Detection of leadership in informal (small) groups based on CCTV information

Jan-Willem Bullée
j.h.bullee@student.utwente.nl
Twente University
Faculty of Electrical Engineering, Mathematics and Computer Science
Chair of Signals and Systems

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Become the kind of leader that people would follow voluntarily; even if you had no title or position.

Brian Tracy

In general, events such as concerts and public celebrations elapse quietly and easy, without problems. The occurrence of an incident, however, may have terrible consequences (Hijum, 2011). Numerous examples from the last 30 years can be given of asphyxia, crushing and stampeding during events (CNN Sports, 2001; Helbing & Johansson, 2009). The number of reported incidents increases each decade. This trend is, not entirely remarkable, accompanied by an increased attention for public safety problems (Fan Weicheng, Liu Yi, 2008 (as cited in Wei, Guo, Dong, & Li, 2012)). The report of Hughes states that in the last decades the number of victims of crowd related incidents is approximately 2000 per year (Lee & Hughes, 2006; Hughes, 2003). Most of the incidents occur at sport matches, concerts, festivals and nightclubs (Langston, Masling, & Asmar, 2006).

To tackle these public safety problems it is important to have an insight into crowds. Especially when a large number of people are gathering at a given time at events, for example a rock concert or a sport event (Smith et al., 2009). Crowds are generally constructed from small groups (Cartwright & Zander, 1968; Ge, Collins, & Ruback, 2009; Johnson, 1987). The unpublished study of McPhail shows that visitors of an event, in 89% of the cases, are accompanied by at least one other person (Ge et al., 2009). So, the crowd at events mainly consist of groups of minimally two persons, who thus also interact with each other. These groups consist mainly of friends or acquaintances who share an interest or like each other. These so-called self-formed

groups are not part of any institutional framework and do not have a leader installed by authority. This form of leadership is called emerging leadership and this kind of leadership has a larger influence over the group, in comparison with a leader installed by some authority (Sanchez-Cortes, Aran, Mast, & Gatica-Perez, 2010). The strength of a leader is his ability to transform individual action into group action (Hogg et al., 2006). Interventions could be more effective if the leader of the group will be addressed (Haslam, Reicher, & Platow, 2011).

The number of cameras in our daily lives increase quickly; in shopping malls, railway stations, concert halls and on the street. In London only tens of thousands of cameras are active in multiple Closed-circuit television (CCTV) systems (Boom, 2010). The main goal of these systems is to detect, prevent and monitor anti-social and obnoxious behaviour. More installed cameras does not directly lead to an increase in public safety. Having more cameras means that there is more information to observe, and thus a higher workload in observing all cameras. To make CCTV contribute to public safety is difficult. Despite the fact that cameras provide a wide angle of view and possibilities to focus and zoom, their intelligence and analytical capacities are limited. The functionality that is missing in a CCTV system is an intelligent tool that helps interpreting data. With the information provided from the tools, you can act right away when arriving at the location (instead of figuring out what the problem is at that moment).

The evolution of technology and the possibility of realtime video processing gives hope and new perspectives, but also leads to more questions. The use of CCTV could be useful in public safety and crowd observation applications. How can this be used to find group leaders, based on visual observable behaviour?

In the preceding research, the psychological aspects of this problem were analyzed. The main focus was the emergence of leadership from within the group. The group of interest is a small informal group. Compared to formal groups, leadership is not assigned by an authority, but has to emerge from within the group. Everybody is equal in an informal group, and has an equal chance to become the leader of the group. The research question was: "*How can CCTV information be used to find group leaders, based on visual observable behaviour?*". Small groups mainly consisting of four people were created and given a task. During this task leadership emerged from within the group. A questionnaire measured the personality characteristics and the perception of dominance and leadership of the team members. Video recordings were shown to multiple observers for interpretation as a validation measure. More details of these results can be found in Section 1.1.5.

In other research the behaviour of four people in a group has been ob-

served (Ashby et al., 2005; Hung & Gatica-Perez, 2010; Sanchez-Cortes et al., 2010). During these studies, people are placed in a chair around a table and given a special role to carry out. The presented study is comparable to these studies, with the exception that the group is free to walk through the room and there are no predefined roles.

The goal of this research is to investigate the possibilities of finding group leadership in a small self-formed group, based on CCTV data. In the psychological research initiative was found as one predictor for leadership within the group. Another proposed predictor from literature is the amount of movement. Where leaders and dominant people tend to move more than non-dominant people and non-leaders. The video recordings that are created during the psychological research will be used as input for this research. The precise question is:"*To what extend is it possible to use gesticulating as a measure for leadership when using CCTV data?*". First, a short theoretical explanation of the field of social groups, leadership and dominance is given, followed by a description of the data collection and the result of the algorithms for leadership detection.

This thesis is the sequel of a psychological study and consists of five chapters. The second part of this introduction provides a short explanation of the preceding psychological research, this will give a short introduction into the topic of social groups, leadership and dominance. In addition to this the collection of data and preparatory to this study is described. Firstly, a literature study is done in Chapter 2. The whole data collection process is described in Chapter 4. This also includes a description of the data validation. In Chapter 3, the algorithm for preprocessing and gesticulation measure is presented. The results of the experiments are described in Chapter 5. The final chapter, Chapter 6, contains conclusions and suggestions for further research.

1.1 Psychological research

This section will give a short introduction into the topics of social groups, leadership and dominance. Besides this, the collection of data preparatory to this study is described. In that study initiative taking is used as a predictor for leadership, it is found that the first person to start walking is more likely to become the leader of the group.

This research is about detecting leaders of small self formed groups. To get a better understanding of the context, short introductions are given about the topics of groups, leadership and visual observable characteristics of leadership. This will be continued with a summary of algorithms that can be used to accomplish the aim of this research.

1.1.1 Theoretical introduction

1.1.1.1 Groups

Groups will be addressed from a psychological point of view. A good definition is given by Sherif: "A social unit consisting of a number of individuals interacting with each other with respect to: Common motives and goals; an accepted division of labor, i.e. roles; Established status (social rank, dominance) relationships; Accepted norms and values with reference to matters relevant to the group; Development of accepted sanctions (praise and punishment) if and when norms were respected or violated" (Sherif, Sherif, & Murphy, 1956, p. 144).

The focus in this research is on the so called informal self formed groups. These groups are originated on a basis of mutual interest and gather on a regular basis. This group is not bound by any formal structure and thus free to do whatever they want. An example of such a group is a subsection of a football team that, after regular training hours, gather to go for a run. Another example is colleagues from different departments who go for a drink after office hours. Important here is that the groups are on basis mutual interest and not bound by some formal framework.

1.1.1.2 Leadership

Leadership is a process whereby an individual influences a group of individuals to achieve a common goal (Northouse, 2009). Because of the structure of informal groups, or lack of it, leadership emerges from within and is not installed by an authority (Côté, Lopes, Salovey, & Miners, 2010; Sanchez-Cortes et al., 2010). The roles within this group are self-organized and flexible, any member can become a leader at any time and thus is leadership context dependent (Vroom & Jago, 2007). This emerged leader has a strong position, his influence over the group is stronger than that of a leader installed by authority (Sanchez-Cortes et al., 2010). For an effective intervention within a crowd, the leader of such a group should be addressed (Haslam et al., 2011).

Close to leadership and associated with leadership is dominance. Dominance refers to the social control over the situation by forcing influence over others (Dovidio & Ellyson, 1982). The dominance personality trait is the tendency to behave assertive, forceful and self-assured (Anderson & Kilduff, 2009). Dominant people are more motivated to lead and to take over control. This is in line with previous research results where people who score high on a dominance scale are more likely to be picked as a leader (Kalma, Visser, & Peeters, 1993; Sanchez-Cortes et al., 2010).

1.1.1.3 Dominance

Dominance refers to the social control over the situation through influence over others (Dovidio & Ellyson, 1982). The personality trait dominance refers to the tendency to behave in assertive, forceful, and self-assured ways (Anderson & Kilduff, 2009; Buss & Craik, 1980; Wiggins, 1979). A high score in the dominance trait means more assertiveness and motivation to lead, which implies taking control. Research shows that taking over leadership by force is not enough, the social competence is an important aspect as well (Anderson & Kilduff, 2009; Van Vugt, 2006). Based on the scores on the social dominance scale, people with high scores on this scale are more likely to be selected as a leader than low scorers (Kalma et al., 1993). A high correlation is found between leadership and sociable dominance (Sanchez-Cortes et al., 2010).

1.1.2 Experimental description

The goal of the experiments is to let leadership emerge in small groups. This with the aim of finding visual observable predictors for leadership. The experiment is purely observable and consist of three parts: 1) get-to-know-games. 2) brainstorm session. 3) questionnaire.

1.1.2.1 Participants

A total of 124 participants, divided over 32 groups, participated in this research. The age of the participants differed between 18 and 25 years, with an average of 20.56 years (SD = 1.51). The distribution information of the sex and nationality can be found in Table 1.1.

Table 1.1: Sex and nationality distribution

	Country		
Sex	Germany	Netherlands	Total
Female	58 (46.8%)	41 (33.1%)	99 (79.8%)
Male	10 (8.1%)	15 (12.1%)	25 (20.2%)
Total	68 (54.8%)	56 (45.2%)	124 (100%)

1.1.2.2 Procedure

The experiment started with the participants gathering in the room where the experiment was conducted. When the group of 3 or 4 people was complete, the session began. First, to get to know each other, four simple team

building games were played. The first two games were get-to-know-each-other-games, the third game involved trust and coordination and the final game revolved around creativity. A more detailed description can be found in Section 1.1.2.3. The duration of this part was about 10 minutes. When the group was finished, the main task began. Here the group had to develop a game, based on certain criteria which were given in the assignment, more details can be found in Section 1.1.2.4. The time limit of this session was 20 minutes, after that, the results had to be presented. The final task for the participants was to fill in a questionnaire, see Section 1.1.2.5.

1.1.2.3 Team building games

A collection of 4 games is described on a piece of paper. The group had to complete each item on the list from top to bottom. The only item that was used, was a tennis ball to throw at each other. The first two games revolved around learning the names of the other team members. In the first game, the team members had to introduce themselves and in the second game the aim was about practicing the names. The third game is a task that requires trust and coordination to be completed. During this game the members had to stand in a circle and face each others back. The final goal of this game was to sit on each others laps. The game combines creativity and knowledge, within this game a series of country and city had to be stated alternately.

