to another person (Westin 1967) or to an organization. According to Burgoon *et al.* (1989), this dimension is closely related to *psychological privacy*: however, the control differs from the individual self-disclosure associated with psychological privacy because it is partly governed by law/custom and as it often extends beyond personal control. Examples of violations to informational privacy include going through another person's mail and sharing personal information with others.

DeCew (1997) also reflects the multidimensional nature of privacy in her definition: however, she distinguishes only three dimensions:

1. *The informational dimension* Informational privacy covers personal information such as finances, medical details and so on that an individual can decide who has access to and for what purposes. If disclosed, this information should be protected by any recipients of it. By protecting informational privacy individuals avoid invasions (or potential invasions) to their privacy.
2. *The accessibility dimension* Accessibility privacy refers to physical or sensory access to a person. It 'allows individuals to control decisions about who has physical access to their persons through sense perception, observation, or bodily contact' (DeCew 1997: 76–7).
3. *The expressive dimension* Expressive privacy 'protects a realm for expressing one's self-identity or personhood through speech or activity. It protects the ability to decide to continue or to modify ones behaviour when the activity in question helps define oneself as a person, shielded from interference, pressure and coercion from government or from other individuals' (DeCew 1997: 77). As such, internal control over self-expression and the ability to build interpersonal relationships improves, while external social control over lifestyle choices and so on are restricted (Schoeman 1992).

Using these multidimensional approaches to define privacy results in some overlap of the features between each dimension. For example, within Burgoon *et al.*'s dimensions some features of informational privacy overlap with psychological privacy, and some features of social privacy overlap with physical privacy. Within DeCew's dimensions there is some overlap between accessibility and informational privacy, and expressive privacy is conceptually linked with both of these dimensions. In addition, there is also some overlap between Burgoon *et al.*'s and DeCew's dimensions. For example, the informational dimension appears in both definitions and Burgoon *et al.*'s physical and social dimensions appear to map onto DeCew's accessibility and expressive dimensions respectively.

Of direct relevance to this chapter are the dimensions of informational and expressive privacy. Central to these dimensions are the desire to keep personal information out of the hands of others, or in other words *privacy concern* (Westin 1967), and the ability to connect with others without interference. In a systematic discussion of the different notions of privacy, Introna and Pouloudi (1999) developed a framework of principles that explored the interrelations of interests and values for various stakeholders where privacy concerns have risen. In this context, concern for privacy is a subjective measure – one that varies from individual to individual based on that person's own perceptions and values. In other words, different people have different levels of concern about their own privacy.

One scheme for categorizing the different levels of privacy concerns is the Westin privacy segmentation (Harris and Associates Inc. and Westin 1998). The Harris Poll is a privacy survey conducted by telephone across the United States among approximately 1,000 people. This survey has been conducted regularly since 1995 and divides respondents into one of three categories depending on their answers to three statements. The three categories of respondents are: *Privacy Fundamentalists* who view privacy as an especially high value which they feel very strongly about.

Currently about a quarter (35%) of all adults are privacy fundamentalists (Computerworld 2005); *Privacy Pragmatists* also have strong feelings about privacy. They weigh the value to them and society of providing their personal information. Currently around approximately 55% of all adults are privacy pragmatists (Computerworld 2005); *Privacy Unconcerned* who have no real concerns about privacy. Approximately 10% of all adults are privacy unconcerned (Computerworld 2005).

Although levels of concern may differ between people, a failure to achieve any level privacy will result in 'costs'. For example, by not obtaining privacy a person will not benefit from the opportunities that the functions of privacy provide – which could results in stress or negative feedback about the self. There are also costs of losing privacy either through privacy invasion – when conditions for privacy are not achieved, for example being overheard – or privacy violation when recipients of personal information, intentionally provided by the discloser or gained through a privacy invasion, pass it on to others – for example, gossip). In the early privacy research described, invasions and violations were not emphasized. Ingham (1978) states that 'In everyday social life most individuals are only rarely confronted with an invasion of their privacy, although the number of potential threats is very large' (1978: 40). However, more recently, technology has fuelled debate and controversy about potential invasions and violations to privacy (Dinev and Hart 2004), as will be described below.

