HOW TECHNOLOGY IS REVOLUTIONIZING OUR UNDERSTANDING OF HUMAN COOPERATION

PROF. DR. PAUL J. TAYLOR

UNIVERSITEIT TWENTE.
HOW TECHNOLOGY IS REVOLUTIONIZING OUR UNDERSTANDING OF HUMAN COOPERATION

INAUGURAL LECTURE GIVEN TO MARK THE ASSUMPTION OF THE POSITION AS PROFESSOR OF

HUMAN INTERACTION IN CONFLICT SETTINGS

AT THE FACULTY OF BEHAVIORAL SCIENCES, AT THE UNIVERSITY OF TWENTE ON THURSDAY 20TH JUNE 2013 BY

PROF. DR. PAUL J. TAYLOR

20TH JUNE 2013
Prelude: Regretfully in English

One of my biggest regrets in life (so far) is not having mastered a second language. Opening this lecture with a pithy Dutch colloquialism would likely have done much to gain your attention and perhaps even your high regard. By way of compensation, I do have a lay theory about the nature of language learning. It is very much a lay theory because it has yet to be tested empirically. But it has an interesting implication and so I want to share it with you. I conjecture that there are two types of language learner. One type focuses their learning on ways to provide information and accomplish transactions with others. Their initial learning focuses on phrases such as “My name is” and “I would like.” By contrast, the second type of learner concerns themselves less with substantive communication and more with learning ways to relate to others. Their focus is learning phrases such as “How are you?” and “I like your dress.”

What is interesting about this observation is that it implies a difference in how people seek to cooperate with others. It implies that there are several distinct ways to cooperate, and that these ways are connected in some meaningful way to a speaker’s predilections and goals. Some principally see language as a vehicle for communicating facts, exchanging information, and arguing a point of logic. Others principally see language as a relational tool that allows for the expression and development of liking and trust. These are two very different ways of using language. They are two very different way of interacting with others.

The purpose of this thought experiment is not to emphasize individual differences in interaction style, although these certainly exist, especially when we consider differences across cultures (cf. Giebels & Taylor, 2009). Rather, the purpose is to illustrate the idea that communication serves multiple functions, each achieved by using a different kind of
behavior, and each important to interaction at different times. This is quite a fundamental assertion because it has two implications about the nature of our interpersonal behavior. First, it implies that the way in which a person communicates is related in a systematic way to their motivations and goals. This may seem like an obvious proposition to the non-psychologist, but the link between cognition (i.e., how we think about things) and behavior (i.e., our actions) has not always been as clear-cut as you might assume. Second, it implies that two people using similar language may also have a common set of motivations and goals. That is, there is a positive association between the similarity of language use and the similarity of peoples’ ‘situational models’ of what is occurring. This is important because the alignment of situational models is the precursor to, if not the definition of, human cooperation.

These observations suggest that cooperation may depend on the degree to which each party is able to decipher and align their behavior with that of their partner. In my work I have often tried to introduce this idea by reference to the term ‘the good stranger.’ This term has been used by the US government to refer to military personnel who are adept at gaining cooperation from civilians who might otherwise be antagonistic or distrustful (DARPA, 2011). They have no doubt observed, in their tours of Afghanistan and Iraq, that some personnel are able to respond to others in a way that is particularly effective at allaying fears and encouraging dialogue. This has led them to ask the question: what is it that these ‘good strangers’ do? Of course, the value of answering this question is not confined to military peacekeeping. It instantiates a fundamental question about the dynamics of human cooperation, from the businessman who wants to sell his latest invention, to the newly anointed professor who wants his audience to believe that he has something important to say! In this lecture I am going to argue for the importance of behavioral alignment to human cooperation. I will do so by drawing on data collected from ‘human touch’ technologies, which are changing our understanding of how human cooperation works at an increasing pace.
The Structure to Human Cooperation

In order to explore the nature of cooperation, it is useful to have an understanding of how people communicate, or more specifically, how peoples’ communication differs. Over 50 years of research has sought to identify the connections between the things we say and the motivations that underlie them. While the nuances of such interrelationships are complex, it turns out that there is an elegant overarching structure to the way in which people communicate (Bilsky, Tebrugge, & Webel-Therhorn, 2010; Taylor & Donald, 2007). This structure may be represented graphically as a cylinder (see Fig. 1).

![Cylinder Diagram](image)

*Figure 1. Schematic representation of the structure of communication (first published in Taylor, 2002)*

The spine of this cylinder captures the overall orientation that people take to interaction. An avoidant orientation is characterized by retractions from substantive discussion and a refusal to accept responsibility for events. It usually occurs because of the overwhelming nature of a situation, but it is sometimes used to stonewall an interaction’s progress (e.g., Loving, Le, & Crockett, 2009). A competitive
orientation is characterized by behaviors that attack the other party’s position (e.g., demands, threats) or credibility (e.g., personal attacks), while restoring personal face (e.g., boasting). This orientation is prevalent in conflicts, due in part to what psychologists call the fixed pie bias: a tendency for people to assume that their interests are directly opposed to those of their counterpart (Donohue & Taylor, 2007). Finally, a cooperative orientation is associated with behaviors such as concessions, compliments, and messages aimed at building relationships. A cooperative orientation is, I hope, the mainstay of your day-to-day interactions. Yet, it is not always a good orientation to have. Sometimes cooperating gives a shrewd counterpart room to take advantage, or allows the cooperator’s viewpoint to be overshadowed by the messages of a single, competitive ‘hawk’ (e.g., Steinel, De Dreu, Ouwehand, & Ramirez-Marin, 2009).

