CHAPTER 19

Digital deception

Why, when and how people lie online

Jeffrey T. Hancock

Deception is one of the most significant and pervasive social phenomena of our age (Miller and Stiff 1993). Some studies suggest that, on average, people tell one to two lies a day (DePaulo *et al.* 1996; Hancock *et al.* 2004a), and these lies range from the trivial to the more serious, including deception between friends and family, in the workplace and in power and politics. At the same time, information and communication technologies have pervaded almost all aspects of human communication and interaction, from everyday technologies that support interpersonal interactions, such as email and instant messaging, to more sophisticated systems that support organizational interactions.

Given the prevalence of both deception and communication technology in our personal and professional lives, an important set of questions has recently emerged at the intersection of deception and technology, or what we will refer to as '*digital deception*'. These questions include issues concerned with deception and self-presentation, such as how the Internet can facilitate deception through the manipulation of identity. A second set of questions is concerned with how we produce lies. For example, do we lie more in our everyday conversations in some media more than in others? Do we use different media to lie about different types of things, to different types of people? Another type of question concerns our ability to *detect* deception across various media and in different online communication spaces. Are we worse at detecting lies in a text-based interaction than we are in face-to-face (ftf)? What factors interact with communication media to affect our ability to catch a liar?

In the present chapter I examine these questions by first elaborating on the notion of digital deception in the context of the literature on traditional forms of deception. The chapter is then divided into two main sections, one concerned with identity-based forms of deception online, and the other focusing on the lies that are a frequent part of our everyday communications.

Digital deception defined

Deception has been studied in a wide variety of contexts (Ekman 2001), including organizational settings (Grazioli and Jarvenpaa 2003a; Schein 2004), forensic and criminal settings (Vrij 2000; Granhag and Stromwall 2004), in power and politics (Ekman 1985; Galasinski 2000) and in everyday communication (DePaulo *et al.* 1996; DePaulo and Kashy 1998; Hancock *et al.* 2004a, b). In the present chapter, we consider deception in the

context of information and communication technology, or what I will call *digital deception*, which refers to *the intentional control of information in a technologically mediated message to create a false belief in the receiver of the message.* While this definition is an adaptation of Buller and Burgoon's (1996) conceptualization of deception, i.e., 'a message knowingly transmitted by a sender to foster a false belief or conclusion by the receiver' (1996: 205), the characteristics of this definition are consistent with most definitions of deception (for review of the many issues associated with defining deception, see Bok 1978; Galasinski 2000). The first characteristic is that an act of deception must be intentional or deliberate. Messages that are unintentionally misleading are usually not considered deceptive, but instead are described as mistakes or errors (Burgoon and Buller 1994). Similarly, forms of speech in which the speaker does not mean what they say but intend for their addressee to detect this, such as irony, joking, etc, are not considered deceptive. The second characteristic is that deception is designed to mislead or create a false belief in some target. That is, the deceiver's goal is to convince someone else to believe something that the deceiver believes to be false. These characteristics can be observed, for example, in Ekman's (2001: 41) definition – 'deliberate choice to mislead a target without giving any notification of the intent to do so' – and in DePaulo *et al.*'s (2003: 74) – 'a deliberate attempt to mislead others.'

Digital deception requires an additional characteristic, namely that the control or manipulation of information in a deception is enacted in a *technologically mediated message*. That is, the message must be conveyed in a medium other than the basic ftf setting. As such, digital deception involves any form of deceit that is transmitted via communication technology, such as the telephone, email, instant messaging, chat rooms, newsgroups, weblogs, listservs, multiplayer online video games etc.

Although a number of different typologies have been proposed for categorizing deception – for example deception by omission vs. by commission, active vs. passive deception, etc. (see Robinson 1996; Galasinski 2000), for the purposes of discussing how the Internet and communication technologies may affect deception and its detection, I break digital deception down into two broad types: those based on a communicator's identity, and those based on the actual messages that comprise a communication. In particular, *identity-based digital deception* refers to deceit that flows from the false manipulation or display of a person or organization's identity. For example, an email designed to look like it originated from someone in Africa that needs a partner to extricate vast sums of money (in order to trick the recipient into providing their bank information) is a case of identity-based digital deception. *Message-based digital deception*, in contrast, refers to deception that takes place in the communication between two or more interlocutors or agents. In particular, it refers to deception in which the information in the messages exchanged between interlocutors is manipulated or controlled to be deceptive. For example, when one friend calls another on his mobile phone to say that he will be late to their meeting because the traffic is bad (when in fact he simply left the office late) is an example of message-based digital deception. The two friends' identities are known to one another, but the information provided by the first friend has been manipulated to create a false belief in the second friend.

Clearly these identity-based and message-based forms of digital deception are not mutually exclusive. Indeed, the messages in a communication may serve to enhance a deception about identity and, when identity-based digital deception is enacted, the messages that make up the communication are more than likely to also be deceptive. For instance, in the email example above, there are several possible relationships between identity- and message-based deceptions. For example, the identity of the sender may be deceptive (i.e., the person is not really someone in Africa), but the message truthful (e.g., the person really does have access to money). Or, the identity of the sender may be accurate (i.e., the person really is in Africa) but the message is deceptive (e.g., the person does not have access to money). Or, both the identity and message could be false. As such, the distinction between identity-based and message-based deception is not intended to be set in stone, but is intended only as a pragmatic distinction that may help us consider how communication technologies may or may not affect deception.

Finally, it should be noted that the definition of digital deception described above includes a number of issues that are beyond the scope of this chapter. For example, the advent of sophisticated and relatively inexpensive digital editing software makes image-based digital deception, such as misleading editing or selection, an important issue (Messaris 1997; Galasinski 2000). Also, the very broad topic of information security, such as attacks and vulnerabilities on information infrastructure (see Schneider 1999), hacking and deceptive intrusion of information networks (see Stolfo *et al.* 2001), will also not be discussed here. Instead, the focus will be on deception in our everyday mediated communication.