Privacy and the Internet

Since the concept of privacy has been applied to technology (e.g. Agre and Rotenberg 1997; Austin 2003) there have been numerous cases reported of the clash between privacy and new technology – how these technologies allow intrusions into private, enclosed spaces, eroding the distinction between public and private space and therefore compromising the very idea of private space. For example, at the end of last year, a body scanning machine was introduced in an airport in the UK. This x-ray machine produces 'naked' images of passengers enabling any hidden weapons or explosives to be discovered. However, this introduction of this technology raised concerns about privacy both among travellers and aviation authorities (*The Sunday Times* 2004).

The concept of privacy has also been applied to the Internet (e.g. Cranor 1999). The increased use of computers and of the Internet now fills many parts of people's lives including online shopping, the sharing of documents and various forms of online communication. It is this increased use of the Internet which raises concerns about privacy, in particular, those described above under informational privacy. There are concerns that the Internet seems to erode privacy (Rust *et al.* 2002) and that offline privacy concerns are magnified online (Privacy Knowledge Base 2005). Indeed, the subject of online privacy has been appearing in newspaper articles regularly over the last few years (e.g. *The Times* 2001; The *Guardian* 2004).

Personal information is fast becoming one of the most important ethical issues of our information age (Milberg *et al.* 1995): personal information has become a basic commodity and users' online actions are no longer simply actions but rather data that can be owned and used by others. Advances in technology and the increased use of the Internet have changed the ways in which information is gathered and used. A wide variety of information data is now collected with increasing frequency and in different contexts, making individuals become ever more transparent. The costs of obtaining and analysing this are also decreasing with the advances in technology. However, the value of the users' information which is collected is increasing.

> At no time have privacy issues taken on greater significance than in recent years, as technological developments have led to the emergence of an 'information society' capable of gathering, storing and disseminating increasing amounts of data about individuals.
>
> (Schatz Byford 1996: 1)

There are a number of specific threats to online privacy. For example, the impact of 'ubiquitous' computing (Weiser 1988) means that we leave data footprints in many areas of

our lives that were previously considered 'offline'. The extremely rapid development of computing power, in terms of greater processing speed, increased storage capacity, wider communication connectivity and lower machine size all impact on privacy (Sparck-Jones 2003). These rapid advances mean that information can be efficiently and cheaply collected, stored and exchanged – even data which may be deemed sensitive by the individuals concerned. Information that is drawn from the physical world is harboured in electronic databases, which give these records permanence, malleability and transportability that has become the trademark of technology. As such, massive databases and Internet records of information about individual financial and credit history, medical records, purchases and so on exist.

Sparck-Jones (2003) labels a number of specific properties of the information collected which have consequences for privacy:

- *Permanence* – once recorded, information rarely disappears. As such, fine-grained, searchable, persistent data exists on individuals and there are sophisticated, cheap, data-mining devices can also be used to analyse this information;
- *Volume* – the ease with which information is now recorded using technology results in huge data sets. Furthermore, storage is cheap, therefore large volumes of information sets can exist indefinitely;
- *Invisibility* – all information collected seems to exist within an opaque system and so any information collected may not be 'visible' to whom it relates. Even if information collected is available to a person they may not be able to interpret it due to the use of incomprehensible coding;
- *Neutrality* – the ease with which information can be collected means that any qualifying information may be lost. So information may be absorbed regardless of its metadata. i.e. there are no distinctions between intimate, sensitive information and non-sensitive information;
- *Accessibility* – there are a number of tools for accessing information meaning that any information collected can possibly be read by any number of people. The ease with which information can be copied, transferred, integrated and multiplied electronically further increases this accessibility;
- *Assembly* – there are many effective tools for searching for and assembling and reorganizing information from many quite separate sources;
- *Remoteness* – information collected is usually both physically and logically away from the users to whom it refers. However, this information can be accessed and used by people who the user does not know.

Each of the above features affects privacy and their effect in combination is even greater. Although massive data collection and storage is possible in many environments, the online privacy problem is further exacerbated by the very structure of the Internet and its additional feature of *connectivity*. The Internet allows for interactive two-way communication and is woven into people's lives in a more intimate way than some other media as it connects people with places and people with people. Accordingly it poses unique information privacy threats that differ from issues previously addressed by research (e.g. Smith *et al.* 1996) therefore, making information collection, sharing and so on even easier.