The distinction between avoidant, competitive, and cooperative orientations captures a person’s overall approach to interaction, but it does not speak to the goal that they pursue while adopting one of these orientations. Each and every one of our utterances is delivered with a purpose, such as to provide information, express empathy, or boast about a personal achievement. These possibilities are referred to as motivational goals, and they are captured in the cylinder model under three broad categories of Instrumental, Identity, and Relational (see Fig. 1). On some occasions a speaker seeks to achieve a substantive goal, such as gaining a commodity, providing information, or reaching agreement on a debated issue. They typically do this by using instrumental behaviors such as demands and offers, threats of action, and concessions. At other times, a speaker may use behaviors such as justifications, repeated interruptions, profanity, and insults. These behaviors are less focused on an instrumental goal and more focused on attacking their interaction partner’s identity while perhaps simultaneously promoting personal esteem. Finally, a person may use behaviors such as apologies, excuses, humor, interruptions and reassurance. These messages are characterized
by their focus on managing the affiliation and power dynamic between the parties, and hence they are relational in their goal.

The distinction between instrumental, identity and relational goals combines with the distinction of orientations to form nine different kinds of message. These are often referred to as communicative frames. At any one time, individuals tend to adopt an avoidant, competitive or cooperative orientation to interaction and pursue either an identity, instrumental or relational goal, with varying degrees of intensity. So, for example, a couple in a custody dispute may yell abuse and insults as they compete over identity issues that stem from their beliefs about how the other acted within the relationship. But they may revert to more cooperative, instrumental behavior such as compromises and promises when discussing what’s best for their child.

What’s quite remarkable about the cylinder structure is how universal it is as a description of the communication that occurs in many contexts. It is possible to take a recording of an interaction and quite literally plot each speaker moving in and out of these different frames over time. This is not to suggest that any one utterance or message is exclusively attached to one of these frames. Language is more dynamic than that. But what it does suggest is that the process of cooperating with others is structured around these different emphases. This has been shown to be true for conversations as diverse as students discussing a project, company teams discussing how they can work more safely, couples fighting over the custody of their child, and police negotiators trying to talk a hostage taker into surrendering (see Taylor, in press, for a review). Experimental comparisons have further shown that the use of these nine frames is affected by person factors such as culture (e.g., Adair & Brett, 2005) and individual differences (e.g., Park, & Antonioni, 2007), and situational factors such as power differences (e.g., Giebels, De Dreu, & Van de Vliert, 2000) and emotions (e.g., van Kleef, De Dreu, & Manstead, 2004). Over the last few years, research has even begun to find neurological correlates of these behaviors. For example, oxytocin dampens the amygdala’s response to threat (the amygdala is a small
brain region believed to be involved in emotion processing), such that high levels of oxytocin lead people to processes the social situation in a way that leads them to be more likely to adopt a cooperative orientation. Our social processing may metaphorically and even quite literally be structured in the way that is illustrated by the cylinder.

The distinctions portrayed in the cylinder model have helped clarify our thinking about a range of applied problems. One example concerns the nature of messages that are designed to promote terrorism. As the recent tragic stabbing of a British soldier in Woolwich, England, illustrates, the world continues to have a problem with terrorism, or more specifically, with those who are willing to use violence to pursue a worldview that they have come to believe. I use the phrase ‘come to believe’ because evidence in this area suggests that people who engage in violent extremism do so after a period in which they are influenced—sometimes referred to as being radicalized—into a particular system of belief (Jacques & Taylor, 2013). In response to such actions, governments around the world have sought to understand what messages lead people to see violence as a legitimate action. In the UK, Sheryl Prentice, Paul Rayson and I addressed this question by examining the language used in online messages that endorsed violence (e.g., Prentice, Taylor, Rayson, Hoskins, & O’Loughlin, 2011). Our goal was to learn how those who are most active in this area approach framing an argument for violence. What do they see as the most effective social leavers?
To answer this question, we collected 250 examples of messages promoting violent extremism and looked for the occurrence of statements that fit into the frames of identity, instrumental and relational behavior. We did this both manually, by reading and marking the occurrence of each frame in the messages, and automatically, by using statistical language models to classify each statement. The result was evidence for various kinds of statements that relate to each of the three frames (see Fig. 2). For example, on some occasions, the messages contained instrumental statements that argued logically and assertively for the reasons why violence was appropriate (i.e., persuasion, ex-changing, and direct pressure). On other occasions, the messages

