

Chapter 16

Social Behaviour in Police Interviews: Relating Data to Theories

Merijn Bruijnes, Jeroen Linssen, Riëks op den Akker, Mariët Theune,
Sjoerd Wapperom, Chris Broekema, and Dirk Heylen

16.1 Introduction

Automatically generated social behaviour is important for human-like interaction with virtual agents (Vinciarelli et al. 2012). Models of human behaviour can be used to make the behaviour of artificial agents more believable to humans (Steunebrink et al. 2012). Believable agents are being employed in systems for training the skills required for successfully conducting negotiations (Swartout 2010) or job interviews (Vaassen and Wauters 2012). While the interaction with virtual agents may elicit some degree of learning by itself, the most important part of the experience lies in the reflection on what has happened (Koops and Hoevenaar 2013). Along the same lines, *explainable artificial intelligence* (Core et al. 2006) advocates the use of virtual agents that can explain their reasoning. These explanations of actions taken by the virtual agents could be used to improve the user's learning process, especially when the virtual agents clarify their actions in terms of the theories a user has to understand. A simple example of such a clarification might be an agent saying "I started shouting because your competitive stance made me angry."

In this paper, we focus on the interaction between police officers and suspects in police interviews. The long-term goal is to model a virtual suspect that can be used in an application for the training of police students. For this, we are developing a computational model that lets a virtual suspect select the behaviour that is most appropriate. The actions of the user are sensed and interpreted to form meaning, for example "the user is angry". Our computational model then uses these interpretations to form a "mental state" of the virtual suspect, for example

M. Bruijnes (✉) • J. Linssen • R. op den Akker • M. Theune • S. Wapperom
C. Broekema • D. Heylen
Human Media Interaction, University of Twente, PO Box 217,
7500 AE Enschede, The Netherlands
e-mail: m.bruijnes@utwente.nl

“the police officer is angry and that makes me sad”. The mental state (or mood) of the virtual suspect helps select the most human-like action that the virtual suspect has available, for example “I am sad so I will make a sad face”. We do not present a completely specified mental model for a virtual suspect in this paper, but provide the groundwork for such a mental model. This work is a continuation of work by Bruijnes (2013) and op den Akker et al. (2013) who looked into interpersonal attitudes in the same domain. This paper has two main contributions; first, it describes how we analysed interpersonal behaviour by validating ad hoc interpretations of factors resulting from a factor analysis. Second, we show which theories from (social) psychology and their underlying concepts are relevant to capture social interactions during police interviews and how these concepts are interrelated.

16.1.1 Police Interviews

Police interactions are a special type of social encounter, primarily because of the role of authority that the police officer has and the often uncooperative stance that a suspect takes—there may be a conflict between the interaction parties. The police officer receives training to resolve or reduce the conflict—to make an uncooperative suspect more cooperative. In this section, we discuss the training that police officers receive to become skilled at police interviews.

A police interview is often a situation of conflict. Suspects often do not cooperate with the police officer and the police interview in general, but behave in a confronting manner. Suspects may be withdrawn, defiant or even aggressive towards the police officer. The police officer has the difficult task to convince the suspect to cooperate and tell the truth in an interview: resolve the conflict. At the start of a police interview course, Dutch police students receive theoretical training on the use of the theory of interpersonal stance, or as they refer to it “Leary’s Rose” (Leary 1957) (see Sect. 16.3.1). In addition, “negotiation” strategies are taught with which it is possible to try and change the behavior of the suspect. The *Table of 10* by Giebels (2002) describes the strategies a police officer can use when, for example, they want to convince the suspect that cooperation will be of mutual benefit. After learning about the theory, some students get the opportunity to apply what they learned in a role-playing exercise with professional suspect-actors.

Training the proper behaviour for interviews is important for the effectiveness of the interview. For example, Holmberg and Christianson (2002) showed that when suspects perceive the police officer’s behaviour during the interview as dominant they tend to deny criminal accusations. Alternatively, when suspects perceive the interview as humane and respectful they gain the confidence and mental space required to admit criminal behaviour (Holmberg and Christianson 2002). Richardson et al. (2014) investigated the relation between the verbal mimicry (known as Language Style Matching) in police interviews and the confession to criminal behaviour. They showed that interviews that lead to a confession have a

higher rate of the suspect matching the verbal language style of the interviewer than interviews that did not lead to a confession. Further, they suggest that language style matching and mimicry can be employed strategically (Richardson et al. 2014; Rogan 2011). The notion that the behaviour of the interviewer is of influence on the outcome of an interview is critical for training and maintaining the skill to conduct a police interview.

16.1.2 Data-Driven vs. Theory-Driven

We work towards a computational model of virtual suspects and their conversational behaviour in police interviews. To build such a model we can follow different paths. One way is to start with a literature study and see what conceptual frameworks in behavioural and cognitive psychology and socio-linguistics—that focus on police interview practice—have to offer. The question is whether these theoretical frames provide insights that help us in building an operational model for suspect behaviour. Another approach is to start with the application of annotation schemes for specific dimensions of conversational behaviour to perform content analysis of the conversational data. The question then concerns the statistical correlations between aspects of behaviour in this type of data. For example, what is the relation between interpersonal stances that suspect and police take in an interview and the way they manage turn-taking (op den Akker et al. 2013)? The question is whether the concepts (labels) are clear enough and applicable to the data so that the inter-rater agreement is acceptable.

In the current paper we take a more holistic approach. The question is what concepts people use to describe what is going on in a police interview when they experience/observe it. How do they describe the interview and the behaviour of the interlocutors? Does the data, consisting of terms used to describe what is going on in a police interview, reveal interesting patterns? Such patterns might resemble the patterns that theories in behavioural and cognitive psychology and socio-linguistics describe, linking them to the observed practice of police interviewing.

16.1.3 Paper Outline

We look at human behaviour in a corpus of police interviews and try to establish which psycho-social theories might explain what happens in these interviews. In Sect. 16.2, we give a more detailed overview of our approach. A factor analysis of the occurrence of terms describing interactions in our corpus yielded the basis for a selection of theories and models from (social) psychology, which are discussed in Sect. 16.3. In Sect. 16.4, we address how the concepts underlying the theories matched the factors used for the factor analysis and how the concepts may be interrelated. To illustrate how these concepts can be used to understand the

behaviour of police officers and suspects, we describe several fragments from our corpus in terms of these concepts (Sect. 16.5). We conclude with our thoughts on the creation of a computational model for our virtual agents based on the combination of models and theories (Sect. 16.6).

16.2 Corpus Analysis

In this section, we outline how we analysed the behaviours of police officers and interviewees in a corpus of police interviews. We look at the behaviour of suspects and police officers because we feel modelling the interaction between both parties is necessary to create a believable virtual suspect. In Fig. 16.1, we show the steps we took in our analysis.

We started with a corpus of police interviews (step 1), the *Dutch Police Interview Training Corpus* (DPIT Corpus), see Sect. 16.2.1. From this corpus, six observers independently selected fragments that they thought were “interesting” in some way (2). For example, these were fragments in which a change in mood or atmosphere took place or fragments in which behaviour could be observed that

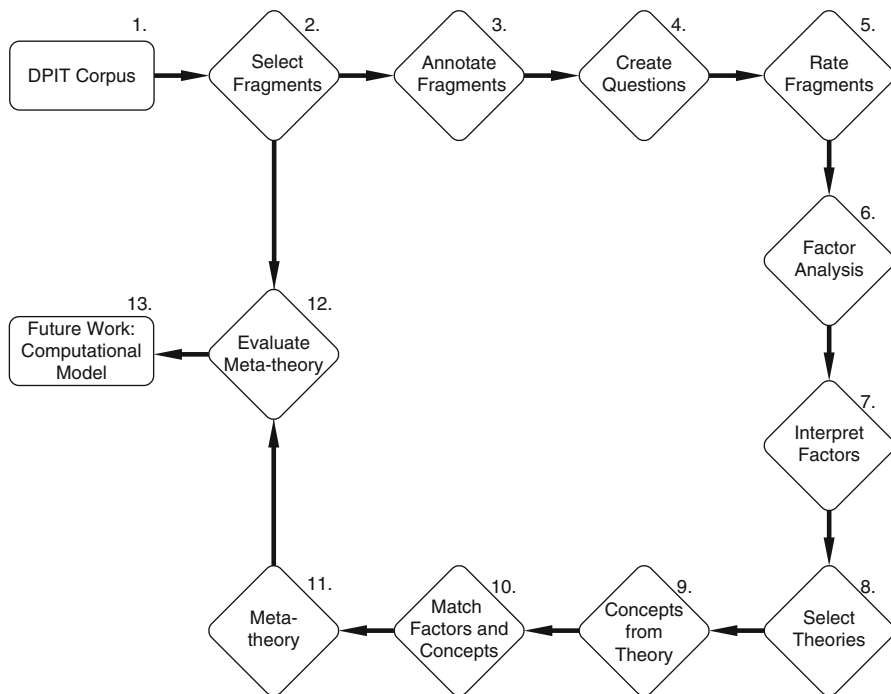


Fig. 16.1 Diagram showing the steps taken towards a computational model of police interviews