There are also *benefits* to the technological advances described, such as personalized services, convenience and efficiency. In this way, the collection of personal information can be considered a 'double-edged sword' (Malhotra *et al.* 2004). Users can trade off providing valuable information about themselves to take advantage of benefits – for example, providing personal details and credit card information in order to have the convenience of completing an online transaction. Jupiter Research (2002) have found evidence that even privacy concerned individuals are willing to trade privacy for convenience or to bargain the release of very personal information in exchange of relatively small rewards. However, consumer concern over disclosing personal information is growing as they realize that data about their internet behaviours is being collected without their knowledge and agreement. These privacy concerns can ultimately reduce the personalization benefits that companies can deliver to consumers. The question is whether the benefits of the advances in technology and

the use of the Internet are diminished by endangering privacy.

Linking models of privacy and CMC

According to Berscheid (1977), privacy is the 'hidden variable' in many social psychological studies. In the years since her article was published, there has been relatively few attempts to expose this hidden variable to scrutiny in the psychological literature. Privacy is particularly important for understanding self-disclosure, since the relationship between privacy and self-disclosure is somewhat paradoxical. Privacy is a prerequisite for disclosure, and yet, the process of disclosure serves to reduce privacy. The Internet may, in some instances, serve to solve this paradox – disclosure and intimacy can be achieved without concurrent increases in vulnerability or losses of privacy (see Ben-Ze′ev 2003). But this introduces a further paradox – the Internet, and new media in general, have tended to erode privacy through, amongst others, the processes we outline above. Often the impression of privacy is a mirage – high levels of personal information are held by a number of gatekeepers – whether it is through the process of registration, caches and logs kept on various servers or even locally based records. It therefore becomes critical to understand the role of these gatekeepers to understand fully disclosure of personal information online. We propose that as well as looking at the micro-level impacts of the media environment on disclosure, one also needs to look at the macro-level – the wider context in which the micro-level behaviour is enacted.

Trust and disclosure

Trust is a critical issue in both FtF and online disclosure of personal information. By disclosing information, we are making ourselves vulnerable – one reason it is often easier to disclose to strangers than to close friends and family (Rubin 1975). This applies equally to disclosure to web-based forms – for instance, Moon (1998) found that people are more willing to disclose personal information to geographically distant servers – presumably because the vulnerability of doing so is reduced. In e-commerce, the issue of trust is also critical – people will generally not disclose personal information to a web service that they do not trust (Hoffman *et al.* 1999).

However, many attempts to establish trust between people and within groups rely on methods that increase the media richness of the interaction – for instance, by introducing video, audio or photographs (see Olsen *et al.* 2002; Chapter 5 this volume). Quite apart from the substantial problems with media richness approaches to understanding online behaviour (see Walther 1996), introducing cues that are supposed to improve trust may well serve to reduce privacy in an interpersonal context.

However, in some instances trust will be critical. For instance, if you register to a discussion board, dating site or other web-based service, you will commonly be required to disclose to the owner of the site your real name, age, location/ZIP or postal code, and email address. It is not uncommon to also be asked questions about salary, occupation, marital status and other marketing-related queries. In the cases of discussion boards and dating sites, this disclosure of personal information purchases access to a pseudonymous interactive environment in which participants can seek help, be intimate or just play, with little concern for the repercussions in their offline lives. In this situation, expressive privacy has been obtained through the loss of informational privacy to a third party. Critically, we would argue that it is this separation between the location of the expressive environment, and the third party, that is important. Obviously too, one would also expect that for this bargain to work, the third party must be trustworthy.

For trust to be established, it is not always necessary for privacy to be reduced. For instance, reputation systems (as used on eBay, the auction site) allow trust to be established through the use of peer-ratings of pseudonyms (Utz 2004). However trust is established, it is clearly critical to understanding online behaviour, and is likely to become more important as we leave our personal data at the door of pseudonymous environments.

Cost and benefits

In the example above, access to an environment in which expressive privacy is enabled has been

effectively purchased with personal information. This commodification of personal information is nothing new – witness the growth in customer cashback or 'loyalty' cards provided by grocery shops – but what is interesting is that one form of privacy is lost to gain another form. Andrade *et al.* (2002) adopt a social exchange framework to study consumers' willingness to disclose personal information, although their results suggest that considering people's decision-making within this framework alone does not explain the results of the study. For instance, while manipulations that seemed to reduce the cost of personal disclosure (e.g. privacy policy) did indeed have the desired effect, the offer of a reward worked to reduce disclosure. Presumably this may be because offering financial rewards opened questions of trust.