1.1.2.4 Group task

The participants were standing around a round table, see Figure 4.1, and got one assignment (see Appendix A), one whiteboard and one marker. In addition each participant was given one piece of plain paper and a pencil. The assignment described a illness, a therapy and a goal. The goal was to develop at least two games that meet the requirements of the therapy in such a way that the therapy gets more interesting for 7 to 10 year olds. After the development session one team member had to give a short presentation of the results.

1.1.2.5 Team Measure

The final measure consisted of six standardized and validated scales from different questionnaires. This final measure is conducted by pencil and paper. More details of these questionnaires can be found in the section 1.1.3 Final Measure.

1.1.3 Team Measure

The analysis consists of four kind of variables: 1) The introspection scales for dominance, measured as Responsibility, Self Esteem and Sociable Dominance Scale. 2) The observed scale for dominance, measured by the observations of the team members, labeled as Team Member Dominance. 3) The ranked observations for dominance and leadership. 4) The observed initiative taking scales Walking, Ball, Paper and Ball. The questionnaire also contains demographic information and the variables age and length.

1.1.3.1 Introspective Dominance

The three introspective scales are shortly described below.

Responsibility This scale contains only the items from the MMPI (Minnesota Multiphasic Personality Inventory)[1] dominance scale (Hathaway, McKinley, & Committee, 1989). A translation to a Dutch version is used (Derksen & de Mey, 1997). This 25 dichotomous item (0 = disagree, 1 = agree) test measures the personality trait dominance. This questionnaire is frequently used in mental health. An example question is 'I definitely have a lack of self confidence'.

Self Esteem The dominance scale from this Dutch Personality Inventory measures: initiative taking, managing other people and self-confidence within a group. The scale consists of 17 yes-no items, with reported Cronbach's alphas between the .70 and the .80. Cronbach's alpha measures internal consistency, for example for questionnaires. This is a value between 0 and 1, a value below .5 are unacceptable, between .5 and .6 is poor, between .6 and .7 is questionable, between .7 and .8 is acceptable, all above .8 is good (Kline, 2000). An example question is 'Within a group, I am mostly in charge' (Luteijn, Starren, & van Dijk, 2000). The test is stable over time, over a time span of 28 months a correlation is reported of r=.72 (Luteijn et al., 2000).

Social dominance This scale measures the dominance, expressed in social activity and attention. A higher score indicates a better relationship with the group members, and higher probability to be leader (Kalma et al., 1993). An example question from this scale is 'I have no problems talking in front of a group'. In other research, a Cronbach's alpha of .79 is found (Kalma et al., 1993).

[1] A personality test that is used in mental health

1.1.3.2 Team Member Dominance

Compared to the previous three scales, which are introspective scales, this scale uses context information. Each team member gives a score for every other individual team member. This scale contains 10 items, where each item consists of an adjective pair where one of the items is the inverse of the other. Each pair has the be scored on a 5-point scale. An example pair is 'dynamic - passive' (Manusov, 2005).

1.1.3.3 Ranking (intern)

Each group member is asked to make a ranking of the level of dominance of all members (including himself). Since the focus of this study is on leadership and not on peck order, the most dominant person is ranked as 1 and all the others as 0. This is based on the relation between leadership and dominance, as suggested in the literature. This variable is defined as Dominance Rank (DRank).

To extend this measure, a distinction could be made by ten dominance points that had to be divided over all group members. More points given indicates a higher level of dominance, this is described in the variable Dominance Points (DPoints). By dividing points, the difference of perceived leadership can be shown. To determine the perception of leadership it is also asked to make a ranking of group leadership, this variable is called Leadership Rank (LRank).

1.1.4 Observed initiative taking (first movers)

One of the observable predictors of leadership is initiative taking. Before the team building games started, the team members were standing literally with their backs against a wall. The instruction is given and the needed materials are put on the ground. This is used as starting point for measuring initiative. Four types of initiative are measured: 1) Walk away from the wall. 2) Pick up the ball. 3) Pick up the paper. 4) Start reading from the paper. To validate this measure, recordings of this were shown to five observers. They made a ranking of each of the initiative behaviours.

1.1.5 Results

The descriptive statistics are shown in Table 1.2. Per questionnaire, the Mean score and Standard Deviation is given for all participants. Based on the given answers, the Chronbach's alpha (α) is calculated. This is a measure for internal consistency for the questionnaire. The scores of the Chronbach's

alpha can be interpreted as follows: Values below 0.5 indicate no consistency at all, values between the 0.5 and 0.7 are questionable and values above the 0.7 are fine. The score range for Self Esteem and Responsibility is between 0 and 1, Sociable Dominance and Team Member Dominance scores range between 1 and 5. Both the Self Esteem and the Responsibility score are below half of the scale score (.5), quite a low score. The Sociable Dominance was around the average, which is with a score of 2.694. In comparison, the Team Member Dominance shows on average a higher score. The alpha's of the Self Esteem is quite low. The scale will not be deleted, but this needs to be taken into account when interpreting the data.

Table 1.2: Descriptive Statistics for al measures (N=124)

Questionnaire	Mean	SD	α
Self Esteem	.411	.165	.554
Responsibility	.341	.183	.777
Sociable Dominance	2.694	.647	.763
Team Member Dominance	3.332	.565	.986

1.1.5.1 Initiative taking (first movers)

The Intraclass Correlation (ICC) of the five observers, scoring the initiative taking of the participants, as described in Section 1.1.4, is calculated as a measure of agreement. This agreement is quantified as .85, with a statistical significant $p < .0001$. Within this analysis, 432 measures are used for each rater.

1.1.5.2 Correlational overview

The correlations in Table 1.3, show the statistical correlation between two variables, expressed in Pearsons r.

Pearsons r is also known as Pearson product-moment correlation coefficient, which is the centered standardized sum of the cross product of two variables. De domain of r is between -1 and +1, where 0 is the neutral point and has no correlation. The closer the value approximates 1, the stronger the relation between the two variables.

In the social sciences the following interpretation is given to the correlational values. A value between 0 and ± 0.09 can be interpreted as no correlation. Small or weak correlations are found between ± 0.1 and ± 0.3. Between the ± 0.3 and ± 0.5 is a moderate or medium correlation. All values greater than +0.5 and below -0.5 indicate a strong correlation (Cohen, 1988).

1.1.6 Discussion

A relation between leadership and dominance found is found in literature. This is result is supported by this study. The data in Table 1.3 shows this relation with the significant correlation between the variables Dominance Rank (DRKN), Dominance Points (DPNT) and Leadership Points (LPNT). A difference is found between the observed dominance behaviour and the introspective measures of dominance (in the variables Responsibility, Self Esteem, Social Dominance and Team Member Dominance). The introspective measures look at behaviour via introspection (self observation), while the Team Member Dominance measures behaviour by how it is perceived by others. One of the reasons for this difference could be context. In small self-formed groups there is no control by an institutional framework or authority. The role of leadership emerges from within the group and everyone has an equal opportunity to be the leader. Besides this, the role of leader can change at the occurrence of an event.

In the experimental setting is initiative taking used as an indicator for leadership. During the analysis four different actions of initiative taking could be distinguished. The best indicator for leadership that is found in this study is when someone starts walking first. This is indicated by the significant correlation between Walking and Leadership Points, Dominance Rank and Dominance Points.

Table 1.3: Correlations for all measures (N=124).

Trait	Dominance							Initiative			
Construct	Introspection			Observe	Ranking						
Variable	Resp	SE	SD	TMD	DRNK	DPNT	LPNT	Walk	Paper	Ball	Read
Responsibility	1.0	.412**	.322**	-.222*	-.077	-.102	.006	-.010	-.086	-.090	-.089
Self Esteem	.	1.0	-.255**	-.033	-.051	-.078	.007	.070	.175	.146	.106
Sociable Dominance	.	.	1.0	-.178*	.017	-.046	-.048	.096	-.135	-.123	-.068
Team Member Dom	.	.	.	1.0	.432**	.483**	.462**	.222*	.089	.099	.147
Dominance Rank	.	.	.	.	1.0	.794**	.693**	.346**	.276**	.305**	.346**
Dominance Points	.	.	.	.	.	1.0	.816**	.213*	.182*	.213*	.213*
Leadership Points	.	.	.	.	.	.	1.0	.245**	.218*	.202*	.158
Walking	.	.	.	.	.	.	.	1.0	.201*	.187*	.195*
Paper	.	.	.	.	.	.	.	.	1.0	.922**	.754**
Bals	.	.	.	.	.	.	.	.	.	1.0	.830**
Reading	.	.	.	.	.	.	.	.	.	.	1.0

Notes: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Chapter 2

Literature Study

Leadership is the capacity to translate vision into reality.

Warren Bennis

The number of Closed-circuit television (CCTV) systems increases quickly in our daily lives, on the street in shopping malls, railway stations and concert halls (Boom, 2010). The main goal of these systems is to detect, prevent and monitor anti-social, aggressive and obnoxious behaviour. The value of automatic analysis in human behaviour is undeniable and essential to safety and security (Burghouts et al., 2013). To make these CCTV systems effective and contributing to safety, intelligent tools are needed to detect unwanted behaviour.

Generally, public events elapse quiet and easy without any problems. In this case, the CCTV system is just used for monitoring the crowd. At the occurrence of an incident, the consequences can be terrible (Hijum, 2011). Before you reach this state you want as intelligent system, based on the CCTV input, to detect the unwanted behaviour that causes incidents. One of the behaviours that could lead to incidents and violence in a public setting is aggression (McEllistrem, 2004). Aggressive behaviour can be detected by a combination of verbal and/or non-verbal information. The actual detection of aggression are the outliers of the baseline of normal behaviour (Lefter, Rothkrantz, Burghouts, Yang, & Wiggers, 2011; Lefter, Burghouts, & Rothkrantz, 2012). When aggression is found during an event, it is not only necessary to intervene to the aggressor, but also the surrounding area.

From the groups of people at public events it is known that they come in 89% of the cases at least with one other person (Ge et al., 2009). These groups consist mainly of friends or acquaintances who share an interest or like each

other. These so-called self-formed groups are not part of any institutional framework and do not have a leader installed by the authority. This form of leadership is called emerging leadership and has a larger influence over the group, in comparison with a leader installed by some authority (Sanchez-Cortes et al., 2010). The strength of a leader is his ability to transform individual action into group action (Hogg et al., 2006). Interventions could be more effective if the leader of the group will be addressed (Haslam et al., 2011).

2.1 Leadership

After an internet literature search, two modalities for predicting leadership were found, verbal and non-verbal. For each modality and a combination of those two, the different aspects are discussed. This section will be concluded with a discussion about the modality chosen for this research.