indicated how the police officers or suspects felt about the interaction. Next, a subgroup of the observers noted as many different *terms* as possible to describe what was going on in the fragments (3)—these could, for example, be adjectives describing the mood or nouns indicating behavioural traits. Based on these terms, we created questions with variations of the format “*To what extent is [term] the case?*” (4), see Sect. 16.2.2. The original six observers then rated a random subset of the same fragments from the corpus on a five-point Likert scale for every question, once for the police interviewer and once for the suspect. For example, a question from the point of view of the suspect was “To what extent is the suspect dominant?” (5). We performed a factor analysis to find a clustering of correlated questions (6) which we discuss in Sect. 16.2.3. Next, a subgroup of the original six observers reached consensus on the interpretation of the factors (7) and we selected well-known psychological and sociological theories that addressed these interpretations (8), see Sect. 16.3. In Sect. 16.4, we discuss how the concepts that these theories employ were matched to the factors by selecting the concepts that fit each factor (9 and 10). This also revealed the relations between the different concepts that were selected (11), as for some factors, concepts from different theories were applicable. This relation between theories allowed us to integrate the different theories into one “meta-theory” that provides the terms and concepts to describe the interactions in a police interview. We checked whether this “meta-theory” can describe what is happening in a police interview in Sect. 16.5 (12). Our final step (13) is to create a computational model from our “meta-theory” (Bruijnes 2013), but this remains a future endeavour (see Sect. 16.6).

16.2.1 Corpus Description

The *DPIT Corpus* is a corpus that consists of police interviews conducted by trainees of the Dutch National Police, recorded in 2012 and 2013. The police officers in the corpus are novice to moderately proficient police interviewers. The suspects they interview in this corpus are professionally trained actors. Due to privacy requirements, the video and audio data of the corpus is not publicly available.

The corpus consists of 32 interviews from 6 scenarios (cases) with a total length of approximately 13 h. The interviews vary in length from about 9 min to almost an hour. Some scenarios were enacted several times (with different students and sometimes also with different actors), while other scenarios were cut into separate interviews. In the latter, the suspect was interviewed multiple times, for example to give the police officers time to check facts. In these scenarios the same actor is interviewed by different police officers. The scenarios are specified as follows:

Bruintjes Ms Bruintjes is suspected of having bought a stolen smartphone from her cousin. She comes across as being not very bright but knew the phone was stolen.

Huls Mr Huls is suspected of the theft of a small amount of cash from a petrol station. He is a professionally trained actor for the police, has financial problems and has difficulty feeding his family.

Motor Actors from this scenario performed (with criminal exemption) an actual theft of an outboard motor and they played themselves. They try to appear innocent but are instructed to admit when the evidence gets too strong to reasonably deny the crime. The police students are not aware the suspects are guilty and treat them as any other suspect.

Remerink Ms Remerink is suspected to have stolen money from her (ex) husband's bank account. She is a full-time mother and gave up her career for their kids. Her (ex) husband is wealthy and he left Ms Remerink for another woman.

Van Bron Mr Van Bron is suspected of arson with the intent to kill his neighbour. Van Bron has an anti-social or bipolar disorder and has a criminal record. His girlfriend made a statement implicating Mr Van Bron.

Wassink Mrs Wassink is suspected to have physically attacked her neighbour over an argument about the dog of that neighbour. Wassink is a working-class mother whose world is as big as the neighbourhood she lives in and she is suspicious of people not originating from her neighbourhood (like her neighbour).

The actors are allowed to change the scenarios according to their preferences and to fill in the details as they see fit. This means that instances of interviews from the same scenario may be different, yet the police officers are always training with the same *persona*.

16.2.2 Observations of the Corpus

Six observers¹ with previous experience in interaction analysis independently viewed a selection of interviews from the corpus. They selected fragments that they thought were “interesting in some way” (step 2 in Fig. 16.1), for example, fragments in which a change in mood or atmosphere took place or fragments in which specific behaviour could be observed. The observers noted as many different *terms* as possible to describe what was going on in their fragments. The following is an example excerpt from the Remerink scenario (translated from Dutch):

POLICE: My name is Bill [*Surname*].² Can we address each other by our first name?

SUSPECT: Well I don't think so.

¹The first six authors of this chapter.

²Names are fake and/or anonymised for privacy.

P: Don't think so? Then I will use Ms Remerink. You can still call me Bill, if you have any questions you can do it like that... eh... eh... Ms Remerink.

One of the observers wrote down:

The suspect is *invited to call the police officer Bill*, even though she *insists* that he calls her Ms Remerink. He is trying to *be nice*, but he might give away *power*. Now there is *asymmetry* in the way they *address* each other (officer has to say “u” while she can say “je”³).

This description provides six descriptive terms for the fragment: *tutoyer*,⁴ *insist*, *be nice*, *power*, *asymmetry*, and *address*. A subgroup of the observers watched all the interesting fragments and also reported as many terms as possible to describe these fragments using a “think-aloud” strategy (step 3 in Fig. 16.1). In the example above, they added the terms *status* and *cold* to the terms selected by the original observer.

The subgroup of observers added fewer new terms to the entire collection of terms with every successively interview fragment. The first observed fragment yielded over 50 unique terms while only 3 new terms were added to the existing collection after observation of the annotations of the final interview. From this we conclude that we have obtained a sufficiently complete collection of terms necessary to describe the interviews included in our corpus: a *semantic frame* (Allan 2001). Eventually, the collection converged on a total of 251 unique terms.

16.2.3 Rating and Factoring Fragments

Based on the semantic frame of 251 terms, we created 227 questions with variations of the format “*To what extent is [term] the case?*” (step 4 from Fig. 16.1). We excluded terms that were not suited to create meaningful questions, for example, “fact” was a term that is too general to yield a sensible question or every question would have to be specific to the scenario. Example questions that were included are “To what extent is aggressive behaviour the case?”, “To what extent is the speaker indifferent?”, and “To what extent is there an uncomfortable posture?”

The original six observers rated fragments from the corpus on a five-point Likert scale for every question (step 5 from Fig. 16.1). The observers scored 14 fragments (with a total running time of 19 min) of the corpus on the 227 questions. The rated fragments were randomly selected from the fragments that were selected at

³In Dutch there is a difference between the second person pronouns “u” (formal) and “je” (familiar), both are translated to “you” in English.

⁴The (French) term for “to thee and thou”, to be familiar, based on the description “[...] invited to call the police officer Bill.”

Table 16.1 The ten items loading highest on the first suspect factor

Item	Factor loading
“Building pressure”	0.96
“Interruptions”	0.95
“Aggressive behaviour”	0.94
“Angry behaviour”	0.94
“Steering a conversation”	0.93
“Accusing the other”	0.92
“Attacking behaviour”	0.92
“Cutting the other off”	0.91
“Worked-up behaviour”	0.90
“Raised voice”	0.88

This factor was interpreted as *dominant and opposed*

step 2 (see Fig. 16.1). The fragments were scored by asking the rating questions explained above for both the police officer and the suspect, resulting in every question (corresponding to a term) being scored 28 times.

We performed a factor analysis to find a clustering of correlated questions which indicated which categories of questions—and, by extension, which terms—are related (step 6 from Fig. 16.1). Questions that were scored with no variation—that is, they always received the same score—for either the police officer or the suspect were excluded from analysis. This resulted in nine questions being removed (two were excluded from analysis for the subject, seven for the police officer). The excluded terms were found to be very role-specific; for example, crying is something the police never does. The factor analyses (extraction method: Principal Component Analysis, rotation method: Varimax with Kaiser Normalisation) revealed 13 factors for both the suspect and the police, see Tables 16.2 and 16.3.