Identity-based digital deception

Perhaps the most obvious deception issue to consider is the affordances provided by information and communication technologies to manipulate or obscure our identity. As Turkle (1995) observed, the relative anonymity and multiple modes of social interaction provided by the many forms of online communication conducted via the Internet provide users with unique opportunities to play with their identity and explore their sense of self. As many have now noted (e.g., Walther 1996; Berman and Bruckman 2001; Bargh *et al.* 2002; Spears *et al.* 2002; Walther and Parks 2002), because online communication typically involves text-based interaction or virtual representations of self (e.g., avatars), people can self-present in ways that they can not in ftf encounters. Boys can be girls, the old can be young, ethnicity can be chosen, 15-year-olds can be stock analysts – and on the Internet no one knows you're a dog.

While this growing body of research has revealed some of the fascinating effects that the relative anonymity of the Internet can have on identity and social interaction, such as the enhancement of group effects (e.g., Postmes *et al.* 1999; Douglas and McGarty 2001) and the potential hyerpersonalization of interpersonal interactions (Walther 1996; Hancock and Dunham 2001a; Walther *et al.* 2001), the affordances of online communication for manipulating identity also have important implications for deception. In one of the first systematic investigations of identity-based deception in online contexts, Donath (1998) observed how different aspects of Usenet newsgroups (asynchronous text-based message exchange systems supporting a wide range of topical discussions) affected participants sense of identity and their abilities to deceive or be deceived by their fellow community members.

Drawing on models of deception from biology (e.g., Zahavi 1993), Donath distinguished between assessment signals, which are costly displays directly related to an organism's characteristics (e.g., large horns on a stag), and conventional signals, which are low-cost displays that are only conventionally associated with a characteristic (e.g., a powerful-sounding mating call). In online communication, conventional signals include most of the information that is exchanged in messages, including what we say (e.g., that I'm very wealthy) and the nicknames we use to identify ourselves (e.g., 'richie-rich'). Assessment signals may be more difficult to come by online, but can include links to a person's 'real-world' identity, such as a phone number or an email address (e.g., emails ending in.ac.uk or .edu suggest that the person works at a university), or levels of knowledge that only an expert could display (e.g., highly technical information about a computer system).

Online, conventional signals are an easy target for deceptive identity manipulation, and Donath notes several types of deceptive identity manipulations in the Usenet communities, including trolling, category deception, and identity concealment. Trolling refers to an individual posing as a legitimate member of a community who posts messages intended to spark intense fights within the community. Category deception refers to deceptions that manipulate our perceptions of individuals as members of social groups, or categories, such as male vs. female, white vs. black, student vs. worker, hockey player vs. squash player. Online, gender deception is perhaps the most commonly discussed example of category deception (e.g., Turkle 1995; Berman and Bruckman 2001; Herring and Martinson 2004). Finally, identity concealment refers to hiding or omitting aspects of one's identity, such as using a pseudonym when posting, in order to shield one's identity.

Research by Whitty and her colleagues (Whitty and Gavin 2001; Whitty 2002) suggests that the notion of using deception to shield one's identity is important for many participants interacting in relatively anonymous online spaces, such as chat rooms. In particular, in one survey of chat room participants, women reported using deception to conceal their identity for safety reasons, such as avoiding harassment. Men, on the other hand, reported using identity deception in order to allow themselves, somewhat paradoxically, to be more expressive and to reveal secrets about themselves (Utz 2005). Indeed, a number of studies have suggested that self-disclosure and honesty tend to increase online when participants' identities are not manifest (e.g., Joinson 2001; Bargh *et al.* 2002).

More recently, however, the Internet has evolved from a virtual space for exchanging information, chatting with others and forming virtual communities into a massive venue for financial and business transactions, with estimates of revenue generated from online transactions in the billions, and an increasing number of businesses and individuals engaging in commerce online. As might be expected, more serious and criminal forms of deception are keeping pace with the increase in money flowing through the Internet (Grazioli and Jarvenpaa 2003b). Indeed, the Internet Fraud Complaint Center (IFCC 2003) reported almost fifty thousand incidences of fraud online, a threefold increase from the previous year, the majority of which involved fraudulent Internet auctions, but also included credit card fraud and identity theft, in which someone's personal information is stolen and used for the gain of another individual.

In their work on deception that takes place in business and consumer contexts, such as touting unsound investments for personal gain or making misleading claims about goods for sale at an auction site, Grazioli and Jarvenpaa (2003a, b) have identified seven common deception tactics. The first three tactics are concerned with obscuring the nature of the goods to be transacted, and include

1. *Masking* – eliminating critical information regarding an item (e.g., failing to disclose that the publisher of a newsletter receives advertisement money from stocks the newsletter recommends)
2. *Dazzling* – obscuring critical information regarding an item (e.g., free trials that lead to automatic enrolment without making that clear to consumers)
3. *Decoying* – distracting the victim's attention from the transaction (e.g., offers of free products that require the revealing of highly detailed personal information).

The four other types of deception tactics involve manipulating information about the transaction itself, and include:

1. *Mimicking* – assuming someone else's identity or modifying the transaction so that it appears legitimate (e.g., the creation of a 'mirror' bank site virtually identical to the legitimate site, inducing users to disclose personal information such as account information)
2. *Inventing* – making up information about the transaction (e.g., Internet auctioneers who advertise merchandise that they do not have)
3. *Relabelling* – describing a transaction expressly to mislead (e.g., selling questionable investments over the Internet as sound financial opportunities)
4. *Double play* – convincing a victim that they are taking advantage of the deceiver (e.g., emails designed to look like internal memos sent by mistake and which appear to contain insider information).

As Grazioli and Jarvenpaa (2000) note, the Internet offers a highly flexible environment for identity-based forms of deception that can make it difficult for even technologically savvy users to detect deception.

While the Internet certainly offers a number of advantages to the deceiver that may not be available face-to-face, an important question is whether one is more likely to encounter identity-based deception online or in more traditional face-to-face social exchanges. While this question is difficult to address for obvious reasons, a recent report comparing identity fraud that took place online or ftf suggests that identity fraud is still much more likely to take place ftf, and that when it does occur online it tends to be much less costly than when it occurs offline (Javelin Strategy and Research 2005).