2.1.1 Verbal (speech)

Verbal communication is a powerful way to expressing yourself. From social psychology it is know that verbal expression is positively correlated to status and dominance (Dunbar & Burgoon, 2005). Vocally expressive people are more dominant and also often have a high-status (Jayagopi, Ba, Odobez, & Gatica-Perez, 2008).

Four types of verbal expression will shortly be discussed. First the relation between dominance and the total contribution to a discussion is discussed. Second the relation with speaking time is discussed, followed by speaker energy. Fourth, the speaker turns and interruptions are explained.

2.1.1.1 Debate Contribution

Within a group debate, the members all contribute differently to the conversation. This is clearly visible in an assignment, where the group had to solve a problem by discussion (Bales, 1953; Bass, 1954). An asymmetric distribution in quantity of the contribution became visible for the group members. The differences remained stable over multiple discussions within a session. In groups with three, five or seven members, the member with the most input, contributes between 40 and 50% of all contributions (Bales, 1953).

Members who contribute a lot, obviously have an influence on the content, direction and outcome of a discussion. These members determine the direction of the group, while low input members tend to listen and follow the

lead of the high input members. Due to this behaviour, the high input members become more dominant over the low input members within the group (Bales, 1953; Bass, 1954).

2.1.1.2 Speaking Time

The feature speaking length is the time that a person speaks (Jayagopi et al., 2008). The literature from social psychology supports the result that speaking time is a strong predictor for dominance and leadership within a group (Mast, 2002). Two meta-analyses support these findings, with a strong effect size by 15 and 25 studies. To explain the relationship between speaking time and dominance, the Expectation States Theory can be used. This theory states that in task-oriented groups, the expected performance of the team members transforms into a self-fulfilling prophecy and becomes the basis for the differences in dominance within the group (Mast, 2002). The relation between dominance and speaking time is not perfect and context dependent. Great amounts of speaking time, does not directly mean a significant dominance. This might have to do with the involvement or personal interest in the topic of discussion. It can be said that high status or high dominant people talk more than their low status or low dominant counter parts (Mast, 2002).

2.1.1.3 Speaker Energy

Speaker energy refers to a set of labels that is used interchangeably; speech loudness, speech energy, speech tempo, pitch and vocal control (Jayagopi et al., 2008). Speaking loud and expressive has a negative connotation and is associated with attempts to dominate and anger (Costanzo, Markel, & Costanzo, 1969). People speak louder when they are trying to express intense anger, a dominant type of expressive behaviour (Kimble, Forte, & Yoshikawa, 1981). Extrovert people who are socially dominant, speak louder than socially introvert people, (Siegman, 1978, as cited in Kimble and Musgrove, (1988)). The loudness of speech seems to be a predictor for dominant behaviour (Kimble & Musgrove, 1988).

Results validate the presumption that speaker energy can be a predictor for dominance. Assertive people talk louder and more than unassertive people. Men also talk louder than female in mixed-sex discussion teams, observed by a team of independent raters (Kimble & Musgrove, 1988). The average speaking energy has as prediction accuracy of 66.7% to predict dominance correct (Jayagopi et al., 2008).

2.1.1.4 Speaker Turns

The number of times someone speaks or the number of times someone takes over the conversation is defines as a speaking turn (Jayagopi et al., 2008). Taking over the conversation is a typical indicator of taking control of the situation, a characteristic of dominant behaviour (Smith-Lovin & Brody, 1989). Although it is clearly visible what is going on, it is difficult to analyze interruptions. Interruptions are rare events, with little occurrences during conversations. Because of their infrequency, long conversations and huge datasets are needed to find them (Smith-Lovin & Brody, 1989). From studies on interruptions is found that men interrupt women more and masculine identities interrupt those with more feminine images more often. As discussed by Kallock et al. (1985, p. 40, as sited in (Smith-Lovin & Brody, 1989)) interruptions are an excellent mechanism for taking over the conversation and a effective measure for dominance. It is a successful mechanism to accomplish leadership and dominance in a discussion.

2.1.2 Nonverbal (movement)

Nonverbal communication contains many aspects for analysis. Only a small subset is discussed here. The relation between dominance and leadership is discussed in the context of the direction of sight, initiative taking and quantity of movement.

2.1.2.1 Direction of sight

Where is someone looking at during a conversation? Is the speaker looking at the ground or at the other people who participate in the conversation? Is the speaker looked at during the conversation and how does this influence the status. High status people receive more visual attention than low status people. People who rarely look at others during a conversation are perceived as weaker (Exline, Ellyson, & Long, 1975; Jayagopi et al., 2008). In line with this is the Visual Dominance Ratio (VDR), the proportion of time someone spends looking at the other while speaking over the the proportion of time spent looking at the other while listening (Dovidio & Ellyson, 1982; Exline et al., 1975). VDR quantifies visual dominance through active or passive participation. When this ratio increases, the strength of dominance also increases (Dunbar & Burgoon, 2005). High power people have a higher VDR than people with low power (Dovidio & Ellyson, 1982; Jayagopi et al., 2008). Dovidio and Ellyson originally defined VDR as a measure for dyads. In order to apply this in a multi-party scenario, M-VDR is developed. The looking-

while- speaking feature is redefined as when a person who is speaking looks at any participant rather than at other objects in the meeting (Hung, Jayagopi, Ba, Odobez, & Gatica-Perez, 2008).

2.1.2.2 Initiative (First movers)

In the evolutionary game theory is inclined that within a group, the person who takes the initiative is more likely to become the leader (Van Vugt, 2006). This theory is developed during World War 2 as an analysis tool for strategies during combat. Nowadays it has become a tool for studying social interactions and processes. The literature review of Van Vugt, Hogan, and Kaiser, is in line with this game theory and they found that initiative taking is positively correlated with leadership (Van Vugt, 2006). High self-esteem shows the same as initiative taking, namely a better chance to be picked as leader. When the self-esteem is high, it is more likely that this person shows initiative to act and emerges as group leader (Andrews, 1984). The opposite is also shown, shy students show a negative correlation with leadership (Judge, Bono, Ilies, & Gerhardt, 2002).

The preceding psychological research shows that initiative taking can be used as a predictor for dominance and leadership. Although statistical significant correlations are found, it needs to be noted that this leadership in self-formed small groups is context dependent. Someone could have the advantage of familiarity with the problem to obtain the leadership position.

2.1.2.3 Movement (Visual activity)

From social psychology it is known that dominant people are visual more active than non-dominant people (Mullen, Salas, & Driskell, 1989; Van Vugt, 2006). Visual activity and body movement contains many facets. Dominance is related to body movement (Coulson, 2004), posture (Carney, Hall, & LeBeau, 2005; Weisfeld & Beresford, 1982), head movement (Jayagopi et al., 2008; Mignault & Chaudhuri, 2003), gaze (Shang, Liu, & Fu, 2008) and facial expressions (Knutson, 1996; Mazur & Mueller, 1996) is found (Dunbar & Burgoon, 2005; Hall, Coats, & LeBeau, 2005; Lance & Marsella, 2007; Ridgeway, 1987). Dominant people have more body movement than non-dominant people. Those dominant people also claim notably more space with their bodies than their non-dominant counterparts (Jayagopi et al., 2008).

Automated leadership prediction in small groups performs well. The data that is used originates from a camera that only captures the head of a team member. The focus here is to capture subtle changes in the facial expression.

This method performs well with a score between the 62 and 83% correct (Hung & Gatica-Perez, 2010; Jayagopi et al., 2008).

The techniques discussed to predict leadership work well and can be applicable in many situations. In the context of public safety, this does not work out well. In a crowded environment, the use of vocal information is unfeasible. It is undoable to filter the voice of each individual. Besides that, a second problem is introduced, the mapping from voice to individual.

The information is useful, but the level of analysis will be coarse-grained. The technological evolution has not reached the state of capturing facial expressions of a crowd at a public event. So the expressed visual behaviour to analyze will also be less fine-grained. For this reason, the choice is made to use body movements. To be sure that everybody is always visible on camera, a fisheye from above is used.

In the next chapter will discuss the algorithm that is created. This algorithm uses the fisheye camera in the ceiling to capture the movements of the group. With the use of this camera, everybody is always visible. It is hereby assumed that the people will not sit on each other's shoulders (Which is due to the height of the experiment room a safe assumption). During the development of the algorithm, some challenges are faced. What to do with people that are taller, or people that wear cloths with a pattern compared to plain clothes?

Chapter 3

Algorithms

Any sufficiently advanced technology is indistinguishable from magic.

Arthur C. Clarke

The recordings from the ceiling camera that were made during the psychological study were analyzed. Before the actual gesticulation measure could take place, the recordings needed to be preprocessed. This preprocessing consisted of a series of steps executed in chain. The following steps can be distinguished and are described in this chapter.

First the starting point of the discussion had to be found and a section of three minutes is made. Second, the sections are cut into segments of ten seconds. The third step is to reduce the frame rate of the segments, where in the fourth step the color from the frames is transformed to gray scale. In step five the background is removed from the frames and only the people stay visible in the room. The quality of the image is increased in step six by removing the noise. The final step is the gesticulation calculation.

3.1 Find the beginning of the recording

The start of the brainstorm session is defined as the moment that the group is finished reading the assignment. This is characterized by flipping the assignment back to the front page of the assignment, see Appendix A.2. From this moment on a segment of three minutes is taken from the recording. Cutting the recording is done by hand with the use of ffmpeg[1], a tool that can cut and encode movies.

[1] `http://www.ffmpeg.org`

3.2 Divide the recordings

From the recording a 3-minute section is created. These sections are cut into segments of 10 seconds, as shown in Figure 3.1. To reduce the data and increase the processing speed, each third segment will be used in further analysis.

3.3 Frame-rate reduction

The video recordings are shot with a frame rate around the 25 frames per second. When analyzing the movie files, 251 frames could be collected. This resulted in 250 comparisons of frames. When comparing all adjacent frames, the time interval between them is around the 40ms. In this time span small movements as shaking and shivering with the hands become visible really well, but the bigger and broader movements got neglected. To avoid this issue and shift from micro to macro movements, the original number of 25 fps is reduced to 1.6 fps in the analysis. The reduction of frames is also done in the study of Jayagopi et al., here the recording is reduced to 5 frames per second (200ms) (Jayagopi et al., 2008). The focus here was on subtile movements with the head. Manual comparison of 251, 32, 16, 6 and 4 frame movies is performed to see which number of frames had the best visual result in showing macro movements. This reduction to 16 frames had the best results, which is equal to a frame rate of 1.6 frames per second, a reduction from 40ms to 625ms. A visual representation of Finding the beginning of the recording, Divide the recording and Frame-rate reduction is shown in Figure 3.1.