Based on the related questions, we determined which terms loaded strongly (having a correlation of more than 0.50) per factor. The observers used these terms to interpret the corresponding factors (step 7 from Fig. 16.1). For example, the first factor (explaining 19.4% of the variance) for the suspect was interpreted as *dominant* and *opposed*. In Table 16.1, we show only the first 10 (of 54) items with factor loadings for the first factor of the suspect.

A subgroup of four of the original six observers interpreted all factors. The consensus on keywords describing the strongly loading factors of the suspects and the police officers is reported in Tables 16.2 and 16.3, respectively. In general, the observers’ interpretations were similar. For example, one of the observers interpreted the first factor of the suspect as “negative, confrontational and dominant”, while another observer interpreted it as describing “dominant behaviour and frustrations”. Discussing the interpretations among the observers generally resulted in agreement on the meaning of the factors. Some factors (suspect factors 9, 11, and 13 and Police factor 8) remain unclear as the observers were unable to reach consensus. We attribute this confusion to the few and diverse items that load on these factors.

Table 16.2 Variance explained by each factor for the suspect, with the interpretation of the factors

Suspect factor	Variance (%)	Cumulative variance (%)	Interpretation
1	19.41	19.41	Dominance/opposed (based on frustration), strategy/face (attack)
2	17.40	36.81	Rapport (building), together
3	15.59	52.39	Submissive/opposed, face
4	6.96	59.35	Together
5	6.65	66.00	Strategy (annoy)
6	6.19	72.19	Information exchange (questions)
7	5.74	77.93	Information exchange (lies)
8	5.27	83.20	Strategy (surround a fact)
9	4.70	87.91	–
10	4.01	91.92	Politeness (face)
11	3.02	94.93	–
12	2.90	97.83	Rapport (present)
13	2.17	100.00	– (one item: thank)

Table 16.3 Variance explained by each factor for the police, with the interpretation of the factors

Police factor	Variance (%)	Cumulative variance (%)	Interpretation
1	14.77	14.77	Rapport (missing rapport, negative emotions), arousal, opposed
2	13.57	28.34	Rapport (present), positivity, together
3	11.16	39.50	Strategy (avoid), information exchange (lies)
4	10.46	49.96	Submissive
5	8.83	58.78	Together
6	8.37	67.15	Arousal, dominance (competitive), strategy (attack)
7	8.01	75.16	Dominance/opposed (based on strategy)
8	4.61	79.77	–
9	4.56	84.33	Dominance/together
10	4.52	88.85	Strategy (confront)
11	4.16	93.01	Strategy (confront)
12	3.68	96.69	Dominance
13	3.31	100.00	Strategy (confront)

16.3 Linking Factors to Theories

In this section, we describe how the interpretation of the factors found in the previous section reflects ideas found in theories from (social) psychology (step 7 from Fig. 16.1). Based on the theories discussed in this section, in the following section we present a meta-theory that describes concepts relevant to the interactions in police interviews. The factors describing interpersonal attitudes are taken together in Sect. 16.3.1 on stance; the factors linked to face and politeness are discussed in Sect. 16.3.2; the factors linked to rapport are captured in Sect. 16.3.3. Additionally, two meta-concepts—*information* and *strategy*—were added to accommodate for the concepts that surfaced in the interpretation of the factor analysis but did not fit easily in a theory. Factors relating to information exchange are discussed in Sect. 16.3.4 and factors linked to strategy are discussed in Sect. 16.3.5. In each of the subsections, we describe the relation between these collections of factors and the theories from (social) psychology, including the concepts underlying those theories (step 8 from Fig. 16.1). We provide examples from the corpus and address work done with virtual agents and the mentioned theories. Also, we give some examples of systems using the concepts.

16.3.1 *Interpersonal Stance*

Several interpreted factors for both the suspect (1, 3 and 4 from Table 16.2) and the police (4, 5, 7 and possibly 6 from Table 16.3) are related to the attitude the suspect and the police officer have toward each other. Taken together, these factors sketch the outlines of *Leary's Rose*, a model for interpersonal behaviour (Leary 1957). Leary's Rose represents such behaviour in categories of interpersonal stance on the dimensions of affect (*x*-axis) and power (*y*-axis), see Fig. 16.2a. That is, the underlying concepts of Leary's Rose are part of these axes: the opposing concepts of dominance and submission constitute power, and the opposing concepts of feeling together (positive affect) or feeling opposed (negative affect) constitute affect.

Theories similar to Leary's Rose are known under names such as the Interpersonal Checklist (LaForge and Suczek 1955) and the interpersonal circumplex (Rouckhout and Schacht 2000), but the differences are often superficial. The model is often pictured as an ordering of the stances on a circle, situated on the two axes, which is called a circumplex. The circumplex can be divided into eight areas: these are interpersonal stances. The circumplex shows that stances that are close together are more related than those that are further apart on the circle, with opposites being negatively related (Fig. 16.2a). Leary suggests that human stances are affected by the interaction with the conversational partner. This means that two conversational partners influence each other with their stance during a dialogue. Leary calls these interactions “interpersonal reflexes” and asserts that acts on the dominance dimension are complementary while acts on the affect dimension are symmetric.

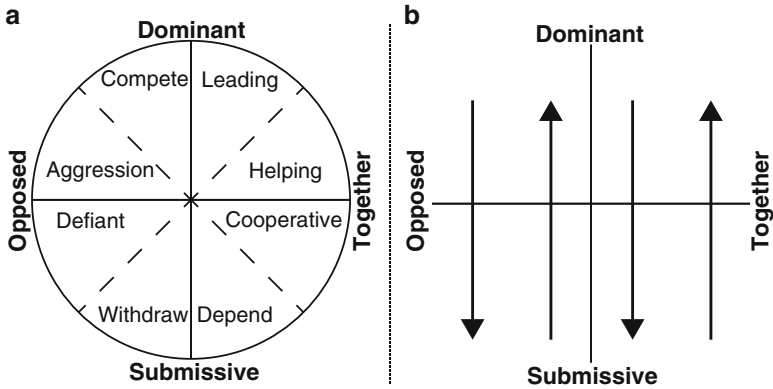


Fig. 16.2 (a) Leary's rose is defined by two axes: a dominance axis (*vertical*), describing the speaker's dominance or submissiveness towards the listener; and an affect axis (*horizontal*), describing the speaker's willingness to cooperate with the listener. The rose can be divided into eight areas that each corresponds to a stance. (b) The *solid arrows* indicate the behaviour-inviting relation between the quadrants according to Leary's (1957) theory. So, dominant-together invites submissive-together behaviour (and vice-versa) and dominant-opposed invites submissive-opposed behaviour (and vice-versa)

This means that a dominant act (for example, power display) elicits submissive acts, whereas an act with positive affect (for example, cooperative behaviour) elicits another positive affect act (see Fig. 16.2b). For example, if someone displays dependent behaviour towards another person (submissive and positive), that other person will feel a tendency to adopt a leading stance (dominant and positive) (Leary 1957).

16.3.1.1 Corpus Examples

In the corpus, we see several examples of different stances. In the Van Bron scenario, the suspect mostly behaves in a detached manner, unwilling to cooperate and expressing this through either competing or defiant behaviour. For example, when Van Bron becomes frustrated about not getting enough time to speak his thoughts and says to one of the officers that they should let him speak, the addressed officer says that he does not need to comply with Van Bron's wishes. As a result, Van Bron becomes somewhat aggressive and acts in a very dominant way, which corresponds to a hostile-dominant stance. On the other hand, the police officer usually displays behaviour with a together stance, for example in the Wassink scenario, in which the police officer does his very best to explain in other words to the interviewee what he was saying just a moment before. In this attempt to help the interviewee, the officer takes a very positive stance towards her by trying to help and cooperate with her.

16.3.1.2 Systems Using This Concept

There have been a few attempts to create virtual agents that act according to the interpersonal circumplex theory. One of these is the serious game *deLearyous*, which focuses on training interpersonal communication skills in a working environment setting, letting users interact with virtual agents through written natural language input (Vaassen and Wauters 2012). However, one of the findings of this project was that determining the stance of dialogue utterances is a very difficult task, even for human annotators. Other work has focused on finding correspondences of non-verbal behaviour with stances (Ravenet et al. 2013). This approach focusses on the generation of upper body movement and of facial animation on a virtual agent, based on human annotation of behaviours.