While this is only one report, it does serve as a reminder that although Internet-based communication provides many features that may facilitate

identity-based digital deception, and that this type of deception appears to be on the rise online, more traditional ftf forms of communication are certainly not immune to identity related deception. Nonetheless, identity-based digital deception is an important area for future research, especially given reports that criminal entities, such as organized crime and terrorist organizations, are increasingly relying on information technologies to communicate (Knight 2004).

Message-based digital deception

Although we typically associate Internet-based communication with relatively anonymous communication spaces, such as chat rooms, newsgroups, online games, etc., most people's everyday use of communication technologies tend to be with people that they know, such as an email to a colleague, an instant message with a friend, or text messaging on the phone with a spouse. In these instances, much like many of our ftf interactions, the identity of our interlocutors is known to us. How do communication technologies affect deception when identities are known? Let us consider first the production of digital deception.

Producing digital deception

Research suggests that deception in general is a fundamental and frequent part of everyday human communication, both in interpersonal settings as well as in work and organizational contexts (Camden *et al.* 1984; Lippard 1988; Metts 1989; DePaulo *et al.* 1996; Hancock *et al.* 2004a, b). Some research suggests that people tell an average of one to two lies a day (DePaulo *et al.* 1996; Hancock *et al.* 2004a, b), and these daily lies range from the trivial, such as a false opinion about someone's appearance, to the more serious matters, such as deception in business and legal negotiations, power and politics, and workplace issues. Indeed, as noted above, some have argued that deception is one of the most pervasive social phenomena of our age (Miller and Stiff 1993).

How do communication technologies affect the frequency with which we produce lies? In particular, are we more likely to lie in some media than in others? Some have speculated that Internet-based communication is rife with deception. For example, Keyes (2004: 198) argues that 'electronic mail is a godsend. With email we needn't worry about so much as a quiver in our voice or a tremor in our pinkie when telling a lie. Email is a first rate deception-enabler'. While this may reflect a popular view of how communication technology might affect deception, theoretical approaches to media effects suggest several possible ways that media may affect lying behaviour.

Media Richness Theory (Daft and Lengel 1986; Daft *et al.* 1987), for example, assumes that users will choose rich media, which have multiple cue systems, immediate feedback, natural language and message personalization, for more equivocal and complex communication activities. Because lying can be considered a complex type of communication, media richness theory predicts that users should chose to lie most frequently in rich media, such as ftf, and least frequently in less rich media, such as email. In contrast, DePaulo *et al.* (1996) argued that because lying makes people uncomfortable, users should choose less rich media in order to maintain social distance between the liar and the target, an argument I refer to as the social distance hypothesis. According to this hypothesis, users should choose email most frequently for lying, followed in order by instant messaging, telephone and finally ftf (see also, Bradner and Mark 2002).

Note that both of these approaches assume that communication technology varies along only a single underlying dimension (i.e., richness, distance) that will influence deception, and ignore other important differences in their design that may have important implications for deception. In our feature-based model of media and deception (Hancock *et al.* 2004a, b), we proposed that at least three features of media are important for the act of deception, including (1) the *synchronicity* of the medium (i.e., the degree to which messages are exchanged instantaneously and in real-time) (2) the *recordability* of the medium (i.e., the degree to which the interaction is automatically documented), and (3) whether or not the speaker and listener are *distributed* (i.e., they do not share the same physical space).

In particular, synchronous media should increase opportunities for deception because the

majority of lies are unplanned and tend to emerge spontaneously from conversation (DePaulo *et al.* 1996). For example, if during a conversation a new friend says to another that his favorite movie is one that she hates, she is now presented with a decision to lie or not about her opinion of the movie. This type of emergent opportunity is less likely to arise when composing an email. Thus, media that are synchronous, such as ftf and telephone, and to a large degree instant messaging, should present more situations in which deception may be opportune.

The more recordable a medium, the less likely users should be willing to speak falsely. Email is perhaps the most recordable interpersonal medium we have ever developed, with copies being saved on multiple computers (including the targets). In contrast, ftf and telephone conversations are typically recordless. Although instant messaging (IM) conversations are logged for the duration of an exchange and can be easily saved, most people do not save their IM conversations. Of course, this may change as IM enters the workplace and companies begin automatically recording IM by employees. In order to avoid being caught, speakers may choose to lie more frequently in recordless media, such as ftf and the telephone, than in more recordable media, such as email and instant messaging.

Finally, media in which participants are not distributed (i.e., co-present) should constrain deception to some degree because they limit deception involving topics or issues that are contradicted by the physical setting (e.g., 'I'm working on the case report' when in fact the speaker is surfing news on the Web). In fact, software is now available that can be downloaded into a phone that plays ambient noise that may be consistent with your lie (e.g., playing the sounds of an office when in fact you are in a car). Because mediated interactions such as the phone, IM and email involve physically distributed participants, this constraint should be reduced relative to ftf interactions. Some support for this notion comes from a study by Bradner and Mark (2002), in which participants were more likely to deceive a partner when they believed their partner was in a distant city than if they were in the same city.

According to our feature-based model, the more synchronous and distributed but less recordable a medium is, the more frequently lying should occur. As such, if these design features of communication media affect deception, then lying should occur most frequently on the telephone, followed by ftf and instant messaging, and least frequently via email.

To test the predictions flowing from the theories described above, we (Hancock *et al.* 2004a) conducted a diary study adapted from DePaulo *et al.*'s (1996) procedures. After a training session on how to record and code their own social interactions and deceptions, participants recorded all of their lies and social interactions for seven days. For each interaction, they recorded in which medium the interaction took place, ftf, phone, IM, email, and whether or not they lied. The results suggested that participants lied most frequently on the telephone (37 per cent of social interactions), followed by ftf (27 per cent) and IM interactions (21 per cent), and that they lied least by email (14 per cent). These data are not consistent with either media richness theory or the social distance hypothesis, which predict that deception will vary linearly along a single dimension, such as richness or social distance. In contrast, the data are consistent with our feature-based model of deception, which predicted that deception production should be highest in synchronous, recordless and distributed media. The data also go against the conventional wisdom that the online world is rife with deception and subterfuge.