3.4 Color reduction

A color is a composition of three prime colors, red green and blue. The amount of contribution of each prime color is expressed in a value between 0 and 255 and can be represented in a range of 8 bits. An intuitive way of merging these these prime colors into one set of gray shades is by taking the average value of each prime color as shown in Equation 3.1.

$$gray = (red + green + blue)/3 \tag{3.1}$$

This method works fine and quick, but has some shortcomings. The transformation of luminosity (brightness) deviates. Pure green is much lighter than pure red and that is more brighter than pure blue. This is solved by

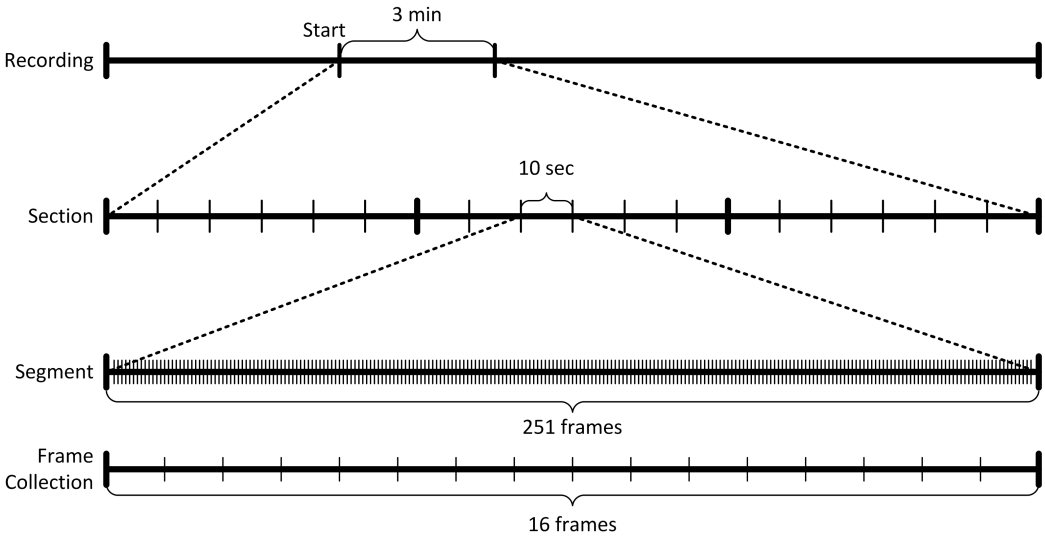


Figure 3.1: Separation and reduction of recording information

adding a weight factor to each color and results in Equation 3.2. Blue is the darkest and gets less weight than the others (Hunt, 2005, p. 408).

$$gray = (blue \times 0.114 + green \times 0.587 + red \times 0.229) \tag{3.2}$$

3.5 Background subtraction

Each frame is now represented in shades of gray. To get a clearer view of the people in the room, the background is removed. This is done by subtracting an empty shot of the room from each frame. The result of this subtraction is applied to a threshold. If the chosen threshold is to low, more pixels will be classified as background. When the threshold is to high, less pixels will be classified as background. For each pixel, if the difference between the current frame and the background image is smaller than a certain threshold, that pixel will be set to zero. For a formal notation see Equation 3.3, here $B_{(x,y)}$ is the background image with x and y coordinates, $F_{(x,y,t)}$ is the foreground image also with x and y coordinates and time indication or frame number t.

$$B_{(x,y)} - F_{(x,y,t)} < \tau = 0 \tag{3.3}$$

This results in a frame with objects that are not in the background image. In the most optimal situation, the whole frame is black, with 3 or 4 white blobs in it. The noise that is left behind can be removed with a noise

reduction filter, as described in Section 3.6. Each blob represents a person. In this case, the threshold is set to 15, this gives well recognizable people.

3.6 Noise reduction

Noise in digital images occurs during the conversion from analog, the real world, to digital conversion. A digital photo camera uses a CCD (Charge-coupled device) array as image sensors, which work on the photoelectric principle. When light reaches the sensor, electrons are produced and captured. Faulty electrons, for example caused by heat, are also captured by the sensor and cause noise. Although the behaviour of this product is uncontrolled, its distribution is gaussian (Bovik, 2005).

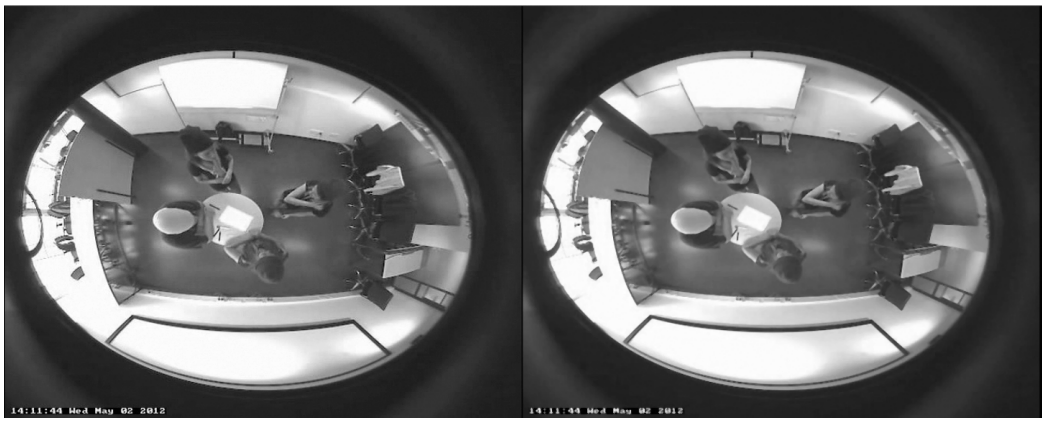
This noise could be removed with a relatively easy and commonly used robust technique, called median smoothing. With a median smoothing filter, the pixel is replaced with the center value of the set surrounding values, after ordering. The original value is included in this set. This filter is robust for extreme outliers of one of the neighbours. Besides this, the new value is a value out of the set, and not a newly calculated one, as shown in Equation 3.4. A graphical representation of an image before and after the noise reduction is shown in Figure 3.2.

$$r(i,j) = median\{x[i,j], (i,j) \in \omega\} \tag{3.4}$$

Computational complexity of median smoothing has an order of $O(n\ log_2 n)$. This is mainly caused by the sorting part of the algorithm. This complexity can be neglected since it is only used for a small set of items that need to be sorted (Huang, Yang, & Tang, 1979).

3.7 Gesticulation

It is found in social psychology that dominant people move more than non-dominant people (Mullen et al., 1989; Van Vugt, 2006). To measure movement, a quantification had to be made. The measurement of movement is defined as the amount of someone moves with their hands, arms and body during a conversation, expressed in changed pixels. This measure is mainly the movement of the limbs, but also the movement of one step off centre, since it is hard to stand still. If someone starts walking through the room, this does not count as gesticulation.



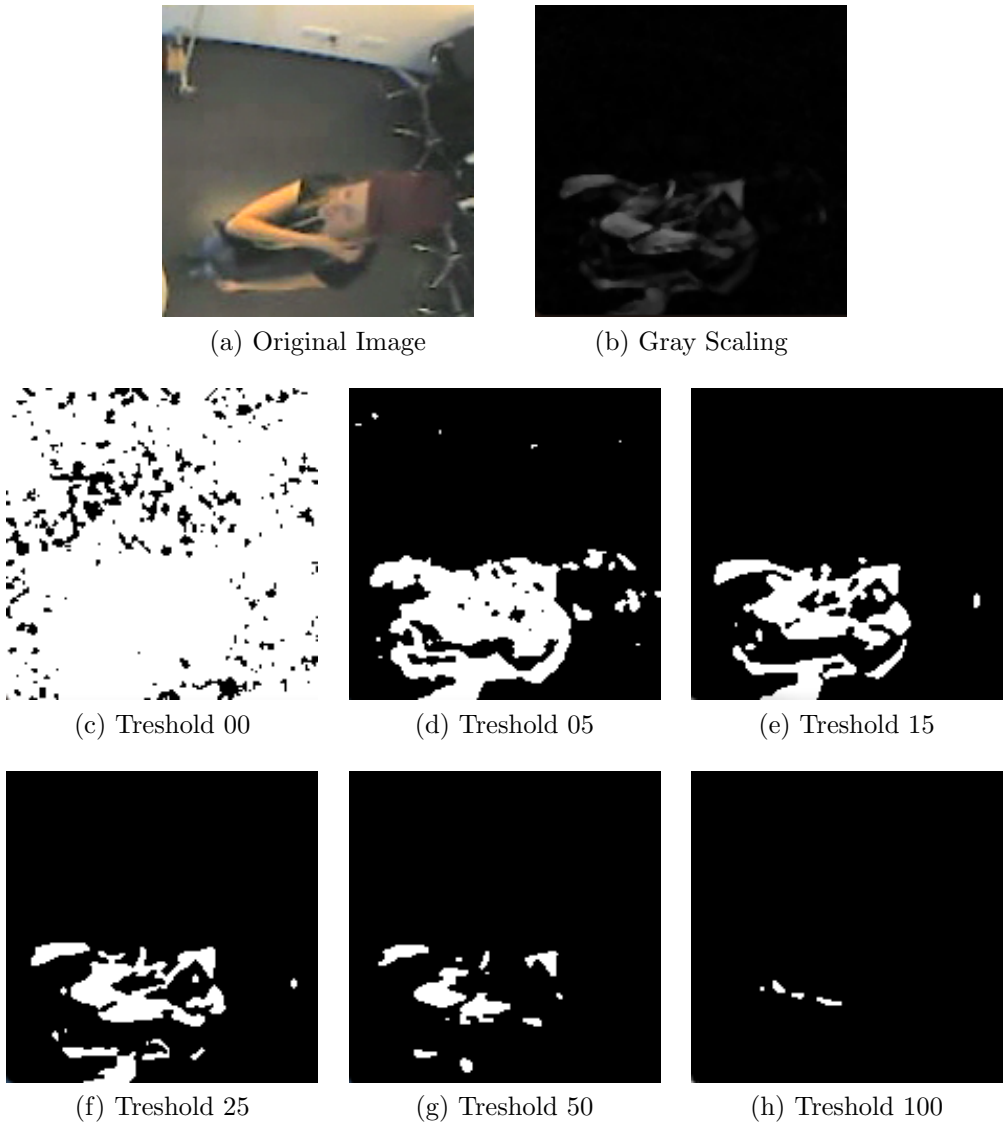
(a) Original image Gray Scale (b) Median Smoothing 3

Figure 3.2: Median Smoothing algorithm

After the frame rate reduction, the gesticulation can be measured. To achieve this, the color reduction is applied to reduce the amount of information in each frame to increase the processing speed. The background subtraction algorithm is applied to handle the issue of dependency on the pattern of the cloths.