16.3.2 Face Threats and Politeness

Informed by Goffman's notion of *face* (Goffman 1959)—a person's public self-image—Brown and Levinson (hereafter, B&L) constructed their theory about politeness strategies (Brown and Levinson 1987). Suspect factors 1, 3 and 10 were interpreted as related to face (and politeness), see Table 16.2. The police factors were not interpreted as having a relation with face.⁵

B&L distinguish between negative and positive face, which denote a person's need for freedom (*autonomy*) and a person's need to be approved of and approving of others (*approval*), respectively. Their approach to politeness revolves around the concept of face-threatening acts (FTAs) which are inherent with actions taken by a speaker, as these actions potentially impose on a hearer's face by threatening their needs. B&L view politeness strategies as ways to redress these FTAs in order to minimise their imposition. The four main politeness strategy types follow below, ordered from least to most polite.

Bald on-record Being straight to the point, e.g., “Tell me where you were that night.”

Positive politeness Taking the other's wants into account, e.g., “Would you like to tell me where you were that night?”

Negative politeness Not hindering the other's autonomy, e.g., “If it's not inconvenient to you, could you tell me where you were that night?”

Off record Being indirect or vague about one's own wants, e.g., “I don't seem to have written down where you were that night.”

Conflict situations often arise in the police domain where people may not have the intention to stay polite—on the contrary, they may have the intention to be

⁵Police factor 4 was considered by some interpreters to have a relation with face but this was not unanimous.

impolite. Complementary to B&L's positive and negative politeness strategies, Culpeper et al. (2003) describe impoliteness strategies.

Positive impoliteness Damaging the addressee's positive face wants by excluding him or her, being disinterested, disassociating oneself from the addressee or using taboo words. For example, "Just bloody tell me where you were that night, so I can go home."

Negative impoliteness Damaging the addressee's negative face wants by being condescending, frightening him or her or invading his or her space. For example, "Tell me right now where you were that night, or I'll lock you up till Monday."

16.3.2.1 Corpus Example

On multiple occasions in the van Bron scenario, the suspect demands that the interview takes place according to his wishes. He does this mostly by using sentences that are short and direct, such as "You have to shut your mouth!" when he does not receive ample time to speak and expressing his disinterest by replying to the police with short answers ("It just is."). The first utterance is an example of an attack on the police officer's negative face, as the suspect invades his space and claims room for himself in the conversation. The second may not come across as a direct attack on the police officer's face, but it does impose on his positive face, as it indicates that the suspect does not want to cooperate and does not approve of the police officer. Impoliteness is not limited to solely being used by suspects: police officers use impolite utterances as well. This happens frequently when the police confront a suspect with a lie or an incriminating fact. For example, in the Huls scenario the officer is bald on-record and says "I think you took the money."

Even though police interviews can be uncooperative dialogues, politeness is still abundantly used. For example, in the Huls scenario a police officer explicitly expresses his approval of the suspect's behaviour: "I think it's decent of you that you try to support your family financially." This can be seen as an example of positive politeness, as the police officer takes the suspect's wants (of being approved) into account. In the Motor scenario, an example of negative politeness can be found, as a police officer tries not to impose too much on the suspect's freedom (his autonomy) by saying "I hope you don't mind too much to have this conversation with me."

16.3.2.2 Systems Using This Concept

Based on B&L's definition of politeness, several systems have incorporated virtual agents that can use utterances that vary in politeness. One of the first of these systems was designed by Walker et al. (1997) and involved asking a waiter for drinks with varying degrees of politeness, based on B&L's theory. Work by Gupta et al. (2007) continued this line of research by creating POLLY, a virtual agent that assisted users in learning English as a second language. This agent took into account

how imposing its requests were to the user by redressing these requests according to B&L's theory of politeness.

16.3.3 Rapport

The feeling of rapport can be described as being “in sync” with another person: communication takes place fluently and both interaction partners are roughly on the same level. In our corpus, we see the effects of both the presence and absence of this feeling. Suspect factors 2 and 12 in Table 16.2 and police factors 1 and 2 in Table 16.3 were interpreted as rapport (rapport-like descriptions).

Tickle-Degnen and Rosenthal (1990) conceptualised rapport in order to identify non-verbal correlates. Their description of the nature of rapport focuses on the interaction process as a whole and relies on three components of rapport: *mutual attention*, *positivity* and *coordination*. To develop and maintain rapport, interaction partners need to be mutually attentive so that they can achieve a focused and cohesive interaction. Moreover, their interest in the other party should remain at a high level during the course of an interaction. Figure 16.3 shows a schematic view of relative importance of mutual attention and the other two factors of rapport over time. Tickle-Degnen and Rosenthal mention that being positive towards each other is important during the build-up of rapport, yet becomes less important as time passes during interaction. An example of this is language usage among teens, where insults (a sign of low positivity) are the order of the day (Wang et al. 2012). Lastly, Tickle-Degnen and Rosenthal describe coordination as having a harmonious relationship between partners—this is the key term related to the feeling of being “in sync” and is the factor that becomes more important over time.

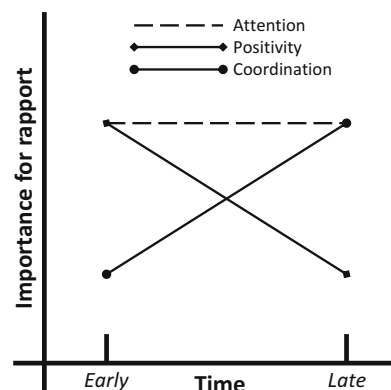


Fig. 16.3 Importance of the three components of rapport over time [from Tickle-Degnen and Rosenthal (1990)]

16.3.3.1 Corpus Examples

In a fragment of the Wassink scenario, the officer and the suspect start to speak more easily and freely to each other after a period of hesitant, slow interaction. The officer starts making gestures and the suspect has her full attention on his comments and responds quickly, without much hesitation. Soon after this, the suspect assumes a more interested body posture and finally both parties start laughing together. What is clear here is that the parties have mutual interest in each other and their coordination increases, resulting in them being “in sync”.

The opposite occurs in a fragment of the Van Bron scenario, in which Van Bron is not listening to what the officers are asking (or does not want to hear what they are saying) and starts making indecent comments about the female officer. In this case, there is little attentiveness of the suspect as well as a lack of intention to be positive towards the officer, resulting in an unpleasant atmosphere in which the police officer is not sure what to say any more.

16.3.3.2 Systems Using This Concept

Huang et al.’s work on the Rapport Agent 2.0—a virtual agent designed to build rapport with users—focuses on backchannelling and turn-taking at the correct moments (Huang et al. 2011). Here, backchannelling and turn-taking are used to inform the user of the attention of the virtual agent. Cassell et al. (2007) address long-term effects of rapport and how these could be modelled by looking at differences in interactions between friends and strangers. In their research, it became apparent that strangers tend to acknowledge each other more—that is, they make sure that the other party understood that they themselves understood what was being said. Friends are much more direct in their interaction, gazing at each other directly and being less explicit about their understanding of each other, which is explained by them having more rapport.

16.3.4 Information Exchange and Framing

Suspect factors 6 (questions) and 7 (lies) and police factor 3 (lies) were interpreted as having to do with information exchange. The discussion between the interpreters revealed more descriptions of information exchange than just “lies” and “questions”, but for these other categories no consensus was reached and they are not included in Tables 16.2 and 16.3. A factor analysis where the questions answered from the points of view of the suspect and the police were taken together revealed more information exchange descriptions during interpretation of these factors, including *give information*, *withhold information*, *lie*, and the notion of *topic or frame*.

Information is exchanged during all conversations between multiple interaction partners. Austin (1975) conceptualised an illocutionary act as the intended

meaning of an utterance, for example a request for information. Based on this notion, Searle (1969) created a classification of five different types of speech acts, namely representatives (informing), directives (requesting), commissives (promising), expressives (expressing a psychological state), and declarations (for official decisions). The main concepts in this theory are those of requesting information (*questioning*) and *giving* (or not giving) *information*.

A special case of giving or withholding information is *lying*: providing information that one knows or believes to be false. Police officers experienced in interviewing have above average lie detecting skills (Mann et al. 2004), mainly because they focus on cues that relate to a suspect's story. In other words, inconsistencies in the information exchange are important during police interviews.