Although the features described in the feature-based model predicted overall rates of digital deception, lies are not homogenous (DePaulo *et al.* 1996; Feldman *et al.* 2002). Deception, for example, can be about one's actions – 'I'm in the library' when in fact the speaker is at the pub – feelings – 'I love your shirt' with regard to a friend's ugly shirt – facts – 'I'm an A student' – and explanations – 'I couldn't make it because my car broke down'. Do people select different types of media for different types of deception? The feature-based model of deception makes several predictions. First, lies about actions should be less likely to occur in non-distributed communicative settings, where the target of the lie can physically see the speaker. Because lies about feelings are most likely to arise in synchronous interactions (e.g., a friend asking whether you like their ugly shirt), lies about feelings were

predicted to occur most frequently face-to-face and on the telephone. Lies about facts should be least likely to be told in recordable media that can later be reviewed, such as email. Finally, explanation type lies were predicted to take place most frequently in asynchronous media, such as email, which provides the liar with more time to construct and plan their explanation than synchronous media.

People also lie differently to different types of people. For example, because people report valuing authenticity and trust in close relationships, people tend to lie less to close relationship partners, such as spouses, family and friends, than to casual relationship partners, such as acquaintances, colleagues and strangers (Metts 1989; Millar and Millar 1995; DePaulo and Kashy 1998). Lies to close and casual relationship targets also seem to differ qualitatively. In particular, lies told in close relationships tend to be more altruistic, in which the lie is told primarily to benefit the target (e.g., false compliments, pretend agreement) than self-serving, in which the lie benefits the liar, while lies in casual relationships tend to be more self-serving than altruistic.

In order to examine whether people used different media to lie about different things or to different people, we conducted another diary-based study in which we also assessed the content and target of the lie (Hancock *et al.* 2004b). While we saw the same pattern of deception frequency across media (i.e., highest rate of deception on the phone, followed by ftf and IM, and least frequently email), the data provided only mixed support for our predictions regarding deception content and target relationship. As predicted, asynchronous interactions involved the least lies about feelings (i.e., email) but involved the most explanation-based lies, which involve explanations about why some event or action occurred – for example 'My dog ate my homework' as an explanation for why a student didn't complete the homework). Distributed media were predicted to involve more lies about actions, but this was only true for lies on the telephone. Finally, lies about facts did not differ across media. With respect to relationships, relative to ftf interactions, phone lies were most likely told to family and significant others. Instant messaging lies were most likely to be told to family. Finally, email lies were most likely to involve lies to higher status individuals, such as a student's professor.

Carlson and George (2004; George and Carlson 2005) have taken a similar approach to examining how the features of a medium, including synchronicity, recordlessness and richness, may affect deception production. While synchronicity and recordlessness are also in the feature-based model described above, Carlson and George (2004) argue that synchronicity may be preferred by deceivers for a somewhat different and very good reason, namely because it increases the deceiver's ability to assess and react to the receiver's behaviour. Richness is considered a positive for deception for the same reason – increased richness should lead to enhanced control over how the receiver perceives the deceiver as truthful. In this approach, however, richness is determined not only be availability of cues and speed of feedback, but also by the participant's experience with that medium (Carlson and Zmud 1999).

In two studies, Carlson and George (2004; George and Carlson 2005) provided a variety of scenarios to business managers that described situations in which they would be required to lie. In general, the results suggested that participants were most likely to choose synchronous and recordless media when they needed to lie, regardless of the severity of the situation. Although these data are generally consistent with the feature-based model, the results in these studies suggested that ftf tended to be the most frequent choice for deception, not the telephone. One possible reason for this difference may be the method employed, which does not control for the different baseline frequencies with which we interact in different media. That is, despite the wide range of communication technologies available to us, the majority of our interactions tend to be ftf. As such, we might expect ftf to be the place that people imagine they will lie most frequently in absolute terms simply because that is where most of their interactions take place.

Regardless of this methodological difference, when considered together, the data from these studies and the ones described above suggest that contrary to some speculations (e.g., Keyes 2004), asynchronous and recordable media, such as email, are unlikely places for people to lie in during their everyday communication.

Instead, more synchronous and recordless forms of media, such as the telephone and ftf settings, appear to be where we lie most.

A final question concerned with how technology might affect deception is whether our language use is different when we lie compared to when we tell the truth online. In groundbreaking work in this area, Zhou and colleagues (Zhou *et al.* 2004a, b, Zhou and Zhang 2004) use computer-assisted, automated analysis of linguistic cues to classify deceptive and non-deceptive text-based communication. In this approach, the language of deceptive and truthful participants' communication are subjected to an automated analysis along a number of linguistic dimensions, including word count, pronoun usage, expressivity, affect and non-immediacy (i.e., less self-reference), among others. For example, in one study examining asynchronous text-based exchanges, Zhou *et al.* (2004) found that, compared to truth-tellers, liars used more words, were more expressive, non-immediate and informal and made more typographical errors. In one of our studies, we (Hancock *et al.* in press a) found similar patterns in synchronous online interaction (i.e., instant messaging), including increased word use and fewer self-references, during deception. Perhaps even more interestingly, we also found that the targets of lies, who were blind to the deception manipulation, also changed systematically depending on whether they were being lied to or told the truth. In particular, when being lied to targets used shorter sentences and asked more questions. These data suggest the fascinating possibility that targets had an implicit awareness or suspicion about the veracity of their partner, despite the fact that when asked whether they thought their partners were lying or not they performed at chance levels. While additional research is required for this novel line of research, these data suggest that how people use language online may change systematically according to whether or not they are being truthful. If this is the case, then the implications for deception detection online are substantial. We turn now to this issue, the detecting of digital deception.

Detecting digital deception

While an extensive literature has examined deception detection in ftf contexts (for review, see Zuckerman and Driver 1985; Vrij 2000; DePaulo *et al.* 2003), the question of how communication technologies affects deception detection has only begun to be addressed. Are we worse at detecting a lie in a text-based interaction than we are in a face-to-face exchange? How do factors that affect deception detection in ftf contexts, such as motivation, suspicion and non-verbal cues, interact with the effects of communication technology?