When wearing a solid colored shirt, the movement is only visible on the edges of the person in the direction of the movement. While wearing a blocked shirt, the movement was also visible within the person, on the side of the blocks. This problem was tackled by introducing background subtraction and thresholding. A person now becomes, where all differences are below the threshold, a white blob on a black screen. Again, the threshold is 15, and a variation of thresholds is shown in Figure 3.3.

People who are taller are closer to the camera. By definition, these people fill a bigger area on the screen. When moving the same distance through the room as a smaller person, the amount of movement and thus gesticulation, is larger. To control this length issue, the values will be normalized. This means that the number of changed pixels is divided by the number of pixels above the threshold. This has some adverse consequences for larger people. The gesticulation movement index is now defined as the absolute difference between two adjacent frames divided by the square root of the number of pixels of the second frame above the threshold. The formula is stated in Equation 3.5, where i is the frame number, (x, y) is the pixel in the frame, and the result is a value between 0 and 1. The whole is multiplied by 100 to make it a percent score.



(a) Original Image (b) Gray Scaling

(c) Treshold 00 (d) Treshold 05 (e) Treshold 15

(f) Treshold 25 (g) Treshold 50 (h) Treshold 100

Figure 3.3: Multiple thresholds to visualizing people at subtraction of the background

$$Gesticulation\ score\ \Delta_i = \frac{\sum_{(x,y)} \mid I_{i+1}^{(x,y)} - I_i^{(x,y)} \mid}{\sum_{(x,y)} I_i^{(x,y)}} \times 100 \qquad (3.5)$$

3.8 Analysis

After processing each video file, per group member, 15 values of gesticulation are returned. For each collection of measures a mean and standard deviation is calculated. These values will be used in further analysis. Combined with this, a ranking per group will be made on the basis of the average score per video. Since this research is about leadership, the value of the leader, relative to the group is of interest. The rest of this section contains the different tests that are performed.

In the analysis the hypothesis and research questions answers will be answers. First, the difference in gesticulation between leaders and non-leaders is discussed. The following hypothesis will be tested:
H_0: $\mu = \mu_0$ In words: *The mean score of leader gesticulation is equal to the non-leader gesticulation.* This is tested for each of the six segments individually and for all of them together. A T-test will be sufficient to test the hypothesis. Besides this the difference between two samples is calculated. This is illustrated by the d'-value, used in the Signal Detection Theory. In case of two equal standard deviations $\sigma_1 = \sigma_2$ the equation

$$d' = \frac{(\mu_1 - \mu_2)}{\sigma}$$

is used. This is only applicable in a few cases. In many others, when the standard deviations are not equal, ($\sigma_1 \neq \sigma_2$) the appropriate measure of sensitivity d_a is used (Simpson & Fitter, 1973; Swets, 1986a, 1986b).

$$d_a = \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}}}$$

Based on the available distribution information, Receiver Operating Characteristics are plotted. This is a graphical representation to show the performance of a classifier system. It shows the True Positive Rate (TPR) as a function of the False Positive Rate (FPR) for different points. The closer the curve to the upper left corner, the better the performance of the classifier (Zweig & Campbell, 1993).

Second, the question how the gesticulation algorithm performs compared to the manual validation, is discussed. To answer this, a series of T-tests is performed. The TPR and FPR are calculated on the basis of a contingency table. The structure of the table is shown in Table 3.1. The TPR or hit rate is calculated by $(A/(A+B))$ and FPR or miss rate is calculated by $(C/(C+D))$. Based on these measures Receiver Operating Characteristics (ROC) can be plotted, which gives a visual representation of the performance.

Table 3.1: Contingency Table Structure

Event forecast	*Event Observed*		
	Yes	No	Marginal total
Yes	A	B	A+B
No	C	D	C+D
Marginal total	A+C	B+D	A+B+C+D = N

In the third question the predictability of leadership by the gesticulation algorithm is discussed. The measures will be included in the correlational table that looks like Table 1.3. This shows how the algorithm correlates with all the other measures of dominance and leadership. On the basis of the assumed ground-truth the TPR and FPR are calculated for the different recording segments. To visualize the performance, ROC-curves are plotted.

Chapter 4

Data Description

The supreme quality of leadership is integrity.

Dwight Eisenhower

4.1 Experimental data

In the conducted experiment from the psychological research, a collection of video recordings is created. During this experiment, the team members are recorded with five camera's. This data will be used as input for the algorithms. Next to the video recordings, a questionnaire from the psychological research is used as input data for analysis.

4.1.1 Video Data

The room where the experiment was conducted is equipped with five cameras. In each corner of the room a standard security camera, with a vision angle of approximate 64°, is mounted. In the center of the room, in the ceiling, a fisheye camera records from above. This angle of sight is nearly 187°. A visual representation of the room is shown in Figure 4.1. The colored slices represent an approximation of the visual area of the cameras. In the centre under the fisheye, a table is positioned where the people could stand around.

4.1.2 Questionnaire

The results from the questionnaire, as described in Section 1.1.3, are used and extended by the addition of the results from the gesticulation algorithm

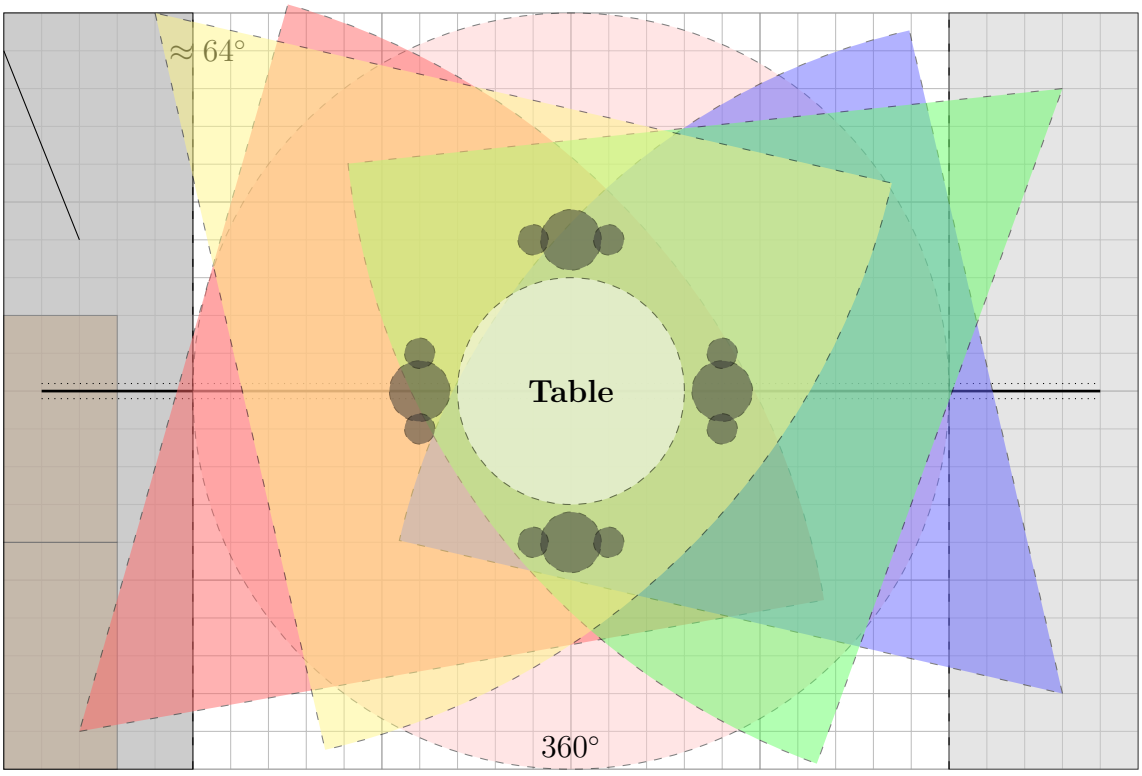


Figure 4.1: Visual representation of the experimental area

(Section 3.7) and manual validation (Section 4.2). This combined result will be used as dataset during the analysis.

4.2 Observed behaviour

It is found in literature that people who score high on the dominance scale, make more movements with their hands and body than people who score low on the scale. This movement of the hands and arms is called gesticulation and will be used as a measure for predicting leadership. Before an automated recognition can be used, a manual validation has to be performed. This validation is necessary to validate that the behaviour of interest is observable in the recordings.

4.2.1 Gesticulation

Recordings of 24 of the participating groups were suitable for the manual validation. During this validation four people ranked the members of each

group in sessions of 10 seconds each. A ranking is made of the amount of movement through the space. The person who had the most gesticulation was ranked as 1, the person with second most gesticulation was ranked as 2, and so on.

4.2.1.1 Procedure

Each observer is given a short description of the context of the experiment. The recordings from the fisheye camera are shown to four observers. In these recordings, all the participants in the experiment were always visible. It was not possible to hide behind each other. From each team six recordings of 10 seconds were shown, within a time frame of 3 minutes. It was possible for the observers to pause and replay the recordings.

The movie started at the moment that the group finishes reading the assignment and start discussing. The group had to turn the paper of the assignment, to see some images that are necessary for the discussion. At the moment that the page is turned, the selection of the recording starts.

4.2.1.2 Ranking

The first three minutes of the group task are analyzed. This section of the recording is cut into pieces of 10 seconds, where every third segment is used in the analysis, see Figure 3.1. Each member of the group is examined and ranked, based on gesticulation, within the group. The focus in this ranking is on the change in space that is used for movements of a person. Is this person gesticulating, making the personal space around as big as possible, active and attendant in the room?

A ranking of this movement will be made. The person that gesticulates most during the 10 second session will be ranked as 1, the second as 2, and so on. The table in the middle can be used as a static reference point.

Since the topic of this study is about leadership, the main focus is on the person who gesticulates most. The data from this manual analysis is transformed in such a way that the person with rank 1 will keep rank 1 and all the others get rank 0.

4.2.1.3 Data Analysis

The recordings that are shown are from the fisheye camera is on the ceiling. With this kind of camera comes transformation of the image to capture the whole area. The videos are shown unedited. The observers who ranked the 10 second segments have backgrounds in different areas. Two of them are

female and two of them are male with respectively ages of 21, 23, 17 and 27 year.