<table>
<thead>
<tr>
<th>Motivational Frame</th>
<th>Tactic</th>
<th>Brief Definition</th>
<th>Example</th>
<th>Frequency</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Pressure</td>
<td>Exchanging</td>
<td>Pressure tactic that includes commands, demands, and threats</td>
<td>“Raise your arms and fight to escape from this humiliation and shame!”</td>
<td>2526 (65%)</td>
<td></td>
</tr>
<tr>
<td>Instrumental (argument-related)</td>
<td>Persuasion</td>
<td>Explicit or implicit promise of a reward or tangible benefits for acting</td>
<td>“When the enemy targets our women and children we should forget theirs”</td>
<td>307 (1%)</td>
<td>9209 (23%)</td>
</tr>
<tr>
<td>Relational (audience-related)</td>
<td>Upright Appeals</td>
<td>Argument that explains reasons, or presents evidence, in support of a position</td>
<td>“Another obstacle is the need for advanced means of resistance to counter the occupation…”</td>
<td>5382 (16%)</td>
<td></td>
</tr>
<tr>
<td>Relational (audience-related)</td>
<td>Social Proof</td>
<td>Use of authority-based comparisons, either religious or political</td>
<td>“Shykh Uhayyimen says, if the enemy kill our children then we are allowed to kill their children”</td>
<td>5319 (13%)</td>
<td></td>
</tr>
<tr>
<td>Relational (audience-related)</td>
<td>Moral Proof</td>
<td>Social comparisons, typically to groups who support a viewpoint</td>
<td>“And we people in Palestine totally reject such a description along with our Arab peoples…”</td>
<td>6457 (13%)</td>
<td>22364 (56%)</td>
</tr>
<tr>
<td>Relational (audience-related)</td>
<td>Activation of commitment</td>
<td>Comparisons that justify a position or highlight an out-group’s immorality</td>
<td>“If a herd of dogs had suffered a tenth of what the Palestinians in Gaza have suffered…”</td>
<td>11556 (28%)</td>
<td></td>
</tr>
<tr>
<td>Identity (speaker-related)</td>
<td>Inspirational appeals</td>
<td>Messages that remind listeners of a debt owed or social identity</td>
<td>“It is incumbent upon us to use all our resources to confront the attack on our ummah”</td>
<td>2292 (6%)</td>
<td></td>
</tr>
<tr>
<td>Identity (speaker-related)</td>
<td>Liking</td>
<td>Request that appeals to positive self-feeling, altruism, or esteem</td>
<td>“And we will be the coming power insha’Allah”</td>
<td>4036 (10%)</td>
<td>8136 (21%)</td>
</tr>
<tr>
<td>Identity (speaker-related)</td>
<td>Liking</td>
<td>Positive message that puts listeners in good frame of mind (e.g., Injilisation)</td>
<td>“For those who asked that I reconsider my view on this, I promise I will review it again”</td>
<td>1796 (4%)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Definitions and examples of nine persuasion behaviors as a function of motivational frame.
contained statements that were primarily relational, combining social and moral comparisons with references to the thoughts of authority figures (i.e., upward appeals, social and moral proof). Finally, some messages contained statements that focused on the identity of the person that they were trying to persuade, emphasizing their role in a collective ‘Ummah’ and arousing enthusiasm by appealing to their self-esteem (i.e., activation of commitment, inspirational appeals and liking).

Although the occurrence of this range of behaviors is interesting, where these data are particularly revealing is when we consider the relative use of the different frames. As can be seen in Fig. 2, over 50% of the statements within the messages (22,364 of 40,917 coded message) were relational in emphases. The authors of these messages appear to believe that the most effective way of inciting violence is to highlight issues of social morality, the behavior of social in-groups, and the opinions of authority figures. This is interesting because it is at odds with what typically characterizes persuasive messages, which is a focus on ‘rational persuasion’ and an instrumental discussion of the pertinent issues. Indeed, it is at odds with what we found when conducting the same analysis on counter-messages produced by officials associated with Western governments. These messages, which are aimed at dissuading people from violence, were primarily instrumental in nature. Thus, if these data are to be believed, such instrumental messages are engaging on a different wavelength to the messages that are being used to promote violence.

When doing this kind of research, a single individual’s use of the different message frames is called a ‘profile.’ We each have a profile, which reflects our idiosyncratic way of communicating with others. For some people the profile is largely set in stone, which is to say that they use the same kinds of messages regardless of who they are interacting with, or what situation they are in. For others their profile adapts depending on the context in which they are communicating. In our violent extremism research, for example, we found considerable
variation in communication profiles, with authors such as Ayman al-Zawahiri (an Al Qa’ida figurehead) showing considerable flexibility in the persuasive leavers that he used across messages (Prentice et al., 2011). Similarly, and perhaps more usefully, message profiles were found to correlate with the degree to which a message resonated and was influential online. This means that, with a clever bit of statistical modeling, it is possible in principle to use an analysis of communication profiles to identify the messages that authorities should prioritize in their investigations. They could pay particular attention to certain kinds of ‘risky’ profiles, in the same way that a medical doctor pays particular attention to certain patterns of symptoms. In an era of resource cutbacks, the ability to efficiently measure the risk associated with different messages, and the authors behind those message, is likely to be valuable.