There are many other occasions when decisions about whether or not to disclose personal information can be interpreted from a social exchange approach. For instance, in many cases a loss of privacy provides benefits in terms of convenience rather than financial gain. Within person-to-person interaction, self-disclosure can also be understood in terms of costs and benefits. As Antaki *et al.* (2005) note, disclosure needs to be 'brought off' – it does not occur without repercussions for both interactants. By disclosing personal information, the cost to a person is increased vulnerability and a loss of privacy. However, in many cases, the benefits – a building of trust, rapport, and reciprocation – will outweigh the costs. However, this is not to say that disclosure is not without risks. For instance, a teenager agonizing about whether to confess to a romantic crush is likely to be acutely aware that disclosure to the object of their desire is a potentially risky business that will lead to either a joyful reciprocation of feelings, or rejection.

In terms of e-commerce, there are also clear cost–benefit issues regarding privacy and disclosure. For instance, imagine the same teenager has successfully arranged their date, and they now wish to purchase prophylactics. They have two options: the first, to pay in their local town with cash, is reasonably high in privacy – there is no data trail, and unless the server behind the counter knows them, they have high information privacy. The alternative is to use a credit card to purchase the desired products online, and to have them delivered at their home address in a plain envelope. In this second case, the level of information privacy is low – they will need to disclose their name, address and credit card details, but expressive or social privacy is high. The method chosen will illustrate the relative costs and benefits our fictional teenager attaches to information and expressive/social privacy.

A critical issue in applying such an economic model to understanding privacy and disclosure is the value placed upon personal information by the individual, and their interpretation of the likely costs of disclosure. As such, people's privacy concerns and the level of trust they have in the recipient of the disclosure will determine the outcome of any cost–benefit analysis.

Control

A further context issue that we believe is important to understanding self-disclosure online is control – that is, control over what information is collected, and how and with whom information is shared.

Information is often collected online with or without the user's knowledge or consent. From a technical standpoint, some types of information are easier to obtain than others. Information can be gathered unobtrusively, which requires little cooperation on the part of the person supplying the information. For example, information may be collected by means of cookies and other software designed to track users' movements over the Internet. Other types of information are less accessible, forcing companies to rely on more intrusive means to obtain important data. This typically involves asking people to engage in some type of self-disclosure.

Individual control over personal information is more difficult than ever before. Even when personal information is voluntarily provided, privacy may still be compromised due to the inability of an individual to control the use of the information. For example, privacy may be comprised on two dimensions (Culnan and Armstrong 1999):

1. *Environmental control* – if personal information is accessed by unauthorized means (e.g. through a security breach or an absence of appropriate internal controls);

2. *Control over secondary use of information* – if information provided for one purpose is used for unrelated purposes, without the individual's knowledge or consent. (e.g. through the duplication and sharing of computerized information).

The secondary use of information and the fact that information may be logged and preserved for future access mean that threats to privacy on the Internet can be immediate as well as future threats.

Most people do not know what information is stored about them or who has access to it. However, there is now a growing awareness, as well as resentment, of the routine practice of collecting and analysing personal information (Nissenbaum 1998). This is partly due to reports in newspapers and on online news sites. For example the 'extent of UK snooping revealed' story reported that 'officials in the UK are routinely demanding huge quantities of information about what people do online and who they call, say privacy experts' (BBC News 2003a). Also the 'Top UK sites 'fail privacy test" story reported '98% do not give enough information about the text files which track user movements, or provide a single-click opt-out option' (BBC News 2003b). A February 2002 Harris Interactive Survey (Harris Interactive 2002) stated that the three biggest consumer concerns in the area of online personal information security were: companies trading personal data without permission, the consequences of insecure transactions, and theft of personal data.