Although the extensive literature on ftf deception detection suggests that our accuracy to detect deception tends to be around chance (Vrij 2000), there are a number of factors that appear to reliably influence an individual's ability to detect deceit, and these factors may have important implications in the context of digital deception. Perhaps the most intuitively obvious factor for digital deception is the reduction of non-verbal cues that are associated with deception in mediated communication. Previous research suggests that there are a small set of reliable verbal, non-verbal and vocal cues to deception (for review, see DePaulo *et al.* 2003). Perhaps the most important of these are 'leakage cues', which are non-strategic behaviours (usually non-verbal) that are assumed to betray the senders' deceptive intentions or feelings, such as a decrease in illustrators, body movements and higher pitch (Ekman 2001).

Given that these types of leakage cues are eliminated in text-based CMC interactions, one might suppose that deception detection would be less accurate in CMC than in ftf interactions (Hollingshead 2000). However, the relationship between communication media and deception appears to be much more complex than a simple reduction of cues. In perhaps the first theoretical framework to consider systematically the detection of message-based digital deception, Carlson *et al.* (2004) draw on Interpersonal Deception Theory (Buller and Burgoon 1996) to identify a number of variables that may interact with the communication medium in the context of deception detection. These factors include the (1) characteristics of the deceiver and receiver, and of their relationship, and (2) aspects of the communication event and the medium in which it takes place.

Characteristics of the deceiver and receiver that are considered relevant to success rates of

deception detection include the motivation to lie or catch a lie, each individual's intrinsic abilities at deceiving or detecting deceit, aspects of the task and the various cognitions and affect that may arise from the discomfort associated with lying. Experience and familiarity are also assumed to play an important role in the model, including the relational experience between the deceiver and receiver, as well as both individuals' experience with the communication medium and context.

Aspects of the communication medium that are considered important include synchronicity, symbol variety (i.e., the number of different types of language elements and symbols available, including letters, basic symbols, fonts, etc.), cue multiplicity (i.e., number of simultaneous information channels supported), tailorability (i.e., ability to customize the message for the audience), reprocessability (i.e., the inverse of recordlessness described above) and rehearsability (i.e., the degree to which it gives participants time to plan, edit and rehearse messages). In this model, the relationships between these variables and deception detection is not assumed to be simple or one-to-one. Instead, the model assumes a 'deceptive potential' that is derived from constellations of these media variables. In particular, Carlson *et al.* propose that media with higher levels of symbol variety, tailorability, and rehearsability increase deceptive potential and reduce the likelihood of deception detection. In contrast, media that have higher levels of cue multiplicity and reprocessability decrease deceptive potential.

An important underlying assumption of this model, derived from the Interperonal Deception Theory, is that deception is a strategic act that is part of an ongoing, interactive communication process, and that all of the factors described above interact in important and predictable ways. A number of the factors described in the model have begun to be examined in several recent studies examining deception detection in online communication (Heinrich and Borkenau 1998; George and Carlson 1999; Hollingshead 2000; Horn 2001; Horn *et al.* 2002; Burgoon *et al.* 2003; George and Marrett 2004, Carlson and George 2004 Study 2; George *et al.* 2004; Hancock *et al.* in press b).

A survey of these studies suggests that, as Carlson *et al.* (2004) predict, the relationship between communication media and deception detection is not a simple one. Some studies, for example, have found more accurate deception detection in richer media (e.g., Heinrich and Borkenau 1998; Burgoon *et al.* 2003), others have found higher accuracy in less rich media (e.g., Horn *et al.* 2002), while still others have found no overall difference between media (Hollingshead 2000; George and Marrett 2004; Woodworth *et al.* 2005). Instead, it appears that a number of factors, such as those described above, interact with the communication medium to determine deception detection accuracy.

Hancock *et al.* (in press b), for example, examined the impact of motivation of the deceiver and the communication medium on deception detection. People who are highly motivated to get away with their deceptive behaviour tend to act differently than those who are less concerned with the outcome, and their non-verbal behaviour (e.g., increased behavioural rigidity) is more likely to give them away (DePaulo *et al.* 1983). The observation that highly motivated liars are more likely to be detected has been referred to as the *motivational impairment effect* (DePaulo and Kirkendol 1989).

Because CMC eliminates nonverbal cues, the motivation impairment effect should be attenuated for highly motivated liars interacting in CMC. In addition, Burgoon and her colleagues (Burgoon and Buller 1994; Buller and Burgoon 1996) argue that moderately motivated liars engage in strategic communication behaviors to enhance their credibility. If that is the case, then there are several aspects of the CMC environment that should be advantageous to a sufficiently motivated liar (Carlson *et al.* 2004): (1) CMC speakers have more time to plan and construct their utterances, and (2) CMC settings enable the sender to carefully edit their messages before transmitting them to their partner, even in synchronous CMC, which affords speakers greater control over message generation and transmission (Hancock and Dunham 2001b). As such, CMC may not only attenuate the motivational impairment effect, but actually reverse it.

To test this possibility, Hancock *et al.* (in press b) examined deceptive and truthful interactions in ftf and CMC environments. Half of the senders were motivated to lie by telling them that research has shown that successful liars tend to

have better jobs, higher incomes and more success with finding a mate (see Forrest and Feldman 2000), while the other half were not. Deception detection accuracy did not differ across ftf and CMC conditions or across motivation levels. However, an interaction between communication environment and motivation was observed. Consistent with the motivation impairment effect, relative to unmotivated liars, motivated liars in the ftf condition were detected more accurately. In contrast, motivated liars in the CMC condition were detected *less* accurately than unmotivated liars. In fact, a comparison across the four conditions in the study reveals that the highly motivated CMC liars were the *most* successful in their ability to deceive their partner.

We refer to this observation as the *Motivation Enhancement Effect*, which has a number of important implications for digital deception. For example, investigators have warned of the increasing number of intrinsically highly motivated sexual offenders (particularly paedophiles) who have been using various online communication forums to lure potential victims (Mitchell *et al.* 2001). This is a particularly important development given the results of the present study, which suggest that highly motivated liars in CMC contexts are not detected very accurately.