The type of information that is exchanged and how it is interpreted is dependent on context—in other words, the interaction's *frame* determines the type of conversation. The notion of frame was first introduced by Bateson in 1955 as he studied the behaviour of monkeys in different situations (Bateson 1955). Bateson stated that no communication could be interpreted without a meta-message about what was actually going on—that is, what the current frame of the interaction between the monkeys was. During a play frame, all monkeys knew that certain behaviours were accepted (such as biting) which would otherwise be interpreted as a hostile act. Fillmore elaborated on this idea by stating that a frame is “a system of linguistic choices associated with a scene, where a scene is any kind of coherent segment of human actions” (Fillmore 1981). According to Tannen, conversational frames are repositories for social cultural norms of how to conduct different types of conversation, such as storytelling, teasing, and small talk (Tannen 1993). A frame tells us something about what we can and cannot say in that particular frame. The frame that is currently active allows us to decide which assumptions we can make, customs or “social scripts” we have (what we can do), and constraints we have (what we should not do).

16.3.4.1 Corpus Examples

Dutch police interviews start with a social frame during which the police officer tries to get to know the suspect and gathers information about the personal life and emotional situation of the suspect. After getting to know each other they continue with a task frame where they discuss the crime that the suspect has been accused of.

Conversational partners do not always agree on the frame that they are using. During the Wassink scenario the suspect does not agree with the social frame the police officer suggests and she asks: “Why do I have to tell you something about myself?”

In the Huls scenario the suspect eventually admits to the crime of stealing money from the gas station. During this confession the police officer uses an empathy frame (Bickmore 2008) in which he comforts the suspect by telling him that he understands his situation because he too has children. He agrees with the suspect that it is hard to provide for two children without a stable income.

16.3.4.2 Systems Using This Concept

Multiple virtual agent systems have been created that are at least partly based on speech or dialogue act theories. For example, the Mission Rehearsal Exercise and Stability and Support Operations systems and their derivatives feature agent decision making using speech acts (Swartout 2010). The same is true of Kopp's virtual museum guide which distinguishes between the performed behaviours and the communicative function of these behaviours (Kopp et al. 2005). This helps the virtual guide to select responses that vary in their performance, yet have the same communicative function.

Bickmore (2008) developed a health counselling agent that bases its reactions on both interpersonal stances and framing. Bickmore uses four different conversational frames to help the agent decide on how to react: the *task* frame, which is used for information exchange; the *social* frame, which is used for social chat and small talk interactions; the *empathy* frame, which is used for comforting interactions; and the *encourage* frame, which is used for coaching, motivating and cheering up interactions. With this information, combined with interpersonal stance, the agent can decide what behaviour to show in different situations.

16.3.5 Strategy Selection

Suspect factors 1, 5 and 8 and police factors 3, 6, 10, 11 and 13 were interpreted as having to do with strategies in interaction, see Tables 16.2 and 16.3. Specifically, the interpreters used the concepts *confront*, *surround*, *evade*, and *annoy*.

During communication, individuals make use of strategies to achieve their desired goals. These strategies play an important role, especially during non-cooperative communication such as in the police domain, as described in Sect. 16.1.1. Traum et al. (2008) describe a set of negotiation strategies—including finding the issue, attacking to aggressively attain a goal, and advocating or proposing solutions—and assert that the negotiating party must balance three goals to be successful in a difficult negotiation. The negotiator has to find an acceptable solution for the problem, gain and maintain the trust of the other participant(s) and manage the interaction by setting the agenda and controlling the topic.

Campos et al. (2012) explain strategies in terms of conflict. In natural conflict situations the agents will respond using emotional reactions. According to Campos et al., conflict varies around five dimensions: participants, causes, initiating action, responses and outcomes. Thomas (1992) argues that participants can take several approaches to resolve conflicts: *accommodation*, *avoidance*, *competition*, *collaboration* and *compromise*. As mentioned in the introduction, the *Table of 10* by Giebels (2002) describes the strategies a police officer can use when, for example, they want to convince the suspect that cooperation will be of mutual benefit.

16.3.5.1 Corpus Examples

The Remerink scenario in our corpus starts with a frustrated suspect who is apparently angry about something. The police officer uses a negotiation strategy to find out what is bothering her in order to resolve this issue. He asks the woman what is bothering her and eventually she says she is angry about the method with which she was picked up from her house. She is ashamed and angry about the way they came to her house and brought her in with all the neighbours watching.

Later in the Remerink interview, the suspect is accused of taking money from her ex-husband and she becomes emotional and silent every time when the topic of her husband comes around. Due to the fact that the topic is undesirable for her to talk about, she tries to avoid going into it any further.

In the Huls scenario, the police officer is surrounding a specific fact during the conversation, so the topic cannot be avoided. He continues to aggressively ask similar questions to the subject to put pressure on him to tell the truth.

16.3.5.2 Systems Using This Concept

The Mission Rehearsal Exercise and Stability and Support Operations systems and their derivatives feature virtual agents in war scenarios with which users have to negotiate (Swartout 2010). These scenarios deal with dilemmas the user has to solve. For example, a user has to convince a local Afghan doctor to move his clinic to another location, as the user has to conduct military operations in the area of the clinic. One of the ways to convince the doctor is using rational arguments such as offering incentives. Furthermore, this system also takes emotional consequences into account when deciding whether to cooperate with or oppose the user.

16.4 Relations Between Factors, Theories and Concepts

In the previous section, we discussed what theories from (social) psychology match the interpretations of the factors found in the factor analyses (see Sect. 16.2.3) and we explained the concepts from these theories (step 9 from Fig. 16.1). In this section, we discuss how these concepts are related to the factors (step 10 from Fig. 16.1). This gives insight into both the relations between the factors and the concepts, and the relations between the concepts themselves (step 11 from Fig. 16.1). Based on our findings on these relations, we describe how the theories from which these concepts originate are connected.

16.4.1 Concepts in Theories

Psychological and sociological theories use concepts and describe the relations between these concepts. Theories provide us with a way to describe an interaction (in our case, a police interview) and they can be used by a virtual tutoring agent (in our case, a virtual suspect) to predict the effects of its behaviour in an interaction with a human user. For example, the central concepts in the interpersonal stance theory are *dominant*, *submissive*, *together*, and *opposed* and the theory describes how the combination of these concepts creates the notion of “stance” and predicts how people are influenced by the stance of others. A virtual tutoring agent can use this knowledge to create an interesting and useful learning experience. For example, a user might learn by experiencing that if he displays opposed behaviour interaction, the conflict escalates. The virtual suspect can display opposed behaviour in an attempt to get the user to also display opposed behaviour and then let the conversation escalate (Bruijnes et al. 2013). Each of the theories we selected in the previous section has such concepts, see Table 16.4. The concepts from face are positive and negative *autonomy*, and positive and negative *approval*. The concepts for rapport are *coordination*, *attention*, and *positivity*. We added two meta-concepts—*information* and *strategy* to accommodate for the concepts that surfaced in the factor analysis interpretation but did not fit easily in a theory. The information concepts we found are *questioning*, *give information*, *withhold information*, *lie* and *frame* or *topic*. The strategy concepts are *confront*, *surround*, *evade* and *annoy*.

16.4.2 Factors: Theories and Concepts

The interpretation of the factors, see Sect. 16.2.3 and Tables 16.2 and 16.3, and the matching of these factors to theories leads to links between theories and factors. To validate these links, four observers indicated with which concept(s) from the theories (Table 16.4) a factor could be explained. This method provided us with a possibility to validate the intuitive (subjective) interpretation of the factors that is common practice in the field of social science. In other words, we used the initial interpretations to select theories that “cover” the factors and we used the concepts

Table 16.4 Concepts within the theories *stance*, *face*, and *rapport* and the meta-concepts *information* and *strategy*

Stance	Face	Rapport	Information	Strategy
Dominant	Autonomy+	Coordination	Questioning	Confront
Submissive	Approval+	Attention	Give info	Surround
Together	Autonomy–	Positivity	Lie	Evade
Opposed	Approval–		Withhold info	Annoy
			Frame/topic	

from these theories to validate the labels for the factors. This matching of factors to concepts is a data-driven interpretation of the factors (see Sect. 16.1.2) and might bring us closer to a “correct” interpretation of a factor.

In Fig. 16.4, we show the cumulative score the observers gave the concepts for each factor. The colour coding in this figure indicates how much the observers agreed that the concept could explain the factor: the dark-coloured cells indicate unanimous agreement, the light-coloured cells indicate that three out of the four observers agreed. The initial factor analysis interpretation of the factors is indicated with an asterisk.