The cylinder model, then, provides a broad model of the structure of human communication from which we can begin to consider the process of cooperation. I use the term process of cooperation because cooperation is not just something that happens at a turn of a switch, but it is something that emerges from the dynamic, back-and-forth of behavior among parties. Understanding the nature of this process, rather than the impact of individual behaviors examined outside the context of the process, remains an important challenge for the field. It is important both theoretically, because it speaks to the social basis of the human condition, and practically, because it provides the basis for being able to train people to be ‘good strangers.’ So, what does cooperation look like when observed through the lens of the cylinder model? Across a range of studies, what has emerged as critical above all other factors is whether or not speakers match one another’s use of communicative frames for sustained periods. This is known within the field as alignment (Garrod & Pickering, 2004). It reflects a mutually reinforcing circuit in which the coming together of communicative behavior leads to an alignment of our cognitive ‘situational model’ of the world, which leads to greater alignment in language, which leads to greater alignment of our model, and so on.
There are many examples of how the alignment of language and frames is connected to cooperation. For example, in analyzing the protracted and volatile interactions of hostage negotiations, Tom Ormerod, Emma Barrett, and I found that frame matching increased over time in negotiations where the perpetrator surrendered, but marginally decreased over time in negotiations that ended violently (Ormerod, Barrett, & Taylor, 2008). The increase in alignment observed in the successful negotiations was four-times the magnitude of the decrease observed in the unsuccessful negotiations. The process of cooperation is, at least in part, the process of getting in sync.

Arguably the more interesting question to ask of this hostage data is why the alignment between negotiator and perpetrator occurred in some negotiations but not in others? To a large extent, the divergence in matching appeared to be due to how the negotiators handled the periods of dialogue in which they were not aligned with the perpetrator. During these transitional periods, successful negotiators tended to adopt a less dominant role in the negotiation, switching their personal framing to match that of the perpetrator more frequently than they did in the non-transitional periods. They achieved this accommodation of the perpetrator’s perspective by reducing the amount they spoke by over 40%! That is quite a dramatic change in behavior. It presumably allowed them to gather social cues and focus on making sense of the perpetrators’ communication before reengaging them with an appropriate frame. By contrast, negotiators of unsuccessful incidents tended to show less change in their behavior during transition periods, which meant that they did not facilitate the re-alignment of frames.

The 1,000GB Elephant in the Room

From what I have covered so far, it is easy to conclude that cooperation is best understood as an intricate tussle of matched and mismatched frames, in which early, turbulent periods of interaction are replaced by aligned dialogue as speakers entrain to one another’s
messages. Such a process may be driven by a deep analytic processing or fast heuristic judgments of social information, or most likely by both (Petty & Cacioppo, 1986). But, this body of research suggests that it will, at all times, be driven by a conscious processing of the higher-level meaning of the other party’s behavior.

For a long time, researchers have realized that there’s something unsatisfactory about this account. That dissatisfaction relates to what might be described as a 1,000GB elephant in the room. Give or take, to encode, interpret and reproduce in a digital form how I have behaved in this lecture so far would take about 1,000GB of computer memory. That’s roughly the equivalent of what could be stored on the earliest Acorn BBC model B computers if everybody in The Netherlands owned one! As humans, however, we ‘analyze’ such quantities and qualities of data effortlessly, and often without conscious awareness. This implies that our interpersonal behavior relies on more than an explicit, conscious processing of social information. There must be more layers to the cooperation problem than the one we have examined so far.

While researchers have appreciated this implication for a long time, it has until recently been difficult to do more than theorize about what those other processes might be. This is a due, in large part, to the methods that have been available to researchers. Traditionally, a researcher wishing to study interpersonal behavior endured long and painful hours in front of a computer while they dutifully watched and re-watched video footage to spot behaviors: “A gesture here. A cough there. And, oh wait, did he glance up to the right?” These efforts—by coders as they are known in the trade—have provided rich data and important insights into social interaction. However, the rapid advance of technological sensors has changed what is possible. A coder may now be replaced by software that uses pattern recognition algorithms, such as n-gram models for language and blob analyses for nonverbal movement, to identify and record patterns that are scarcely discernable to the coder. Similarly, we now have technology, sometimes no larger than a matchbox, that captures every flinch and twitch of a limb in
a way that a coder would take years to record, and even longer to record without error. When these technologies are brought together, it becomes possible to measure a range of dynamics of interpersonal behavior with exciting accuracy and important consequences.

The Role of Alignment

One example of what technology has allowed us to examine is a process known as language style matching. Over time, speakers come to match one another’s lexical, syntactic and semantic choices, and through this process they begin to converge on the same situational framing (Garrod & Pickering, 2004). Cooperation thus occurs not only from high-level alignment of orientation and goals, as I explored in the previous section, but also from a basic coordination of language use. A bottom-up process as well as a top-down one. The link between language and these higher up cognitive processes can be quite striking. To give just one example from the literature, waitresses who repeat back to customers their order using equivalent language receive, on average, more tips than those who do not (Van Baaren, Holland, Steenaert, & Van Knippenberg, 2003).