As this study suggests, and the Carlson *et al.* (2004) model predicts, the effect of communication technologies on how humans detect deception is complex. Another interesting line of detection research, however, involves computer-assisted detection of deception (Burgoon *et al.* 2003; Burgoon and Nunamaker 2004). As described above, research on automated textual analysis suggests that there are detectable differences in linguistic patterns across deceptive and non-deceptive text-based communication (e.g., Zhou *et al.* 2004a; Hancock *et al.* in press a). Can a tool be developed that exploits these differences to detect digital deception in real time, as an interaction unfolds? While the prospect of creating this type of tool is appealing, the task of automating the detection of such a complex communication process as digital deception is a clearly daunting one (Burgoon and Nunamaker 2004). Nonetheless, the research findings from the studies described above, which suggest a high diagnostic value of text-based cues (e.g., word quantity, pronoun use, etc.) in digital deception, and the tremendous advances in computing power and statistical classification techniques, lay a foundation for the development of such a tool.

Conclusions

Given the degree to which information and communication technologies pervade many aspects of our lives, it is perhaps difficult to overestimate the impact such technologies may have on one of the oldest aspects of human life, deception. The present chapter provides an overview of the state-of-the-art on the early stages of research on digital deception. Additional research is needed to examine systematically the wide variety of factors that the literature has identified as affecting deception face-to-face, including, among others, the motivation to detect deception, the relationship between deceiver and target, the type and magnitude of the deception, the role of suspicion (e.g., George and Marrett 2004) and experience with the medium.

Similarly, as new technologies are developed and employed, their features and affordances with respect to deception will need to be identified. For example, how do online dating sites, on which people post profiles of themselves, affect deception and its perception (Cornwell and Lundgren 2001; Ellison *et al.* 2004)? How frequently do people lie in their profiles, and what kinds of lies are considered acceptable?

While further studies are needed, the research to date suggests that the questions posed at the beginning of this chapter concerning the intersection of deception and technology have complex answers, but the research also suggests that communication technologies do indeed affect how frequently we lie, about what and to whom. The data also suggest that deception detection will be as complicated, if not more so, online as it is face-to-face, although the potential for computer-assisted deception detection may create new avenues for this age-old issue.

References

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

Berman, J. and Bruc
exploring identity
Convergence 7, 83
Bradner, E. and Mar
effects on cooper
Proceedings of the
Supported Cooper
New York.
Buller, D. B. and J. K.
deception theory.
Burgoon, J. K. and B
deception: III. Eff
communication a
Journal of Nonverl
Burgoon, J. K. and N
computer-aided s
Group Decision an
Burgoon, J. K., Stonel
(2003). Trust and
communication. *P*
International Conf
IEEE Computer Sc
Burgoon, J. K., Stoner
(2003). Trust and
communication. *P*
International Conf
Bok, S. (1978). *Lying:*
New York: Pantheo
Bradner, E. and Mark,
effects on cooperat
Proceedings, Compu
(CSCW 02) (pp. 22
Camden, C., Motley, M
lies in interpersona
preliminary investi
Journal of Speech C
Carlson, J. R. and Geo
appropriateness in
deceptive communi
richness and synch
Negotiation 13, 191-
Carlson, J. R., George,
White, C. H. (2004)
communication. *Gr*
Carlson, J. R. and Zmu
theory and the expe
perceptions. *Acaden*
153–170.
Cornwell, B. and Lund
Internet: involvemer
romantic relationshi
Computers in Huma
Daft, R. L. and R. H. Le
information require
structural design. *M*
554–571.
Daft, R. L., R. H. Lengel
equivocality, media s
performance: implic
MIS Quarterly 11(3),

Berman, J. and Bruckman, A. (2001). The Turing game: exploring identity in an online environment. *Convergence 7*, 83–102.

Bradner, E. and Mark, G. (2002). Why distance matters: effects on cooperation, persuasion and deception. *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work* (pp. 226–235). ACM Press: New York.

Buller, D. B. and J. K. Burgoon. (1996). Interpersonal deception theory. *Communication Theory 6*, 203–242.

Burgoon, J. K. and Buller, D. B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior 18*, 155–184.

Burgoon, J. K. and Nunamaker, J. F. (2004). Toward computer-aided support for the detection of deception. *Group Decision and Negotiation 13*, 1–4.

Burgoon, J. K., Stoner, G. M., Bonito, J. A. and Dunbar, N. E. (2003). Trust and deception in mediated communication. *Proceedings of the 36th Annual Hawaii International Conference on System Sciences* (10 pages). IEEE Computer Society Press: Washington, D. C.

Burgoon, J. K., Stoner, G. M., Bonito, J. A. and Dunbar, N. E. (2003). Trust and deception in mediated communication. *Proceedings of the 36th Hawaii International Conference on Systems Sciences*, Maui, USA.

Bok, S. (1978). *Lying: Moral choice in public and private life*. New York: Pantheon.

Bradner, E. and Mark, G. (2002). Why distance matters: effects on cooperation, persuasion and deception. *Proceedings, Computer Supported Cooperative Work (CSCW 02)* (pp. 226–235). November, New Orleans, LI.

Camden, C., Motley, M. T. and Wilson, A. (1984). White lies in interpersonal communication: a taxonomy and preliminary investigation of social motivations. *Western Journal of Speech Communication 48*, 309–325.

Carlson, J. R. and George, J. F. (2004). Media appropriateness in the conduct and discovery of deceptive communication: the relative influence of richness and synchronicity. *Group Decision and Negotiation 13*, 191–210.

Carlson, J. R., George, J. F., Burgoon, J. K., Adkins, M. and White, C. H. (2004). Deception in computer-mediated communication. *Group Decision and Negotiation 13*, 5–28.

Carlson, J. R. and Zmud, R. W. (1999). Channel expansion theory and the experiential nature of media richness perceptions. *Academy of Management Journal 42*(2), 153–170.