The fit of the factors and the concepts determines the validity of the interpreted theory for this factor. The observers unanimously matched most factors to the concepts corresponding with the initial factor analysis interpretation of the factor; see Fig. 16.4 in which the asterisks indicate the initial interpretation and the dark cells indicate unanimous matching. The factors where the observers disagreed (not unanimous) with the initial factor interpretation are police factors 1, 9, 11, and 13 and suspect factors 1, 3, and 10 (see Fig. 16.4).

We see several explanations for this disagreement; first, the initial subjective interpretation of the factors might have been wrong. The factors with a higher number had fewer items loading on them (and less explanatory power), which might have made it more difficult to interpret them. Four of these higher factors (suspect factors 9, 11, and 13 and police factor 8) had few and diverse items loading on them, which resulted in disagreements during the initial factor interpretation. It is likely the disagreement persisted in the current analysis for suspect factors 9, 11, 13 and police factor 10. Second, factors could initially have been interpreted as having an “absence of something”. This was the case for police factor 1 which was initially interpreted as “missing rapport”. In the subsequent “mapping concepts to factors” task, the observers did not unanimously match the concepts of rapport to police factor 1. This might be because the instructions were unclear what to do when a concept was explicitly absent: some observers said that this factor contains information about the concept rapport (i.e. rapport is missing) while others said there is no rapport so the concept of rapport is not present. For suspect factor 3 and 10 we can give no alternative explanation and conclude that our initial factor interpretation was incorrect (see Table 16.5).

16.4.3 *Relations Between Theories*

The observers unanimously matched factors to concepts from theories that were not initially included in the factor analysis. In other words, more concepts than initially come to mind might play a role in explaining a factor. For example, suspect factor 4 was interpreted during factor analysis as a *together* stance, but the factor was not only matched to the concept *together* (from theory of stance), but also to *positive approval* (from the theory of face), and *attention* and *positivity* (from rapport) (see Fig. 16.4).

Factor	Stance		Face		Rapport		Information			Strategy		
	Dom Sub	Tog Opp	Aut+ App+ Aut-	App-	Coor Att Pos	Quest G.	Inf Lie W.	Inf Frame	Confr	Sur	Evade	Annoy
P 1	2	4*	2	4	2*	1*	2*	1	1	1	2	4
P 2	4*	4*	4	4	2*	4*	4*	1	2	1	3*	3
P 3	1	3	1	3	1	2	1	4*	4	3	1	3*
P 4	4*	1	1	2	1	2	1	1	2	3	1	3
P 5	2	1	2	4	3	3	4	1	1	1	1	1
P 6	4*	1	4	2	1	1	1	2	1	4*	1	1
P 7	3*	1	1	2	1	1	1	1	1	2	1	1
P 8	4*	1*	2	3	1	1	1	1	1	1	1	1
P 9	4*	1*	2	2	1	1	2	1	1	1	1	1
P 10	1	1	2	2	1	1	1	1	1	3*	3	1
P 11	1	2	2	1	2	1	1	1	2*	2	1	1
P 12	4*	1	1	3	2	1	1	3	2	2	1	1
P 13	1	3	3	2	1	3	1	2	1*	1	1	1
S 1	4*	4*	* 3*	4*	1	2	1	1	1	3	2	4
S 2	4*	4	3	4	4*	4*	4*	1	1	1	2	4
S 3	4*	1*	1*	2*	3*	1	1	1	2	2	3	3
S 4	1	4*	2	4	3	4	4	2	2	2	3	3
S 5	2	1	1	2	1	1	1	1	2	3	1	1
S 6	2	1	2	4	2	3	3	4*	1	1	2	4*
S 7	1	1	3	1	1	1	1	1	3	1	4	2
S 8	1	1	1	3	2	1	2	1	2	1	4*	2
S 9	1	2	3	1	2	2	1	1	1	1	1	1
S 10	1	2	3*	1*	2	2	2	1	1	1	1	1
S 11	2	1	2	3	1	2	1	1	1	1	1	1
S 12	1	4	2	3	4*	4*	4*	1	1	1	1	1
S 13	1	1	2	3	1	1	1	1	1	1	1	1

Fig. 16.4 The matching of concepts from the models with the factors derived from the clustering of terms. The numbers indicate how many (out of 4) observers thought the concept from the theory (column) fit the factor (row). Asterisks denote the initial interpretation of the corresponding factor

Table 16.5 The factors from the factor analysis with the concepts that were unanimously matched to the factors

	Factor	Interpretation based on concepts
Police	1	Opposed, negative approval, annoy
	2	Together, positive approval, attention, positivity
	3	Lie, withhold info
	4	Submissive
	5	Together, positive approval, positivity
	6	Dominant, negative autonomy, confront
	7	Opposed, negative approval
	8	–
	9	Dominant
	10	–
	11	–
	12	Dominant
	13	–
Suspect	1	Dominant, opposed, negative approval, annoy
	2	Together, positive approval, coordination, attention, positivity
	3	Submissive
	4	Together, positive approval, attention, positivity
	5	Opposed, negative approval, annoy
	6	Questioning
	7	Evade
	8	Surround
	9	–
	10	–
	11	–
	12	Together, coordination, attention, positivity
	13	–

A dash means no concepts were unanimously matched to the factor

Our methodology makes clear how the theories are related to each other. For each factor, concepts from different theories can be applicable. This *co-occurrence of concepts* suggests that the corresponding theories are related. In Table 16.6 we show in how many factors concepts co-occur. For example, the interpersonal stance *together* co-occurs in more than one factor to: *positive approval*, a concept underlying face (4 co-occurrences), and *coordination* (2), *attention* (4), and *positivity* (5), which underlie rapport (see Table 16.6). This is indicative of a strong link between these concepts. The concept of *dominance* co-occurs with the concepts: *opposed stance*, *negative autonomy*, *negative approval*, *confrontation* and *annoy* (see Table 16.6). To investigate this relation further, we look at Fig. 16.4, which shows that suspect factor 1 was matched with the concepts dominance, opposed, negative approval, and annoy. For the police factors dominance co-occurs with negative autonomy and confront. This is likely due to the different roles the

Table 16.6 The observed relations between the theories based on the co-occurrence of their concepts in the factors

	Stance			Face			Rapport			Info			Strategy	
	Dom.	Tog.	Opp.	App+	App-	App+	App-	Att.	Pos.	Lie	W. info	Confront	Annoy	
Stance	x		1		1							1		1
Dominant														
Together		x		4		2	4	5						
Opposed	1		x		4									3
Face														
App+		4		x		1	3	4						
App-	1		4		x									3
Rapport														
Coordination		2		1		x	2	2						
Attention		4		3		2	x	4						
Positivity		5		4		2	4	x						
Info										x	1			
Lie														
Withhold info										1	x			
Strategy													x	
Confront	1													
Annoy	1		3		3									x

The numbers indicate how often these concepts co-occurred. Only the concepts rated unanimously present in the factors are shown

interactants have. The police officer assumes a dominant stance when he confronts the suspect with an incriminating fact: the act of confronting is dominant. This might be strengthened by the power the police officer has, as he dictates the course of the interview: a concern for the autonomy of the suspect. A dominant suspect might use the strategy annoy to intentionally thwart the progress of an interview and this could negatively impact the approval of the officer. Linssen et al. (2013) proposed that interpersonal stance and politeness (face) are related. They suggest that the dimensions of power and affect used in the model of Leary's Rose can be mapped to the dimensions of face: autonomy and approval. For example, when a person is very dominant, she does not take the other's autonomy into account. A similar relation holds for the dimension of affect, as a person who is opposed to someone else expresses disapproval of that person.

We further investigate the relations between the different concepts in the next section and illustrate them using examples from the corpus. The related theories will be integrated to form the basis for a computational model. As each theory describes relations between the cause and effect of behaviour in an interaction, a virtual tutoring agent (virtual suspect) could use a computational model of these theories to predict the effects of its behaviour in an interaction with a human user (see Sect. 16.6 on future work).

16.5 Illustration of Relations

In the previous section, we showed that certain concepts underlying the theories appear to be related based on the data from our corpus (see Fig. 16.4 and Table 16.6). Here, we illustrate several of these links with example fragments from the corpus that were not used for annotation and the subsequent factor analysis. We illustrate the *co-occurrence of concepts* (see previous section) that shows the relation between the concepts of different theories. Also, we illustrate how our findings might be extended to explain the dynamic aspects in a police interview.