The development of computer-based text analytic methods has allowed researchers to measure a person’s language style through the rapid assessment of their use of key categories of words. It is possible to take lengthy interactions and track the degree of language style matching across utterances; indeed, as the accuracy of speech-to-text engines increases, so it has been possible to develop software than goes straight from speech to analysis. By far the most utilized measure of language style matching, proposed by Jamie Pennebaker and colleagues, draws on word categories known as function words. These include word groups such as adverbs (e.g., very, well), articles (e.g., a, the), auxiliary verbs (e.g., am, have) and pronouns (e.g., I, us). They are interesting because they make up over half of our daily speech and yet they are not tied to the discussion of particular content and they are often used in a way that is outside
of people’s conscious awareness (Ireland & Pennebaker, 2010). By using a measure of language style matching (LSM) that is based on function words, a number of studies have shown a positive relationship between LSM and interaction outcome. For example, cooperation in small group activities goes hand-in-hand with greater matching of function words (Gonzales, Hancock, & Pennebaker, 2010), and so does success in hostage negotiation (Taylor & Thomas, 2008). A recent illustration of this process occurring comes from a study, led by Beth Richardson, on the coordination of interrogator and suspects’ verbal behavior over time (Richardson, Taylor, Snook, Conchie, & Bennell, submitted). Fig. 3 plots the relative LSM between interrogator and suspect over time in sixty-four interrogations, as a function of whether or not the suspect ended up confessing. Values below .00 in Fig. 3 indicate that the interrogator was matching the suspects’ language style, and values above .00 indicate that the suspect was matching the interrogators’ language style. As one might

Figure 3. Turn-by-turn LSM difference score as a function of outcome. Positive values indicate a suspect matching the interrogators’ language style. Negative values indicate an interrogator matching a suspects’ language style.
expect, over time there are evident ebbs-and-flows in the nature of matching across the parties. However, these localized fluctuations add up to an important aggregate difference in LSM across confession and no-confession interrogations (as shown by the regression lines in Fig 3). Interrogations ending in a confession were characterized by a linear increase in the degree to which the suspect matched the language of the interrogator, while no-confession interrogations were characterized by a sustained oscillation between interrogator-led and suspect-led matching. In the confession interactions, then, the suspect was increasingly cooperating with his or her interrogator.

Another example of language style matching comes from research that sought to develop an innovative way of identifying insiders (Taylor, Dando, Ormerod, Ball, Jenkins, Sandham, & Menacere, in press). An insider is somebody who misuses his or her legitimate access and knowledge of an organization for personal gain. Usually these offenders provide significant information to an outsider in exchange for a financial reward. For example, in 2001, Robert Hanssen, a FBI Supervisory Special Agent with 25 years’ service, pleaded guilty to spying for the Soviet and Russian intelligence services for over 22 years. During this period, he provided his handlers with significant information about U.S. counterintelligence operations, including the identity of three Russian double agents who were secretly helping the US. He did this all in exchange for money.

The act of conducting this kind of attack is characterized by cognitive and social challenges that may affect an offender’s day-to-day work behavior. Insiders must keep up the quality of their regular work while simultaneously achieving their insider goals, and they must do so without raising the suspicions of their coworkers. This requires not only deceitful behavior but also the constant monitoring of coworkers reactions. These challenges may leak into the day-to-day work behavior of an insider in several ways, but notably for our purposes, into the naturalness of the their interpersonal behavior with coworkers. Specifically, perhaps the degree to which insiders matched
the language style of their coworkers will be significantly different from what is typical across coworkers, since they may unintentionally be less cooperative and they may pay less attention to their relationships with colleagues.

To test this idea, we ran a series of one-day work simulations in which 3 teams of 4 employees worked together in four, one-hour sessions in which they were required to assimilate information, problem solve, and make inferences to determine a solution. The first of the four sessions was a ‘control’ to get everybody used to working within the new environment. After the first session, we approached one person in each team and offered them £20 to sneak out some information for us. We then continued to offer them further money in the third and fourth sessions in exchange for information that was more complicated to obtain. Nobody refused to undertake this task. Perhaps more worrying, only two failed to accomplish it!

What these participants did not know is that during the sessions we were examining the degree of LSM shown within their work emails. We observed exactly the decrease in LSM that we predicted. The insiders showed a reduction in the extent to which they matched the language style of coworkers, which meant that their matching was lower than the average level of LSM found among non-malicious coworkers. This reduction in LSM, which suggests an inadvertent social distancing by the insiders, increased over time, such that by the end of session four, it was possible to use this metric in a statistical model to identify 92.6% of insiders (the AUC = .959). This is exciting because the LSM metric utilizes behavioral data that is quite separate from the existing, resource-oriented technologies for catching insiders, which work by detecting anomalies in document access and usage (Thompson, 2004).

It is clear from these studies that LSM is associated with cooperation, but is it the cause of cooperation, or is it the product of people aligning their communicative frames? This is a fundamental question because it is seeking to determine whether our behavior is the result of top-down cognitive intervention, or whether it is, at least
in part, the product of some more basic, bottom-up response to others’ behavior. In recent unpublished work, Stacey Conchie and I attempted to answer this question by examining four levels of behavioral alignment in a set of crisis negotiations (Taylor & Conchie, 2009). Specifically, we measured the extent to which negotiators and perpetrators showed lexical matching (i.e., common word choice), syntactic matching (i.e., common structuring of language), semantic matching (i.e., common matching of meaning) and situational matching (i.e., common framing of the situation) over time. By examining these four levels of matching across six periods of interaction, we showed that early periods of interaction were primarily driven by alignment of basic lexical and syntactic form, rather than any coordination of situational frames. The alignment of the higher language levels of semantic and situational frames only came to the fore during later stages of interaction, when negotiators began to deal with the substantive issues surrounding the crisis. This is some of the first evidence to suggest that cooperation emerges from a bottom-up process of alignment as much as it does a top-down coming together of cognitive representations of the interaction.