Cornwell, B. and Lundgren, D. C. (2001). Love on the Internet: involvement and misrepresentation in romantic relationships in cyberspace vs. realspace. *Computers in Human Behavior 17*, 197–211.

Daft, R. L. and R. H. Lengel. (1986). organizational information requirements: media richness and structural design. *Management Science 32*(5), 554–571.

Daft, R. L., R. H. Lengel, and L. K. Trevino. (1987). Message equivocality, media selection, and manager performance: implications for information systems. *MIS Quarterly 11*(3), 355–366.

DePaulo, B. M. and Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology 74*, 63–79.

DePaulo, B. M. Kashy, D. A., Kirkendol, S. E., Wyer, M. M. and Epstein, J. A.(1996). Lying in everyday life. *Journal of Personality and Social Psychology 70*, 979–995.

DePaulo, B. M. and Kirkendol, S. E. (1989). The motivational impairment effect in the communication of deception. In J. C. Yuille (ed.), *Credibility assessment* (pp. 51–70). Dordrecht, Netherlands: Kluwer Academic.

DePaulo, B. M., Lanier, K. and Davis, T. (1983). Detecting the deceit of the motivated liar. *Journal of Personality and Social Psychology 45*, 1096–1103.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K. and Cooper, H. (2003). Cues to deception. *Psychological Bulletin 129*, 74–118.

Donath, J. S. (1998). Identity and deception in the virtual community. In M. A. Smith and P. Kollock (eds) *Communities in Cyberspace* (pp. 29–59). New York: Routledge.

Douglas, K. M. and McGarty, C. (2001). Identifiability and self-presentation: computer-mediated communication and intergroup interaction. British Journal of Social Psychology *40*, 399–416.

Ekman, P. (2001). *Telling lies: Clues to deceit in the marketplace, politics and marriage*. New York: W. W. Norton.

Ellison, N. B., Heino, R. D. and Gibbs, J. L. (2004). Truth in advertising? An explanation of self-presentation and disclosure in online personals. Paper presented at the Annual Convention of the International Communication Association, New Orleans, LA.

Feldman, R. S., Forrest, J. A. and Happ, B. R. (2002). Self-presentation and verbal deception: do self-presenters lie more? *Basic and Applied Social Psychology 24*, 163–170.

Forrest, J. A. and Feldman, R. S. (2000). Detecting deception and judge's involvement: lower task involvement leads to better lie detection. *Personality and Social Psychology Bulletin 26*, 118–125.

Galasinski, D. (2000). *The language of deception. A discourse analytic study*. Thousand Oaks, CA: Sage.

George, J. F. and J. R. Carlson. (1999). Group support systems and deceptive communication. *Proceedings of the 32nd Hawaii International Conference on Systems Sciences*, Maui, HI.

George, J. F. and J. R. Carlson. (1999). Group support systems and deceptive communication. *Proceedings of the 32nd Annual Hawaii International Conference on System Sciences* (10 pages). IEEE Computer Society Press: Washington, D. C.

George, J. F. and Carlson, J. R. (2005). Media selection for deceptive communication. *Proceedings of the of the 38th Annual Hawaii International Conference on System Sciences* (10 pages). IEEE Computer Society Press: Washington, D. C.

George, J. F. and Carlson, J. R. (2005). Media selection for deceptive communication. *Proceedings of the 38th Hawaii International Conference on System Sciences*. Big Island, HI.

George, J. F. and Marrett, K. (2004). Inhibiting deception and its detection. *Proceedings of the 34th Hawaii International Conference on System Sciences.*

George, J. F. and Marrett, K. (2004). Inhibiting deception and its detection. *Proceedings of the 34th Annual Hawaii International Conference on System Sciences* (10 pages). IEEE Computer Society Press: Washington, D. C.

George, J. F., Marrett, K. and Tilley, P. (2004). Deception detection under varying electronic media and warning conditions. *Proceedings of the 34th Hawaii International Conference on System Sciences.*

George, J. F. and Marrett, K and Tilley, P. (2004). Deception detection under varying electronic media and warning conditions. *Proceedings of the 34th Annual Hawaii International Conference on System Sciences* (10 pages). IEEE Computer Society Press: Washington, D. C.

Grazioli, S. and Jarvenpaa, S. (2000). Perils of Internet fraud: an empirical investigation of deception and trust with experienced Internet consumers. *IEEE transactions on Systems, Man, and Cybernetics 3,* 395–410.

Grazioli, S. and Jarvenpaa, S. (2003a). Consumer and business deception on the Internet: content analysis of documentary evidence. *International Journal of Electronic Commerce 7,* 93–118.

Grazioli, S. and Jarvenpaa, S. (2003b). Deceived! Under target on line. *Communications of the ACM 46,* 196–205.

Hancock, J. T., Curry, L., Goorha, S., and Woodworth, M. (in press a). On Lying and Being Lied To: A Linguistic Analysis of Deception in Computer-Mediated Communication. *Discourse Processes.*

Hancock, J. T. and Dunham, P. J. (2001a). Impression formation in computer-mediated communication revisited: an analysis of the breadth and intensity of impressions. *Communication Research 28,* 325–347.

Hancock, J. T. and Dunham, P. J. (2001b). Language use in computer-mediated communication: the role of coordination devices. *Discourse Processes 31,* 91–110.

Hancock, J. T., Thom-Santelli, J. and Ritchie, T. (2004a). Deception and design: The impact of communication technologies on lying behavior. *Proceedings, Conference on Computer Human Interaction* (pp. 130–136). New York, ACM.

Hancock, J. T., Thom-Santelli, J. and Ritchie, T. (2004b). What lies beneath: the effect of the communication medium on the production of deception. Presented at the Annual Meeting of the *Society for Text and Discourse,* Chicago, IL.

Hancock, J. T., Woodworth, M., and Goorha, S. (in press b). See no evil: The effect of communication medium and motivation on deception detection. *Group Decision and Negotiation.*

Heinrich, C. U. and Borkenau, P. (1998). Deception and deception detection: the role of cross-modal inconsistency. *Journal of Personality 66,* 667–712.