Within the context of person-to-person interaction, clearly control is also a critical issue. Walther (1996) argues that hyperpersonal social interaction online occurs, at least in part, because of the increased control afforded by synchronous, visually anonymous CMC. For instance, we can control what information we choose to disclose, in what manner, and how we disclose it. If privacy and self-disclosure are viewed as dynamic processes, then the removal of control affects the ability of people to effectively regulate their social interactions. We may wish to control the flow of personal information for a number of reasons – for instance, to maintain a desired level of intimacy and privacy, or to maintain social harmony by not disclosing specific information – but without control over what is disclosed and to whom, this is not possible.

Conclusions

Self-disclosure is one of the few widely replicated and noted media effects of online interaction. However, despite the evidence that self-disclosure occurs in a number of different contexts online, including CMC, weblogs and submission of web forms, most approaches to understanding the phenomenon confine themselves to considering the impact of a single factor – anonymity. We argue that by focusing solely on this micro-level media effect, the wider context in which disclosure is given, or required, is ignored – and that ignoring this context limits how we can conceptualize online behaviour. By considering the wider context, and in particular its implications for privacy, it is possible to develop a more nuanced picture of online behaviour across situations.

Acknowledgements

The writing of this chapter was supported by a grant from the Economic and Social Research Council (RES-341–25–0011).

References

Agre, P. E. and Rotenberg, M. (eds) (1997). *Technology and privacy: The new landscape*. Cambridge, MA: MIT Press.

Altman, I. (1975). *The environment and social behavior*. Monterey, CA: Brooks/Cole.

Altman, I. and Taylor, D. (1973). *Social penetration: The development of interpersonal relationships*. New York: Holt, Rinehart and Winston.

Andrade, E. B., Kaltcheva, V. and Weitz, B. (2002). Self-disclosure on the web: the impact of privacy policy, reward, and company reputation. *Advances in Consumer Research 29,* 350–353.

Antaki, C., Barnes, R. and Leudar, I. (2005). Self-disclosure as a situated interactional practice. *British Journal of Social Psychology 44,* 181–200.

Austin, L. (2003). Privacy and the question of technology. *Law and Philosophy 22,* 119–166.

Bargh, J. A., McKenna, K. Y. A. and Fitzsimons, G. M. (2002). Can you see the real me? Activation and expression of the true self on the Internet. *Journal of Social Issues 58,* 33–48.

BBC News (2003a). Extent of UK snooping revealed. 16 May. Available at http://news.bbc.co.uk/1/hi/technology/3030851.stm. Accessed 20 June 2005.

BBC News (2003b). Top UK sites 'fail privacy test'. 11 December. Available at: http://news.bbc.co.uk/1/hi/technology/3307705.stm. Accessed 20 June 2005.

Ben-Ze′ev, A. (2003). Privacy, emotional closeness, and openness in Cyberspace. *Computers in Human Behavior 19,* 451–467.

Berendt, B., Gunther, O. and Spiekerman, S. (2005). Privacy in E-commerce: stated behaviour versus actual behaviour. *Communications of the ACM 48,* 101–106.

Berger, C. R. and Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: toward a developmental theory of interpersonal communication. *Human Communication Theory 1,* 99–112.

Buchanan, T., Joinson, A. N. and Ali, T. (2002). Development of a behavioural measure of self-disclosure for use in online research. Paper presented at German Online Research 2002, Hohenheim, Germany.

Burgoon, J. K., Parrott, R., LePoire, B. A., Kelley, D. L., Walther, J. B. and Perry, D. (1989). Maintaining and restoring privacy through communication in different types of relationship. *Journal of Social and Personal Relationships 6,* 131–158.

Chesney, T. (2005). Online self disclosure in diaries and its implications for knowledge managers. In *UK Academy for Information Systems Conference proceedings*, 22–24 March, Northumbria University, UK.

Computerworld (2005). Available at http://www.computerworld.com/printthis/2005/0,4814,101766,00.html. Accessed 20 June 2005.

Cranor, L. F. (1999). Internet privacy. *Communications of the ACM 42,* 29–31.

Culnan, M. J. and Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: an empirical investigation. *Organization Science 10,* 104–115.

DeCew, J. (1997). *In pursuit of privacy: Law, ethics, and the rise of technology*. Ithaca, NY: Cornell University Press.

Des Jarlais, D. C., Paone, D., Milliken, J., Turner, C. F., Miller, H., Gribble, J., Shi, Q., Hagan, H. and Friedman, S. (1999). Audio-computer interviewing to measure risk behaviour for HIV among injecting drug users: a quasi-randomised trial. *The Lancet 353*(9165), 1657–1661.