Based on the rankings from the observers, the Intraclass Correlation Coefficient (ICC) is calculated to measure the inter-rater reliability within the observers. ICC is a quantitative measure that describes the strength of consistency between the observers. To make a single measure of all observations, the most common value is taken. In case of a conflict, first the non-conflicting values are processed.

The Intraclass Correlation (ICC) of the four raters is calculated as a measure of agreement. This agreement is quantified in .816, with a statistical significant $p < .0001$. Within this analysis, 336 measures are used for each rater.

Chapter 5

Results

A leader is one who knows the way, goes the way, and shows the way.

John Maxwell

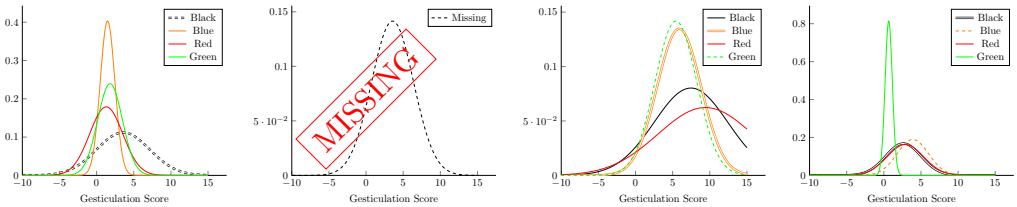
In this chapter the results of the analysis, as described in the previous chapter, will be presented. First the results of the hypothesis *'The mean score of leader gesticulation is equal to the non-leader gesticulation'* will be discussed. Second, the performance of the gesticulation algorithm is shown. Third the predictability of leadership is shown for the gesticulation algorithm.

In each adjacent frame, the mean and standard deviation was calculated for each individual team member. A visual representation of this is presented in Figure 5.1 - 5.6. Each column in the figure represents a group, where each row represents a moment in time. When reading the figure from top till bottom, the evolution of movement per group can be seen. Each graph has 3 or 4 lines in different colors, per person this color is unique. The line type differs per color. The '=' line represents the leader of the group on basis of the ground-truth. The dotted line '–' represents the leader on basis of the manual video observations. The combination of these two, a dotted double line' ('=='), is a leader both on basis of the ground-truth and on the basis of the observers. The curve with the highest mean score is the leader on basis of the algorithm.

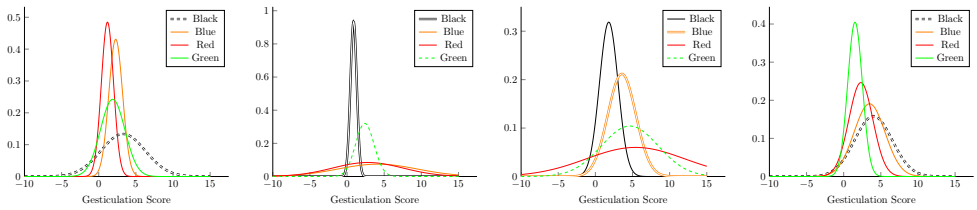
When analysing the plots in Figure 5.1 - 5.6, there are some peaks near the 0-value on the x-axis visible. This are people standing still and barely move. The lower and broader waves are people who move more. This behaviour, based on a quick visible observation is quite stable. In a few cases, the behaviour of almost no movement changes to movement. During the period of no movement, someone could be reading the assignment again or listening

with deep focus to someone talking

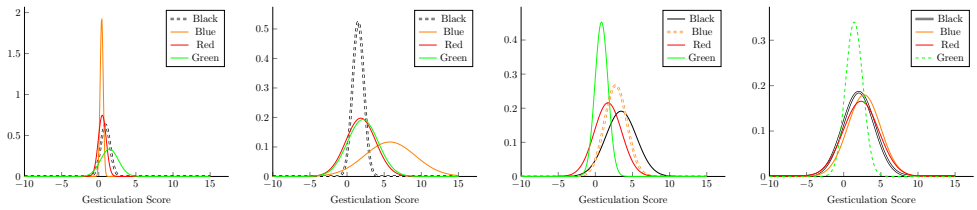
The graphs shows that in most of the analyzed videos, the leader, determined either from the observations or the ground-truth, has the highest score. The T-test confirms that on basis of the ground-truth, the leaders score significantly higher than non-leaders. These results are shown in Table 5.1, with the distribution information, p-value and d'. For interpretation, all the p-values smaller than 0.05 are statistically significant. All time slots, except for six, show a statistical significant relation. A graphical representation of the d' is given in Figure 5.7. From the table and the figure we learn that all the segments are close together. Except for segment 5, here the difference between leaders and non-leaders is bigger.



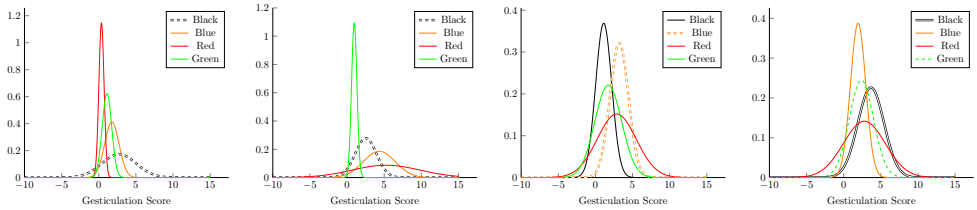
(a) 0401 - 1-1 (b) 0501 MISSING (c) 0801 - 3-4 (d) 0901 - 3-1



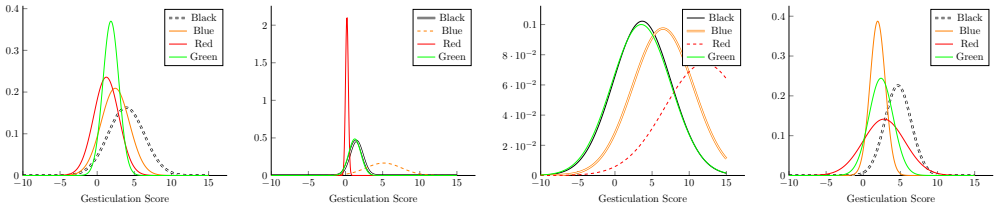
(e) 0404 - 1-1 (f) 0504 - 4-3 (g) 0804 - 3-2 (h) 0904 - 1-1



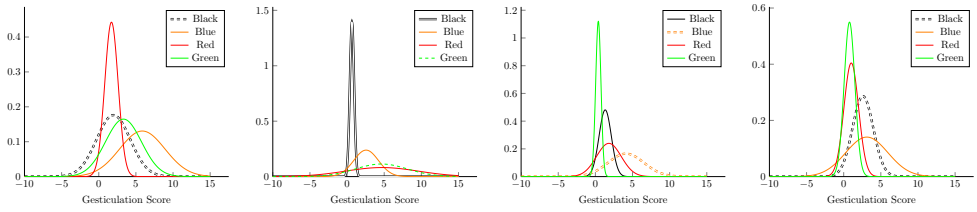
(i) 0407 - 2-2 (j) 0507 - 4-4 (k) 0807 - 2-2 (l) 0907 - 3-4



(m) 0410 - 1-1 (n) 0510 - 3-3 (o) 0810 - 1-1 (p) 0910 - 1-3

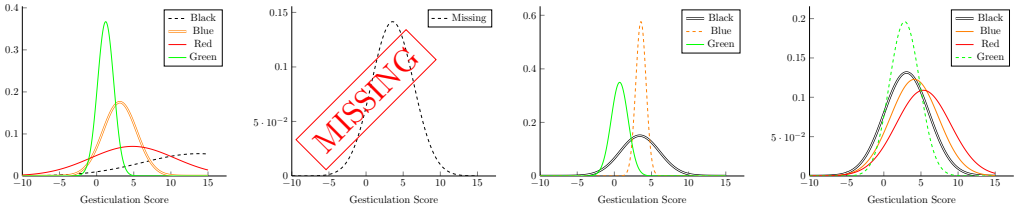


(q) 0413 - 1-1 (r) 0513 - 3-1 (s) 0813 - 2-1 (t) 0913 - 1-1

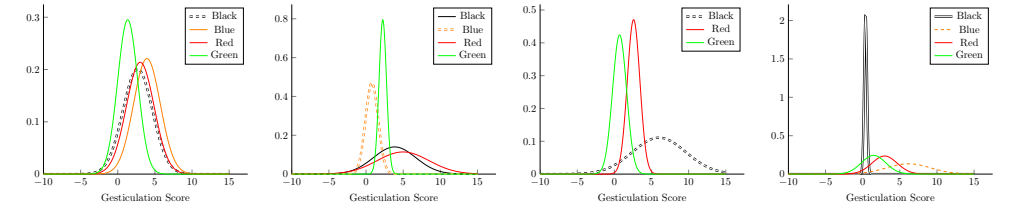


(u) 0416 - 3-3 (v) 0516 - 4-2 (w) 0816 - 1-1 (x) 0916 - 2-2

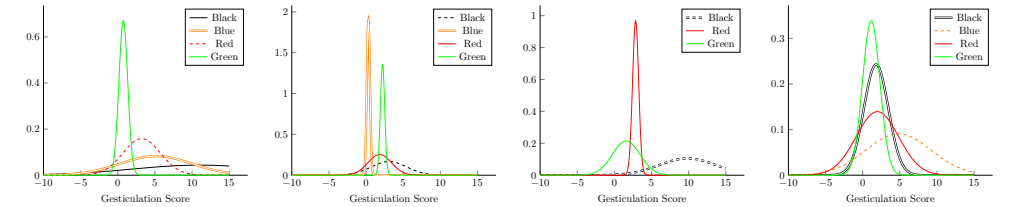
Figure 5.1: Groups 04, 05, 08 and 09



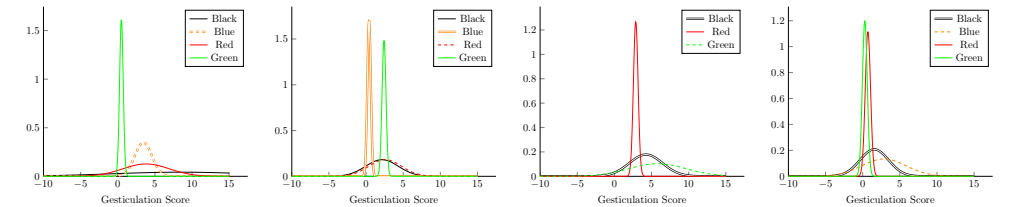
(a) 1001 - 3-1 (b) 1101 MISSING (c) 1201 - 2-1 (d) 1301 - 3-4