16.5.1 Co-occurrence of Concepts

We found the strongest links between the *together* stance and *positive approval* concepts and between the *opposed* stance and *negative approval* concepts (see Table 16.6). An example from the Brintjes scenario (see transcript below) shows a together stance occurring together with positive approval. In this fragment, the police officers are asking questions about the suspect's leisure time, to which the suspect responds that she spends most of her time at the mall with her girlfriends. The police officers respond to this by indicating that they understand what she means ("Just chilling.") and they all start laughing about this. In this moment, the police officers are very much trying to sympathise with the suspect, thereby

adopting a together stance. They are also expressing approval by saying that they understand the suspect's wish to stay at the mall. In the preceding section, we also showed that there is a strong link between the together stance, concepts underlying rapport (particularly attention and positivity) and positive approval. In the corpus fragment about the suspect staying at the mall, it is clear that both interaction parties are paying a high degree of attention to each other. One of the police officers is asking questions about the suspect's activities which yield immediate responses from the suspect. There is, however, no uncomfortable atmosphere during this part of the conversation, as both the officers and the suspect start laughing about this topic. Thus, the concepts of a together stance, positive approval, and both attention and positivity are displayed in this part of the conversation.

POLICE OFFICER: Those girlfriends, eh, 'cause you said you go shopping with your girlfriends...

SUSPECT: Mm mm. [Confirmatory.]

P: Do you have good friends? Tight friends?

S: [Nods enthusiastically.] Yes.

P: Yes?

S: Yes.

P: So what do you go and do with your friends?

S: Yeah, well, basically, we are often at the mall.

P: At the mall?

S: Yeah, one of those indoor malls.

P: And what do you do there?

S: [Softly:] Kind of hanging around. [Laughs.]

P: [Laughs.] Just chilling.

S: [Laughs.] Yeah!

An example of the strong link between an opposed stance and negative approval can be found in the Wassink scenario (see transcript below). In this excerpt, Mrs Wassink, the suspect, is asked whether she wants to cooperate with the police officer by answering some of his questions, because he wants to form a picture of her situation. Mrs Wassink does not comply and indicates that she does not see the point of doing so.

POLICE OFFICER: I don't know you and you don't know me either.

SUSPECT: No.

P: But maybe it would be convenient if we would first discuss some things about you—about who you actually are. Do you think that's OK?

S: Well... Why?

P: You don't think that's useful?

S: [Shrugs, shakes her head.] No, I don't know why I should tell you who I am.

P: Yeah. [Short pause.] Well, I would like to know.

S: But what for?

- P: Because I would like to know who I'm talking to before I can talk with you—what you just indicated, that you might have physically abused someone.
- S: [Stares blankly.]
- P: Do you understand what I'm saying?
- S: Yes, I do, but that you would like it, well... [Shrugs.] Are you going to tell me something about yourself as well or...? [Shakes head.]
- P: Well, I don't know, would you be interested in what I would have to say?
- S: No, but I also don't understand why you would be interested in me.

In the above fragment, the suspect gives snappy replies to the police officer's questions. This is a first indication that she is not trying to cooperate: she seems to have a very opposed stance to the officer, his approach and his proposals. With her behaviour, the suspect also expresses disapproval (negative approval), repeatedly shaking her head and shrugging, indicating that she does not agree with the police officer or just does not care. Moreover, Mrs Wassink goes a bit beyond simply disagreeing, as she seems to intentionally annoy the police officer. She does this by questioning the police officer's approach (repeatedly asking of what use it is), by expressing that she does not understand what is going on, and by asking a counter-question to the police officer (whether he will say something about himself as well). Here we see that an opposed stance, negative approval, and an annoy strategy occur together.

We found several other relations between concepts from the theories we used that occurred less frequently in our collection of factors. On some occasions, the police officer, but mostly the suspect, used a confront strategy which was accompanied by a dominant stance and negative approval. In these cases, the suspect was trying to lead the conversation by confronting the police officer(s) with his or her own opinions (which were negative in nature most of the time). Another striking co-occurrence of concepts is that of the concepts underlying rapport. As Tickle-Degnen and Rosenthal (1990) assert, coordination, attention and positivity generally occur together to form the feeling of rapport and this is confirmed by our observations.

16.5.2 Concept Dynamics

Our approach to analysing the corpus of police interviews hinges on the annotation of short fragments. However, our annotation did not capture the dynamic aspects of the interviews, for example how and why people change stances or how their feelings of rapport increase or decrease. Here, we illustrate how this may work by describing a change in a situation in the Van Bron scenario (see transcript below) in terms of the concepts from Table 16.6. In this case, the suspect is asking the police officers their name and surname in a dominant way. One of the police officers immediately agrees to give his surname, but the other only gives his first name. To this the suspect replies by making a small gesture with his hand, implying that

he also wants to know the officer's surname. The officer then responds in a laconic way by saying "Oh, you want my *surname*?" which leads to the suspect imitating the officer's response and adding "Wise guy." What we witness here is an exchange in which the police officer has an opposed stance towards the (dominant) suspects and acts disapprovingly of the suspect's behaviour by not granting a full answer. This, in turn, elicits a similarly disapproving response by the suspect. Thus, over time, the suspect changes from having a dominant stance towards one that is more opposed to the officer because of the latter's behaviour.

- SUSPECT: [Points at police officer 1.] What was your name again?
 POLICE OFFICER 1: **** [Gives his first name.]
 S: And...? [Makes a gesture with his hand for the police officer to complete his name.]
 P₁: **** [Gives his surname.]
 S: ****? [Repeats the name.]
 P₁: Yeah.
 S: [Chuckles.] And you? [Points at police officer 2.]
 POLICE OFFICER 2: **** [Gives his first name.]
 S: [Makes the 'completion' gesture again.]
 P₂: What is that? [*Mimics the gesture.*]
 S: [To P₁.:] He doesn't understand? **** [P₁'s surname.] And you? [Points at P₂.]
 P₂: Ah, you want to know my last name?
 S: [Stares at P₂.]
 P₂: Yeah, if you could just be clear in your questions...
 S: He understands [Points at P₁.] are you a bit stupid or something like that?
 P₂: Yeah, I'm a bit more stupid than him, OK...
 S: That's clear.
 P₂: [Softly:] All right.
 S: No, it's not all right. What is your surname?
 P₂: **** [Gives his surname.]
 S: Ah... **** [Repeats P₂'s surname.] [Softly:] Wise guy.

16.6 Conclusion and Discussion

In this paper, we presented our methodology for analysing the behaviour of police officers and suspects in a corpus of enacted police interviews. Taking a holistic approach, we described fragments of this corpus in short terms that captured the behaviour of the participants in, and the atmosphere of, the interviews. We used a factor analysis to cluster related terms based on ratings of observers who annotated to what extent these terms were applicable to each fragment. Based on the factors we found, we selected theories from (social) psychology that we intuitively thought could explain these factors. We included theories about interpersonal stance, face

and politeness, and rapport and defined two meta-concepts, namely “information” and “strategy”, to account for the interpretations that remained. To determine whether these theories matched the factors, we investigated whether these factors could be explained by the concepts underlying the theories. We found that our initial factor interpretation and the match between factors and concepts overlapped broadly. We also found that many factors were matched to more concepts than initially were associated with the factors. We used this finding to create a collection of interrelated concepts that gives insight into how the different theories relate to each other. With this collection, we are able to (at least partly) describe the behaviour of both police officers and suspects in an interview setting.

Our combination of holistic and theory-driven methodology does, however, have its limitations. As is the case in most observational studies, our annotations of the police interview corpus were based on our interpretations of the behaviour of the interacting parties and thus subjective. For future work, our methodology may be repeated to include more observers (and more independent observers) which may lead to a broader semantic frame, possibly alleviating problems inherent with interpretation of behaviour.

A second limitation of our approach is that it currently focuses on describing *short* fragments from the corpus. In the previous section however, we illustrated how our findings may be extended to explain changes in the behaviour of interacting parties over longer periods of time. We based these examples on how temporal aspects are explained by the theories from which we drew our concepts. We wish to continue this line of research by investigating how the interplay of these concepts influences the dynamics in police interviews. This may, for instance, be done by locating moments in our corpus in which a person’s behaviour changes. For example, there may be moments when a person changes his or her stance or becomes less polite. Someone may also consciously change his behaviour to evoke desired behaviour of the other party. This may, for example, be the case when a police officer adopts a “together” stance to build rapport with a suspect. Thus, the communicative contexts before and after this type of changes in behaviour should be compared, to discover what may have caused the change in behaviour. This causality is of vital importance for the creation of a virtual suspect agent, as such an agent needs to be capable of taking logical (and explainable) actions. Thus, this calls for an extended empirical study of the corpus. Such a study may also validate the links we found between concepts, as our current work only investigated a number of fragments from our corpus.