Nonverbal Alignment

So far in this lecture I have mainly been talking about dialogue, but another example of where technology has revolutionized our understanding is in relation to nonverbal behavior. The process of nonverbal mimicry may be defined as movements by one person that coincides with the timing and rhythm of the movements of their interaction partner. For example, behavioral mimicry may take the form of discrete movements such as matched touching of the face (Stel et al., 2010), or more continuous patterns of behavior such as mutual changes in posture (Cappella, 1997). This mutual coordination of behavior typically occurs unconsciously and it is associated with increased cooperation and liking (Chartrand & van Baaren, 2009).

To capture nonverbal mimicry, my colleagues and I have used wireless motion capture sensors that measured the similarity in
interviewer and suspects’ arm, head and torso movement over the course of an interaction.\(^1\) There are various types. Our early worked utilized Sparkfun’s WiTilt sensors, a matchbox sized package of accelerometers that was available from an amateur electronics store and designed for a very different purpose than measuring nonverbal behavior. However, with a bit of tweaking and some software development, it was possible to obtain from these devices 120 per second measurements of the extent to which, and direction in which, the device moved in space. This made it possible to measure a person’s behavior over time by strapping the WiTilts onto various limbs and recording the output. We then compared these movement data across participants in order to measure nonverbal mimicry.

Things have developed since then. Today we tend to use a technology known as Xsens MVN, which was, coincidentally, first developed here at Twente University. The Xsens MVN utilizes a series of sensors distributed across the main limbs and joints of the body, which are worn using either a suit or a set of straps. The distribution of sensors tracks movement across the entire body and allows for almost perfect recording and reproduction of movement. Indeed, this is the technology that they use to make animated films such as the recent comedy Ted. To measure nonverbal mimicry, then, we use two (or more) Xsens MVN suits, capturing the relative movement of each individual within the scenario, and comparing this behavior to establish a measure of similarity. As an aside, it is enlightening to record and watch your own nonverbal behavior using the Xsens MVN system, since our posture and gestures are not something we are typically aware of doing. The Xsens MVN system can thus be useful in training because it allows those tasked with being a ‘good stranger’ to review how they present to others.

Before describing some of our research on nonverbal mimicry, it is worth saying just a little about how to measure mimicry in this context, since it is far from straightforward and raises some fundamental, largely

---

\(^1\) In this work I have often referred to the behavior as nonverbal coordination rather than nonverbal mimicry, in order to recognize the fact that electronic measurement can determine whether the limb has moved in the same direction, but not technically whether the same qualitative gesture was mimicked.
unanswered questions about how humans perceive the movement of others. When people match the behavior of others, they do not do so instantaneously. Nor do they do so in a way that is delayed or lagged in a constant manner. This makes it necessary to adopt a statistical methodology that measures the alignment of individuals’ behavioral streams in a way that accommodates an inconsistent, though not necessarily random, lag in time. The solution comes in the form of dynamic time warping (DTW). DTW takes data on the movement of the same limb in two people, let’s say the vertical movement of their right arm, and examines the extent to which they match one another (i.e., correlate). However, DTW does this while allowing each of the streams to contract or stretch a little. The DTW algorithm operates iteratively by contracting and stretching the behavioral streams until it finds the maximum overlap. The relative degree of overlap provides a measure of mimicry that can be compared across cases and conditions.

The DTW technique, while a useful step forward in what is possible, makes a number of assumption that highlight the unanswered questions that I mentioned. The most obvious of these questions relates to gesturing. If I gesture with my left arm, do we consider your behavior to be mimicry when you gesture with your left arm or when you gesture with your right arm? The former would suggest an imitation that is based on a mirroring of anatomical movement, while the latter would suggest a system that is based more on looking into a mirror (cf. Press, Ray, & Hayes, 2009). For me the answer to this question has important implication for measurement, since I want to ensure that we measure nonverbal mimicry appropriately. However, the answer also has far reaching implications for neuroscientists who are interested in understanding how our brain is wired to perceive human interaction. Thus, this technology is shaping our ability to answer this kind of question using research designs that are more naturalistic than is possible with current brain imaging technologies.

Back to cooperation. One of the established things about nonverbal mimicry is that it is mainly an automatic behavior; it occurs largely
outside of our conscious awareness. With something like nonverbal behavior, which can come under conscious control, the automatic aspect of the behavior becomes prominent when we are under cognitive load. That is, when we are in situations that require us to think hard and remember things (i.e., high cognitive load), our capacity to monitor our behavior reduces because we do not have the capacity to manage the cognitive demand. In relation to cooperation, the one circumstance where people are known to experience significant cognitive load is when they are being uncooperative by lying. Lying takes great cognitive effort. It requires us to imagine a hypothetical account, suppress our true memory of what really happened, control our tendencies to appear nervous, and perhaps even monitor the behavior of the listener to determine whether or not they are suspicious. It stands to reason, then, that when we are lying we may find ourselves becoming more reliant on our automatic nonverbal behavior. As a consequence, changes in the degree to which nonverbal mimicry occurs may reflect an unconscious response to the cognitive load associated with lying.