Dinev, T. and Hart, P. (2004). Internet privacy concerns and their antecedents – measurement validity and a regression model. *Behaviour and Information Technology 23,* 413–423.

Epstein, J. F., Barker, P. R. and Kroutil, L. A. (2001). Mode effects in self-reported mental health data. *Public Opinion Quarterly 65,* 529–550.

Ferriter, M. (1993). Computer-aided interviewing and the psychiatric social history. *Social Work and Social Sciences Review 4,* 255–263.

Frick, A., Bächtiger, M. T. and Reips, U.-D. (2001). Financial incentives, personal information and drop-out in online studies. In Reips, U.-D. and Bosnjak, M. (eds), *Dimensions of internet science* (pp. 209–219). Lengerich: Pabst.

Galegher, J. Sproull, L. and Kiesler, S. (1998). Legitimacy, authority and community in electronic support groups. *Written Communication 15,* 493–530.

Garfinkel, S. (2000). *Database nation: The death of privacy in the 21st century*. Sebastopol, CA: O'Reilly and Associates, Inc.

Greist, J. H., Klein, M. H. and VanCura, L. J. (1973). A computer interview by psychiatric patient target symptoms. *Archives of General Psychiatry 29,* 247–253.

Guardian, The (2004). The privacy debate: this time it's personal. 26 April.

Harris and Associates Inc. and Westin, A. (1998). E-commerce and privacy: what net users want. Privacy and American Business and Pricewaterhouse Coopers LLP. Available at http://www.pandan.org/ecommercesurvey.html. Accessed 20 June 2005.

Harris Interactive (2002). First major post-9/11 privacy survey finds consumers demanding companies do more to protect privacy; public wants company privacy policies to be independently verified. Available at http://www.harrisinteractive.com/news/allnewsbydate.asp?NewsID=429. Accessed 20 June 2005.

Hoffman, D. L., Novak, T. P. and Peralta, M. (1999). Building consumer trust online. *Communications of the ACM 42,* 80–85.

Holbrook, A. L., Green, M. C. and Krosnick, J. A. (2003). Telephone vs. face-to-face interviewing of national probability samples with long questionnaires: comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly 67,* 79–125.

Home Office (1972). *Report of the Committee on Privacy*. Rt. Hon. Kenneth Younger, Chairman. London: HMSO.

Ingham, R. (1978). Privacy and psychology. In Young, J. B. (ed.) *Privacy* (pp. 35–59). Chichester: Wiley.

Introna, L. D. and Pouloudi, A. (1999). Privacy in the information age: stakeholders, interests and values. *Journal of Business Ethics 22,* 27–38.

Jenson, C., Potts, C. and Jenson, C. (2005). Privacy practices of internet users: self-reports versus observed behaviour. *International Journal of Human Computer Studies*. Special issue on HCI Research in Privacy and Security.

Joinson, A. N. (1999). Anonymity, disinhibition and social desirability on the Internet. *Behaviour Research Methods, Instruments and Computers 31,* 433–438.

Joinson, A. N. (2001a). Self-disclosure in computer-mediated communication: the role of self-awareness and visual anonymity. *European Journal of Social Psychology 31,* 177–192.

Joinson, A. N. (2001b). Knowing me, knowing you: reciprocal self-disclosure on the internet. *Cyberpsychology and Behavior 4,* 587–591.

Joinson, A. N. (2003). *Understanding the psychology of internet behaviour: Virtual worlds, real lives*. Basingstoke and New York: Palgrave Macmillan.

Joinson, A. N. (2004). Self-esteem, interpersonal risk and preference for e-mail to face-to-face communication. *CyberPsychology and Behaviour 7*, 472–478.

Joinson, A. N. (2005). Internet behaviour and the design of virtual methods. In C. Hine (ed.), *Virtual methods: Issues in social research on the internet* (pp. 000–000). Oxford: Berg.

Joinson, A. N., Woodley, A. and Reips, U-R. (in press). Personalization, authentication and self-disclosure in self-administered Internet surveys. *Computers in Human Behavior*.

Jourard, S. M. (1971). *Self-disclosure: An experimental analysis of the transparent self*. New York: Krieger.