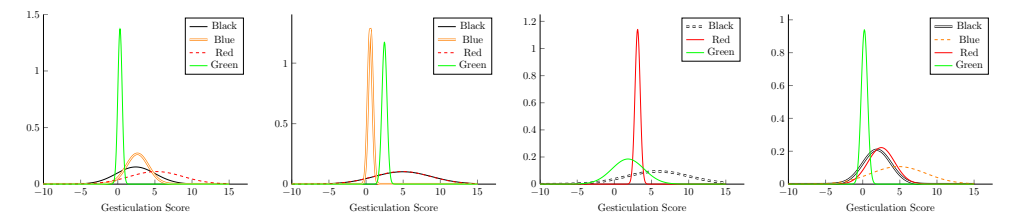
(e) 1004 - 3-3 (f) 1104 - 4-4 (g) 1204 - 1-1 (h) 1304 - 4-1



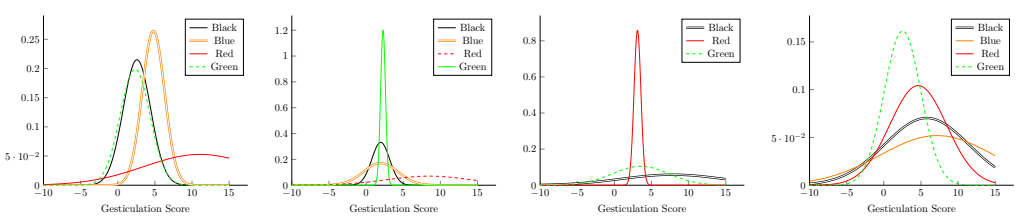
(i) 1007 - 2-3 (j) 1107 - 4-1 (k) 1207 - 1-1 (l) 1307 - 3-1



(m) 1010 - 3-3 (n) 1110 - 4-1 (o) 1210 - 2-1 (p) 1310 - 2-1

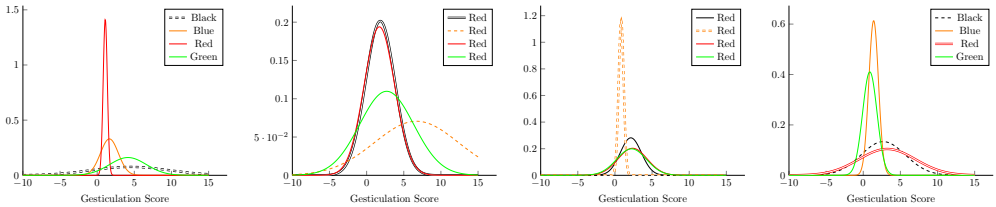


(q) 1013 - 2-1 (r) 1113 - 4-2 (s) 1213 - 1-1 (t) 1313 - 3-1

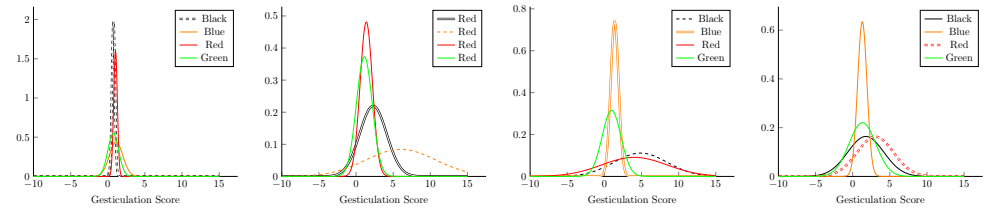


(u) 1016 - 2-4 (v) 1116 - 4-1 (w) 1216 - 1-2 (x) 1316 - 2-4

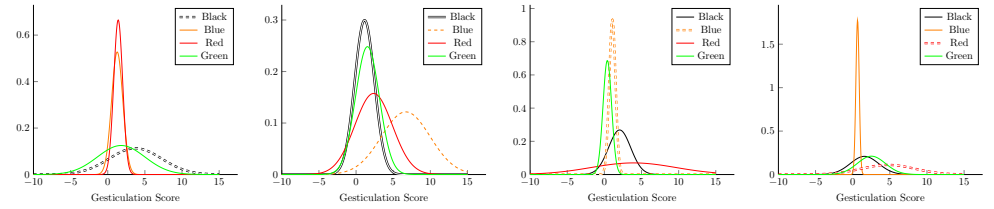
Figure 5.2: Groups 10, 11, 12 and 13



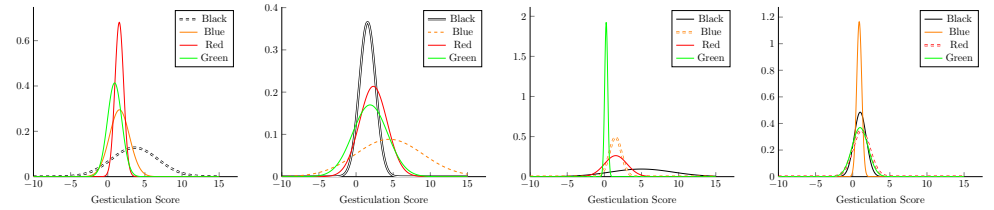
(a) 1401 - 1-1 (b) 1601 - 3-1 (c) 1701 - 4-4 (d) 1801 - 1-2



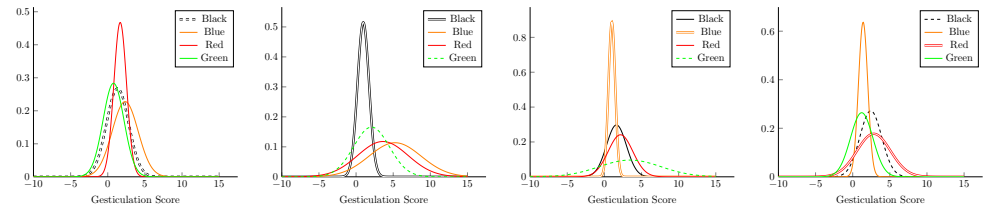
(e) 1404 - 3-3 (f) 1604 - 2-1 (g) 1704 - 3-1 (h) 1804 - 1-1



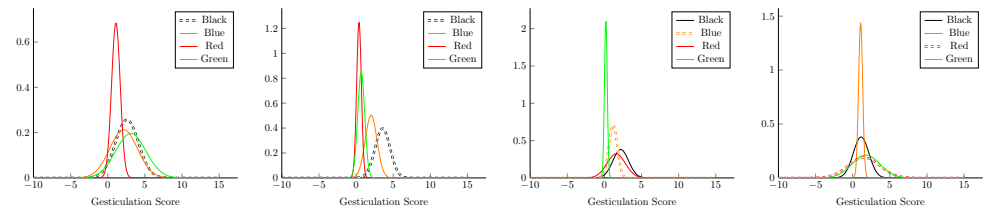
(i) 1507 - 1-1 (j) 1607 - 4-1 (k) 1707 - 3-3 (l) 1807 - 1-1



(m) 1510 - 1-1 (n) 1610 - 3-1 (o) 1710 - 3-3 (p) 1810 - 1-1

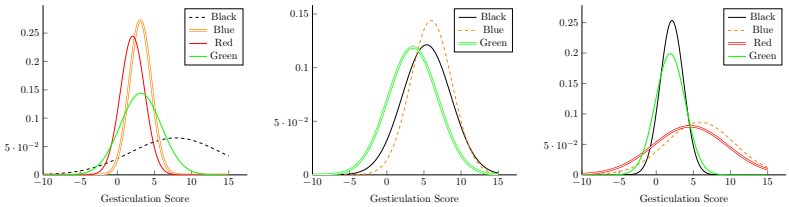


(q) 1513 - 3-3 (r) 1613 - 4-3 (s) 1713 - 4-1 (t) 1813 - 1-2

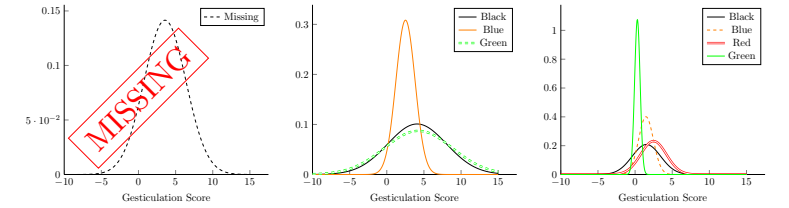


(u) 1516 - 2-2 (v) 1616 - 1-1 (w) 1716 - 3-3 (x) 1816 - 2-2

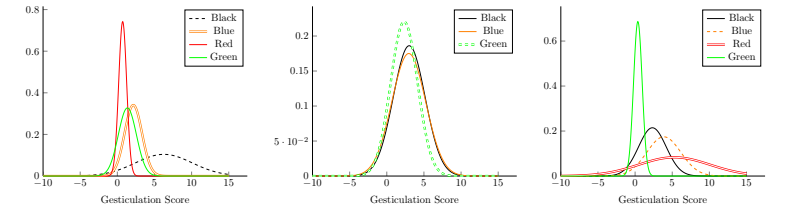
Figure 5.3: Groups 14, 15, 16, 17 and 18



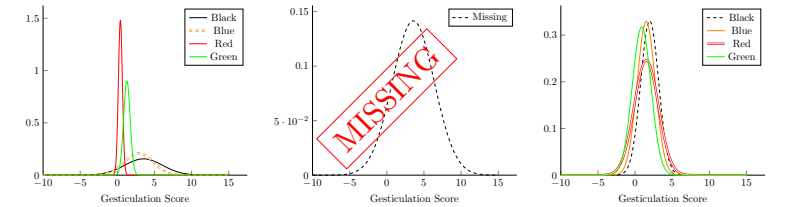
(a) 1901 - 3-1 (b) 2101 - 3-1 (c) 2201 - 2-1



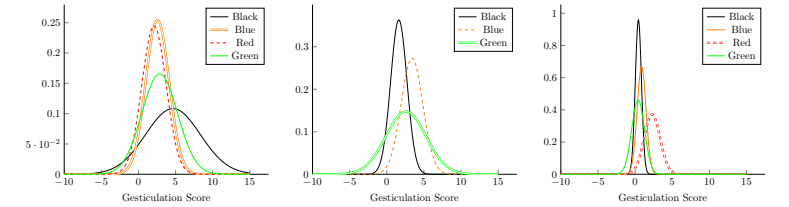
(d) 1904 - MISSING (e) 2104 - 1-1 (f) 2204 - 1-3