In turns out that this prediction largely holds true. For example, Sophie van der Zee and I examined the impact of telling truths, easy, and difficult lies on the extent of nonverbal mimicry between interviewer and interviewee. In such experiments it is important to ensure that interviewees have a broadly similar episodic memory of an event to recall. Accordingly, our interviewees met and chatted with another participant—who was actually a confederate working for us—before being instructed by Sophie to complete a wooden puzzle. Half way through this task our confederate noticed the instructions for the puzzle were left on a table, and she proceeded to persuade the interviewee to cheat. On discovering the cheating, Sophie made it clear that the interviewee could not disclose this behavior to the interviewer, and she also had them construct a story about playing the board game Cleudo, since there was no time remaining to play the actual game. At interview, our interviewees were required to tell the truth about the conversation, conceal their cheating during the
puzzle task (easy lie), and convince the interviewer that they did in fact play Cluedo (difficult lie). Interviewers asked standard questions of interviewees, but half requested the game experience be described in reverse order, which is a technique known to increase the experienced cognitive load, particularly for liars (very difficult lie; Vrij et al., 2008).

By using technology to track nonverbal behavior, we were able to examine the extent to which mimicry was affected by deceit. Our analysis found an impressive, positive correlation between the degree of imposed cognitive load and the amount of behavioral mimicry. The more difficult the lie, the more the interviewee and interviewers’ nonverbal behavior was coordinated. Moreover, this effect was difficult to undo. In a subsequent study, we directed interviewees to pay attention to the nonverbal behavior or the interviewer, on the assumption that this instruction would focus their attention on nonverbal behavior and lead them to direct resources to monitoring and controlling their behavior. But the instruction had no impact. On average, interviewees did not bring their behavior under conscious control and they continued to mimic the interviewer more when they were lying.

Those who are practically minded may be concerned that full Xsens MVN suits, while good for research, will be less practical when it comes to measuring the behavior of a real suspect. However, the same technology that drives the measurements within an Xsens MVN suit are now in many mobile phones and, perhaps more importantly, they are readily integrated into devices that are no larger than an employee’s name badge (Pentland, 2010). While the resolution of such devices is less sensitive and accurate than the Xsens MVN, and this inevitably limits the precision with which mimicry can be measured, their sensitivity is sufficient for capturing broad changes in movement and movement mimicry. This not only makes them an attractive compromise for use in applied environments, but their relative low cost makes it possible to consider behavior on a larger scale, such as within a group or across a whole organization. A number of researchers have used these devices to record behavior in large groups, finding that nonverbal activity and
nonverbal mimicry are important predictors of cooperative outcomes such as task performance and sales pitch success (Pentland, 2010). For example, using such devices, David Ellis, Charlotte McClelland, Stacey Conchie and I having demonstrated that verbal and nonverbal mimicry predicted performance on a group problem-solving task, and that these dynamics were predictably influenced by discrete events such as the loss of a team member or the introduction of erroneous information.

**The Role of Social Norms**

Thus far in this lecture I have examined the nature of the interpersonal process, showing that an alignment of situational frames and verbal and nonverbal behaviors both play a role in the emergence of cooperation. In doing so, however, I have ignored discussing the impact of the wider social context. Many of the most critical questions of cooperation are questions about groups, such as families, communities, and organizations. Indeed, many psychologists have argued that third parties play a special, and not particularly positive, role when it comes to cooperation and conflict. They have argued that the social group leads to a deregulation, or deindividuation, of a person’s behavior (Zimbardo, 1969). When group size increases, they argue, the anonymity of the crowd leads a person to decrease their self-evaluation and self-awareness, and under these conditions antisocial and violent behavior become more permissible. The proponents of such an account point to riots and tragedies such as Abu Ghraib as examples of this deregulation.

But are we so prone to violence? As social beings, is it really the case that being social is the one thing that leads us away from cooperation?

Mark Levine, Rachel Best and I had the unique opportunity to study this question by examining the occurrence of night-time violence as recorded by closed-circuit television surveillance (CCTV) of public spaces (Levine, Taylor, & Best, 2011). It is often said that every country is the first in something, and the UK is cited as having the dubious honor of being first for the number of CCTV per head of population: 1 per every 14 people. One, presumably unintended, advantage of
CCTV technology is that it provides a unobtrusive mechanism for examining the nature of third party involvement in real—as opposed to experimentally contrived—incidents of violence. We took advantage of this opportunity by obtaining recordings of 312 people involved in 42 incidents, some of which were severe and some of which ended with minimal violence. We then identified the perpetrator, the victim and the third parties in each clip, and coded their behaviors as either escalatory (e.g., punches, kicks) or conciliatory (e.g., blocking contact, open-hand gestures). This provided a sequence of escalatory and de-escalatory acts across the duration of the incident, which allowed us to examine the role of third parties in fueling or extinguishing the aggression between protagonist and victim.