Jourard, S. M. and Lasakow, P. (1958). Some factors in self-disclosure. *Journal of Abnormal and Social Psychology 56*(1), 91–98.

Jupiter Research (2002). Security and privacy data. Presentation to the Federal Trade Commission Consumer Information Security Workshop. Available at http://www.ftc.gov/bcp/workshops/security/0205201leathern.pdf. Accessed 20 June 2005.

Kenny, D. A. and Judd, C. M. (1986). Consequences of violating the independence assumption in the analysis of variance. *Psychological Bulletin 99,* 422–431.

Knapp, H. and Kirk, S. A. (2003). Using pencil and paper, Internet and touch-tone phones for self-administered surveys: does methodology matter? *Computers in Human Behaviour 19,* 117–134.

Kupritz, V. W. (2000). The role of the physical environment in maximising opportunities for the aging workforce. *Journal of the Industrial Teacher Education 37,* 66–88.

Lau, J. T. F., Tsui, H. Y. and Wang, Q. S. (2003). Effects of two telephone survey methods on the level of reported risk behaviours. *Sexually Transmitted Infections 79,* 325–331.

Laurenceau, J. P., Barrett, L. F. and Pietromonaco, P. R. (1998). Intimacy as an interpersonal process: the importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. *Journal of Personality and Social Psychology 74,* 1238–1251.

Lessler, J. T., Caspar, R. A., Penne, M. A. and Barker, P. R. (2000). Developing computer-assisted interviewing (CAI) for the National Household Survey on Drug Abuse. *Journal of Drug Issues 30,* 19–34.

Malhotra, N. K., Kim, S. S. and Agarwal, J. (2004). Internet users' information privacy concerns (IUIPC): the construct, the scale and a causal model. *Information Systems Research 15,* 336–355.

Margulis, S. T. (2003). On the status and contribution of Westin's and Altman's theories of privacy. *Journal of Social Issues 59,* 411–429.

Marshall, N. J. (1974). Dimensions of privacy preferences. *Multivariate Behavior Research 9,* 255–272.

McKenna, K. Y. A and Bargh, J. (1998). Coming out in the age of the Internet: identity demarginalization through virtual group participation. *Journal of Personality and Social Psychology 75,* 681–694.

Metzger, M. J. (2004). Privacy, trust and disclosure: exploring barriers to electronic commerce. Available at http://www.jcmc.indiana.edu/vo19/issue4/metzger.html. Accessed 20 June 2005.

Milberg, S. J., Burke, S. J. and Smith, H. J. (1995). Values, personal information privacy, and regulatory approaches. *Communications of the ACM 38,* 65–74.

Moon, Y. (1998). Impression management in computer-based interviews: the effects of input modality, output modality, and distance. *Public Opinion Quarterly 62,* 610–22.

Moon, Y. (2000). Intimate exchanges: using computers to elicit self-disclosure from consumers. *Journal of Consumer Research 27,* 323–339.

Musch, J., Broder, A. and Klauer, K. C. (2001). Improving survey research on the World-Wide Web using the randomized response technique. In U. D. Reips and M. Bosnjak (eds), *Dimensions of Internet science* (pp. 179–192). Lengerich, Germany: Pabst Science Publishers.

Nissenbaum, H. (1998). Protecting privacy in an information age: the problem of privacy in public. *Law and Philosophy 17,* 559–596.

O'Neill, D. (2001). Analysis of Internet users' level of online privacy. *Social Science Computer Review 19,* 17–31.

Olivero, N. (2001). Self-disclosure in e-commerce exchanges: relationships among trust, reward and awareness. Paper presented at the Psychology and the Internet: A European Perspective conference, DERA, Farnborough, UK.

Olson, J. S., Zheng, J., Bos, N., Olson, G. M. and Veinott, E. (2002). Trust without touch: jumpstarting long-distance trust with initial social activities. In *CHI2002 Conference Proceedings* (pp. 141–146). New York: ACM Press.

Parent, W. (1983). Privacy, morality and the law. *Philosophy and Public Affairs 12,* 269–288.

Parks, M. R. and Floyd, K. (1996). Making friends in cyberspace. *Journal of Communication 46,* 80–97.

Pederson, D. M. (1979). Dimensions of privacy. *Perceptual and Motor Skills 48,* 1291–1297.