Herring, S. C. and Martinson, A. (2004). Assessing gender authenticity in computer-mediated language use: evidence from an identity game. *Journal of Language and Social Psychology 23,* 424–446.

Hollingshead, A. (2000). Truth and lying in computer-mediated groups. In M. A. Neale, E. A. Mannix, and T. Griffith (eds), *Research in managing groups and teams, vol. 3: Technology and teams* (pp. 157–173). Greenwich, CT: JAI Press.

Horn, D. B. Is seeing believing? Detecting deception in technologically mediated communication. *Proceedings, Extended Abstracts of CHI'01.*

Horn, D. B., Olson, J. S. and Karasik, L. (2002). The effects of spatial and temporal video distortion on lie detection performance. *Proceedings, Extended Abstracts of CHI'02.*

Horn, D. B., Olson, J. S. and Karasik, L. (2002). The effects of spatial and temporal video distortion on lie detection performance. *Extended Abstracts of the CHI' 02 Conference on Human Factors in Computing Systems* (pp. 714–715). ACM: New York.

Internet Fraud Complaint Center (2003). *Internet Fraud Report.* The National White Collar Crime Center. Washington, D. C.

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

Keyes, R. (2004). The post-truth era: Dishonesty and deception in contemporary life. *New York:* St. Martin's Press.

Knight, J. (2004). The truth about lying. *Nature 428,* 692–694.

Messaris, P. (1997). *Visual persuasion.* Thousand Oaks, CA: Sage Publications, Inc.

Lippard, P. V. (1988). 'Ask me no questions, I'll tell you no lies': situational exigencies for interpersonal deception. *Western Journal of Speech Communication 52,* 91–103.

Metts, S. (1989). An exploratory investigation of deception in close relationships. *Journal of Social and Personal Relationships 6,* 159–179.

Millar, M. and Millar, K. (1995). Detection of deception in familiar and unfamiliar persons: the effects of information restriction. *Journal of Nonverbal Behavior 19,* 69–84.

Miller, G. R. and Stiff, J. B. (1993). *Deceptive communication: Sage series in interpersonal communication, vol. 14.* Thousand Oaks, CA: Sage Publications, Inc.

Mitchell, K. J., Finkelhor, D. and Wolak, J. (2001). Risk factors and impact of online solicitation of youth. *Journal of the American Medical Association 285,* 3011–3014.

Postmes, T., Spears, R. and Lea, M. (1999). Social identity, group norms, and 'deindividuation': lessons from computer-mediated communication for social influence in the group. In N. Ellemers, R. Spears and B. Doosje (eds), *Social identity: Context, commitment, content* (pp. 164–183). Oxford: Blackwell.

Robinson, W. P. (1996). *Deceit, delusion, and detection.* Thousand Oaks, CA: Sage Publications Inc.

Schein, E. H. (2004). Learning when and how to lie: a neglected aspect of organizational and occupational socialization. *Human Relations 57,* 259–273.

Schneider, F. B. (Ed.) (1999). *Trust in cyberspace.* Washington, DC: National Academy Press.

Spears, R., Postmes, T. and Lea, M. (2002). The power of influence and the influence of power in virtual groups: a SIDE look at CMC and the Internet. *The Journal of Social Issues. Special Issue: Social impact of the Internet 58,* 91–108.

Stolfo, S. J., Lee, W., Chan, P. K., Fan, W. and Eskin (2001). Data-mining based intrustion detectors: an overview of the Columbia IDS project. *SIGMOD Record 30,* 5–14.

Turkle, S. (1995). *Life on the screen: Identity in the age of the Internet.* New York: Simon and Schuster.

Utz, S. (2005). Types of deception and underlying motivation: what people think. Social *Science Computer Review 23,* 49–56.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and its implications for professional practice.* Chichester: John Wiley and Sons.

Walther, J. B. (1996). Computer-mediated communication: impersonal, interpersonal, and hyperpersonal interaction. *Communication Research 23,* 1–43.

Walther, J. B. and Parks, M. R. (2002). Cues filtered out, cues filtered in: computer-mediated communication and relationships. In M. L. Knapp and J. A. Daly (eds), *Handbook of interpersonal communication,* 3rd edn. (pp. 529–563). Thousand Oaks, CA: Sage.

Walther, J. B. and Slovacek, C. and Tidwell, L. C. (2001). Is a picture worth a thousand words? Photographic images in long term and short term virtual teams. *Communication Research 28,* 105–134.

Whitty, M. T. (2002). Liar, Liar! An examination of how open, supportive and honest people are in chat rooms. Computers in Human Behavior *18*(4), 343–352.

Whitty, M. and Gavin, J. (2001). Age/sex/location: uncovering the social cues in the development of online relationships. *CyberPsychology and Behavior 4*(5), 623–630.

Zahavi, A. (1993). The fallacy of conventional signaling. *The Royal Society Philosophical Transaction 340,* 227–230.

Zhou, L., Burgoon, J. K., Nunamaker, J. F. and Twitchell, D. (2004a). Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communication. *Group Decision and Negotiation 13,* 81–106.

Zhou, L., Burgoon, J. K., Twitchell, D., Qin, T. and Nunamaker, J. F. (2004b). A comparison of classification methods for predicting deception in computer-mediated communication. *Journal of Management Information Systems 20,* 139–165.

Zhou, L. and Zhang, D. (2004). Can online behavior unveil deceivers? An exploratory investigation of deception in instant messaging. *Proceedings of the 37th Hawaii International Conference on Systems Sciences,* Maui, USA.

Zhou, L. and Zhang, D. (2004). Can online behavior unveil deceivers? An exploratory investigation of deception in instant messaging. *Proceedings of the 37th Annual Hawaii International Conference on System Sciences* (10 pages). IEEE Computer Society Press: Washington, D. C.

Zuckerman, M. and Driver, R. E. (1985). Telling lies: verbal and nonverbal correlates of deception. In A. W. Siegman and S. Feldstein (eds), *Multichannel integrations of nonverbal behavior* (pp. 129–147). Hillsdale, NJ: Erlbaum.