Our results came as a surprise to those who link the existence of third parties with a greater likelihood of violence. As group size increased from 2 to 12 third parties, the average number of escalating behaviors within an incident did increase slightly, as proponents of deregulation would predict. However, the average number of conciliatory behaviors also increased, and this increase was greater than that observed for escalating behaviors. Thus, third parties tended to de-escalate rather than escalate violence, and they did so more frequently as they increased in number.

But how precisely were third parties encouraging cooperation within the incidents? To explore this question, we examined the nature of the behaviors that occurred in the three acts that followed each instance of a perpetrator’s aggressive behavior. We examined when and how third parties intervened, and how their behavior collectively shaped the future trajectory of the conflict. Our analysis revealed an important group-level dynamic about the nature of cooperation. The occurrence or otherwise of violence was a consequence of how many third parties intervened. If the three interventions were made by the same person, even though other third parties were present and could intervene, then the fourth act in the sequences of behaviors was more likely to be escalatory than conciliatory. If two different third parties made the three interventions, then
regardless of the combination, no significant preference for conciliatory or escalatory intervention was observed in the fourth act. However, critically, when the first three interventions involved three separate third party actors, then the fourth turn was significantly more likely to be conciliatory than escalatory. In other words, collective and coordinated group intervention is more likely to lead to nonviolent outcomes than are repeated actions from the same individual. This suggests that cooperation also emerges from an alignment of behavior with the social majority.

**Epilogue: Alignment and Interact Theory**

In this lecture I have argued that three forms of behavioral alignment—the behavioral, cognitive and social—are critical to how we manage our interactions. Such alignment simplifies the problem of 1,000GB into something that is far more tractable. It works because it is self-reinforcing—each of the layers of behavior complement one another—and it is self-correcting—there is redundancy in the system because signals from layers can moderate the input of other layers. Thus, alignment is how we make interaction look easy.

The astute amongst you, however, will realize that my case for alignment comes without an explanation for why it is that alignment, rather than some other condition, that links behavior to cooperation. To that question I leave you with just the beginnings of an answer: familiarity. Through countless hours of experience our social selves become accustomed to certain sequences of behavior occurring in certain contexts. It is this familiarity that underpins our normative responses in different contexts; since the behavioral sequences that typically occur across contexts can be different, so our normative responses in each context can be different (Mischel, 2004; Taylor & Donald, 2003). Alignment works because it establishes a localized set of behavioral sequences that are understood by both parties. This makes interacting with the other party familiar and predictable, which in turn gives us the confidence to trust and cooperate. At least that’s my hypothesis.
Acknowledgements

It is a pleasure for me to conclude this lecture by acknowledging some of the people who have contributed to my development and worked with me on the new challenges of our discipline. I owe a great deal to the relaxed and supportive PhD environment that Ian Donald created for me at Liverpool, where, together with my longstanding collaborators Craig Bennell, Louise Porter, and Brent Snook, I was able to explore ideas and try things out without the fear of falling too far. That type of environment can be hard to work in when you are young and looking for direction, but it breeds independence and creativity, and I wouldn’t change a thing.

I am now the one who orchestrates the research environment of my own lab, though, ironically, I continue to be reliant on those around me. The hard work and creativity of my past and present MSc and PhD students, postdoctoral colleagues, software engineers, and administrator is humbling, and it makes work life fun. Take a moment to look around the room, for amongst us are, I suggest, the new generation of researchers in this area.

I am particularly indebted to a number of practitioners who have had a profound influence on me personally and on my career. Some of them are in this room. They act with modesty but they are the real masters of our discipline. They have taught me many things, but I will share one that has particular Dutch resonance. Working to deliver evidence-based, behavioral science solutions to practical problems taught me that there is a footnote to Kurt Lewin’s famous saying “There is nothing so practical as a good theory.” The theory does have to be related to something that occurs in the real world to be practical.

I have a new set of colleagues at Twente, both in the Department of Conflict Risk and Safety and the Research Group in Human Media Interaction. They are superb, inventive, and nice; a combination that is not all that common in academia. I will mention just two by name: prof. dr. Karen van Oudenhoven-van der Zee, thank you for allowing me to be part of this wonderful Faculty and University.
It is exciting and stimulating for me to be part of a place where everybody seems to be doing the type of research that interests me!; and, prof. dr. Ellen Giebels, without whom I would not be here, not simply because she is my Head of Department but because we have had an enjoyable and fruitful research collaboration. I recall that our first meeting required some adept sense-making and nonverbal mimicry to avoid an awkward outcome. Long may we stay in sync.

It has been wonderful to share this lecture with my family. My parents are both teachers, and I remember them encouraging me to pursue whatever career gave me enjoyment, so long as it wasn’t teaching! Well I’ve kept to the spirit if not the letter of that advice. I am graced with doing something that I enjoy, and I owe this to their unwavering support and love over the years.

A final word for Stacey. When people are in situations of particular importance they tend to resort to type. So, I have gotten you a present. Well, when it comes to cooperation, as you know, I’m an instrumental type.

Ik heb gezegd.
REFERENCE