Pennebaker, J. W., Kiecolt-Glaser, J. K. and Glaser, R. (1988). Disclosure of traumas and immune function: health implications for psychotherapy. *Journal of Consulting and Clinical Psychology 56,* 239–245.

Privacy Knowledge Base (2005). Available at http://privacyknowledgebase.com. Accessed 20 June 2005.

Quittner, J. (1997). Invasion of privacy. *Time Magazine*, 25 August, pp. 30–31.

Raab, C. Bellamy, C., Staylor, J., Dutton, W. H. and Peltu, M (1996). The information polity: electronic democracy, privacy and surveillance. In Dutton, W. H. (ed.) *Information and communication technologies: Visions and realities*. Oxford: Oxford University Press.

Rheingold, H. (1993). *The virtual community*, revised edn. London: MIT Press.

Robinson, R. and West, R. (1992). A comparison of computer and questionnaire methods of history-taking in a genito-urinary clinic. *Psychology and Health 6,* 77 – 84.

Rubin, Z. (1975). Disclosing oneself to a stranger: reciprocity and its limits. *Journal of Experimental Social Psychology 11,* 233–260.

Rust, R. T., Kannan, P. K. and Peng, N. (2002). The customer economics of internet privacy. *Journal of the Academy of Marketing Science 30,* 455–464.

Schatz Byford, K. (1996). Privacy in cyberspace: constructing a model of privacy for the electronic communications environment. *Rutgers Computer and Technology Law Journal 24,* 1–74.

Schoeman, F. (1984). Privacy and intimate information. In Schoeman, F. (ed.) *Philosophical dimensions of privacy* (pp. 403–417). Cambridge: Cambridge University Press.

Smith, H. J., Milberg, S. J. and Burke, S. J. (1996). Information privacy: measuring individuals' concerns about organizational practices. *MIS Quarterly 20,* 167–196.

Smyth, J. M. (1998). Written emotional expression: effect sizes, outcome, types, and moderating variables. *Journal of Consulting and Clinical Psychology 66,* 174–184.

Sobel, D. L. (2000). The process that 'John Doe' is due: addressing the legal challenge to Internet anonymity. *Virginia Journal of Law and Technology* 5.

Sparck-Jones, K. (2003). Privacy: what's different now? *Interdisciplinary Science Reviews 28,* 287–292.

Sproull, L., Subramani, M., Kiesler, S., Walker, J. H. and Waters, K. (1996). When the interface is a face. *Human–Computer Interaction 11,* 97–124.

Tidwell, L. C. and Walther, J. B. (2002). Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: getting to know one another a bit at a time. *Human Communication Research 28,* 317–348.

Times, The (2001). Governments seem determined to overthrow online privacy – even if they have to behave like hackers to do so. Technobabble, 3 December.

Times, The (2001). Junk e-mails cost £6.4bn. 3 February.

Times, The (2004). Plane passengers shocked by their x-ray scans. 7 November.

Tourangeau, R. (2004). Survey research and societal change. *Annual Review of Psychology 55,* 775–801.

Tourangeau, R. and Smith, T. W. (1996). Asking sensitive questions: the impact of data collection mode, question format, and question context. *Public Opinion Quarterly 60,* 275–304.

Tourangeau, R., Couper, M. P. and Steiger, D. M. (2003). Humanizing self adminstered surveys: experiments on social presence in web and IVR surveys. *Computers in Human Behaviour 19,* 1–24.

Utz, S. (2004). Trust at eBay – influenced by the reputation of the seller or the description of the product? Paper presented at German Online Research 2004, University of Duisburg-Essen, Germany.

Wallace, P. (1999). *The psychology of the internet*. Cambridge: Cambridge University Press

Walther, J. B. (1996). Computer-mediated communication: impersonal, interpersonal, and hyperpersonal interaction. *Communication Research 23,* 3–43.

Warren, S. and Brandeis, L. D. (1890). The right to privacy. *Harvard Law Review 4,* 193–220.

Weisband, S. and Kiesler, S. (1996). Self-disclosure on computer forms: meta-analysis and implications. *Proceedings of CHI96*. Available at http://www.acm.org/sigchi/chi96/proceedings/papers/Weisband/sw_txt.htm. Accessed 20 June 2005.