Lastly, we also wish to investigate how the methodology we present in this paper translates to other domains. Whether our approach can be used to analyse communicative behaviour in other domains depends on the availability of a corpus and theories on interaction that explain (parts of) the behaviour. A related domain in which we are also involved is that of street interventions by police officers with loitering juveniles. This domain features a different setting and the environment imposes other restrictions on the interaction, such as an easier “way out” for the juveniles because they are not kept in a room like the interviewees. Still, this domain does not differ strongly from the domain discussed in this paper, as they are both

related to police work and display the unique features this work has, such as the status of the police officer. It remains to be investigated whether our approach would allow for feasible analysis of behaviour a completely different domain. However, given a sufficiently rich corpus of such interactions, we expect that our methodology can be used to analyse corpora from other domains as well.

In future work, we will construct a mental model for virtual agents in a police interview setting. As indicated above, we will focus on the dynamics of such interviews, establishing a computational model that enables a virtual agent to perform causal reasoning. This system will go beyond being an “autonomous sensitive artificial listener” as in Schroder et al. (2012). The system will be able to use the “mood” of its mental model to select the most appropriate action it has available. The current work will inform the creation of this model, which will in turn be used for virtual agents in a tutoring application. Thus, having related data to theories, our next step will be to relate theories to practice.

Acknowledgements This publication was supported by the Dutch national program COMMIT.

References

- Allan K (2001) *Natural language semantics*. Blackwell Publishers, Oxford
- Austin JL (1975) *How to do things with words*, vol 1955. Oxford University Press, Oxford
- Bateson G (1955) A theory of play and fantasy. *Psychiatr Res Rep* 2(39):39–51
- Bickmore T (2008) Framing and interpersonal stance in relational agents. In: *Functional markup language workshop at AAMAS 2008*
- Brown P, Levinson SC (1987) *Politeness: some universals in language usage*. Cambridge University Press, Cambridge
- Bruijnes M (2013) Affective conversational models: interpersonal stance in a police interview context. In: *Humaine Association conference on affective computing and intelligent interaction*, pp 624–629
- Bruijnes M, Kolkmeier J, op den Akker R, Linssen JM, Theune M, Heylen DKJ (2013) Keeping up stories: design considerations for a police interview training game. In: *Social believability in games workshop at ACE 2013*
- Campos H, Campos J, Martinho C, Paiva A (2012) Virtual agents in conflict. In: *Intelligent virtual agents*, pp 105–111
- Cassell J, Gill AJ, Tepper PA (2007) Coordination in conversation and rapport. In: *Proceedings of the workshop on embodied language processing*. Association for Computational Linguistics, Prague, pp 41–50
- Core MG, Lane HC, Van Lent M, Gomboc D, Solomon S, Rosenberg M (2006) Building explainable artificial intelligence systems. In: *Proceedings of the national conference on artificial intelligence*, vol 21
- Culpeper J, Bousfield D, Wichmann A (2003) Impoliteness revisited: with special reference to dynamic and prosodic aspects. *J Pragmat* 35(10):1545–1579
- Fillmore C (1981) Pragmatics and the description of discourse. In: Cole P (ed) *Radical pragmatics*. Academic Press, New York, pp 143–166
- Giebels E (2002) Beïnvloeding in gijzelingsonderhandelingen: De tafel van tien. *Nederlands tijdschrift voor de psychologie* 57:145–154
- Goffman E (1959) *The presentation of self in everyday life*. Garden City, New York

- Gupta S, Walker M, Romano D (2007) How rude are you? Evaluating politeness and affect in interaction. In: Proceedings of the international conference on affective computing and intelligent interaction, pp 203–217
- Holmberg U, Christianson SÅ (2002) Murderers' and sexual offenders' experiences of police interviews and their inclination to admit or deny crimes. *Behav Sci Law* 20(1–2):31–45
- Huang L, Morency LP, Gratch J (2011) Virtual rapport 2.0. In: Intelligent virtual agents. Springer, Berlin, pp 68–79
- Koops M, Hoevenaar M (2013) Conceptual change during a serious game: using a Lemniscate Model to compare strategies in a physics game. *Simul Gaming* 44(4):544–561
- Kopp S, Gesellensetter L, Krämer NC, Wachsmuth I (2005) A conversational agent as museum guide-design and evaluation of a real-world application. In: Intelligent virtual agents, Springer, pp 329–343
- LaForge R, Suczek RF (1955) The interpersonal dimension of personality. *J Personal* 24(1):94–112
- Leary T (1957) Interpersonal diagnosis of personality: functional theory and methodology for personality evaluation. Ronald Press, New York
- Linssen JM, Theune M, Heylen DKJ (2013) Taking things at face value: how stance informs politeness of virtual agents. In: Workshop on computers as social actors at IVA 2013
- Mann S, Vrij A, Bull R (2004) Detecting true lies: police officers' ability to detect suspects' lies. *J Appl Psychol* 89(1):137
- op den Akker R, Bruijnes M, Peters R, Krikke T (2013) Interpersonal stance in police interviews: content analysis. *Comput Linguist Netherlands J* 3:193–216
- Ravenet B, Ochs M, Pelachaud C (2013) From a user-created corpus of virtual agent's non-verbal behavior to a computational model of interpersonal attitudes. In: Intelligent virtual agents. Springer, Berlin, pp 263–274
- Richardson BH, Taylor PJ, Snook B, Conchie SM, Bennell C (2014) Language style matching and police interrogation outcomes. *Law Human Behavior* 38(4):357–366
- Rogan RG (2011) Linguistic style matching in crisis negotiations: a comparative analysis of suicidal and surrender outcomes. *J Police Crisis Negot* 11(1):20–39
- Roukhout D, Schacht R (2000) Ontwikkeling van een nederlandstalig interpersoonlijk circumplex. *Diagnostiekwijzer* 4:96–118
- Schroder M, Bevacqua E, Cowie R, Eyben F, Gunes H, Heylen D, Ter Maat M, McKeown G, Pammi S, Pantic M et al (2012) Building autonomous sensitive artificial listeners. *IEEE Trans Affect Comput* 3(2):165–183
- Searle JR (1969) Speech acts: an essay in the philosophy of language, vol 626. Cambridge University Press, Cambridge
- Steunebrink BR, Dastani M, Meyer JJC (2012) A formal model of emotion triggers: an approach for BDI agents. *Synthese* 185(1):83–129
- Swartout W (2010) Lessons learned from virtual humans. *AI Mag* 31(1):9–20
- Tannen D (1993) What's in a frame? Surface evidence for underlying expectations. In: Tannen D (ed) Framing in discourse. Oxford University Press, New York, pp 14–56
- Thomas KW (1992) Conflict and conflict management: reflections and update. *J Organ Behav* 13(3):265–274
- Tickle-Degnen L, Rosenthal R (1990) The nature of rapport and its nonverbal correlates. *Psychol Inquiry* 1(4):285–293
- Traum D, Marsella SC, Gratch J, Lee J, Hartholt A (2008) Multi-party, multi-issue, multi-strategy negotiation for multi-modal virtual agents. In: Intelligent Virtual Agents, Springer Berlin Heidelberg, pp 117–130
- Vaassen F, Wauters J (2012) deLearyous: training interpersonal communication skills using unconstrained text input. In: Proceedings of the 6th European conference on games based learning, pp 505–513

- Vinciarelli A, Pantic M, Heylen DKJ, Pelachaud C, Poggi I, D'Ericco F, Schroeder M (2012) Bridging the gap between social animal and unsocial machine: a survey of social signal processing. *IEEE Trans Affect Comput* 3(1):69–87
- Walker MA, Cahn JE, Whittaker SJ (1997) Improvising linguistic style: social and affective bases for agent personality. In: *Proceedings of the international conference on autonomous agents*, pp 96–105
- Wang WY, Finkelstein S, Ogan A, Black AW, Cassell J (2012) “Love ya, jerkface”: using sparse log-linear models to build positive (and impolite) relationships with teens. In: *Proceedings of the 13th annual meeting of the special interest group on discourse and dialogue*, pp 20–29