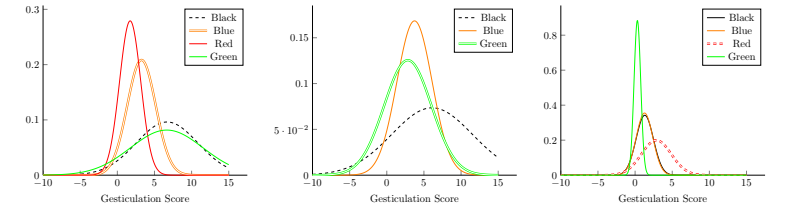
(g) 1907 - 2-1 (h) 2107 - 3-3 (i) 2207 - 1-2



(j) 1910 - 2-2 (k) 2110 - MISSING (l) 2210 - 2-1

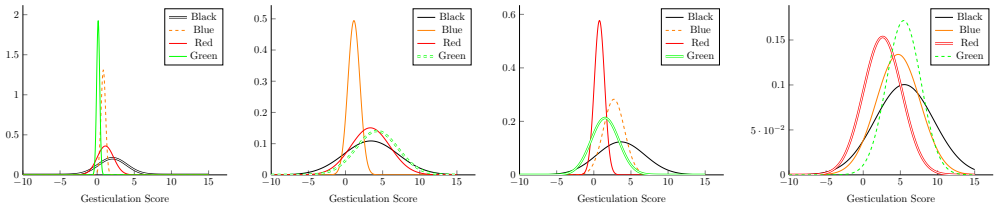


(m) 1913 - 3-4 (n) 2113 - 2-1 (o) 2213 - 1-1

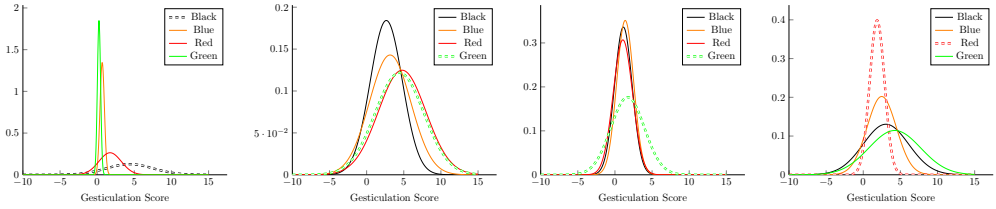


(p) 1916 - 2-1 (q) 2116 - 3-1 (r) 2216 - 1-1

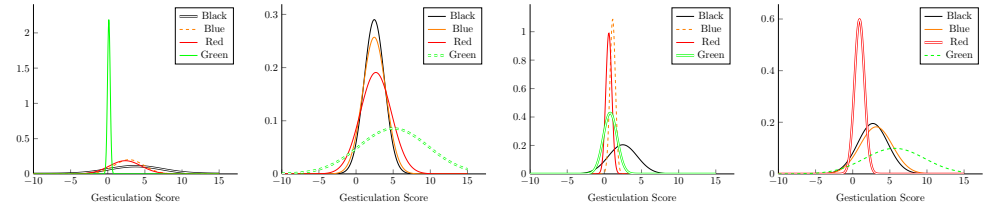
Figure 5.4: Groups 19, 21 and 22



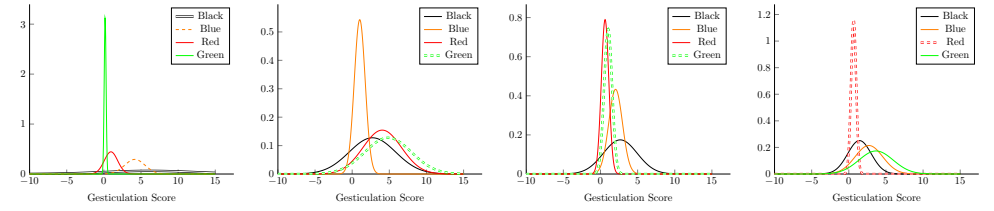
(a) 2301 - 1-3 (b) 2401 - 1-1 (c) 2601 - 3-2 (d) 2701 - 4-2



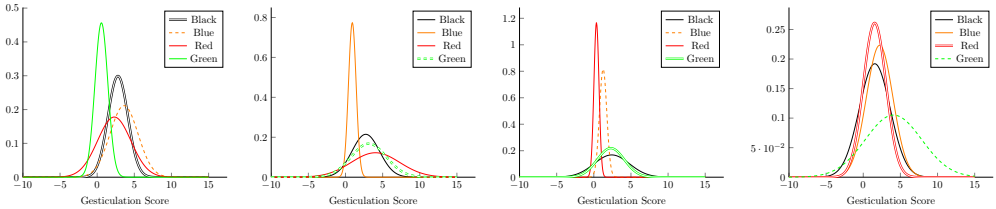
(e) 2304 - 1-1 (f) 2401 - 1-1 (g) 2604 - 1-1 (h) 2704 - 4-4



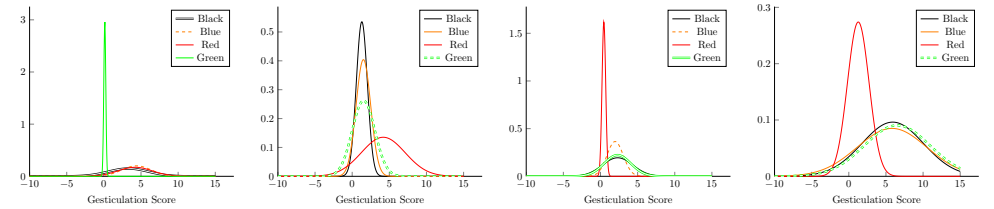
(i) 2307 - 1-2 (j) 2407 - 1-1 (k) 2607 - 3-2 (l) 2707 - 4-1



(m) 2310 - 1-2 (n) 2410 - 1-1 (o) 2610 - 3-3 (p) 2710 - 4-4

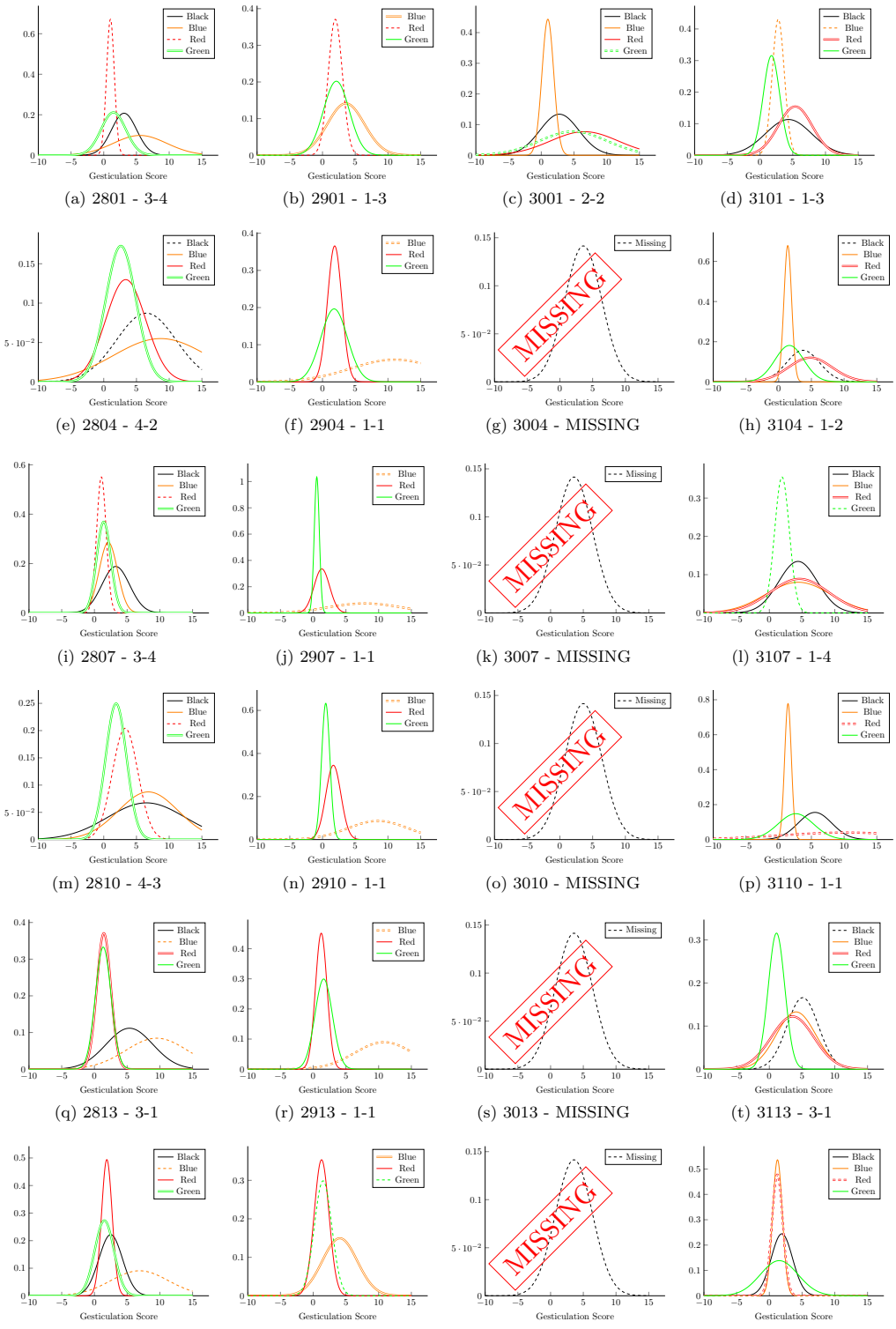


(q) 2313 - 2-1 (r) 2413 - 2-2 (s) 2613 - 2-3 (t) 2713 - 3-1



(u) 2316 - 3-1 (v) 2416 - 2-2 (w) 2616 - 2-3 (x) 2716 - 1-1

Figure 5.5: Groups 23, 24, 26 and 27



(a) 2801 - 3-4 (b) 2901 - 1-3 (c) 3001 - 2-2 (d) 3101 - 1-3

(e) 2804 - 4-2 (f) 2904 - 1-1 (g) 3004 - MISSING (h) 3104 - 1-2

(i) 2807 - 3-4 (j) 2907 - 1-1 (k) 3007 - MISSING (l) 3107 - 1-4

(m) 2810 - 4-3 (n) 2910 - 1-1 (o) 3010 - MISSING (p) 3110 - 1-1

(q) 2813 - 3-1 (r) 2913 - 1-1 (s) 3013 - MISSING (t) 3113 - 3-1

(u) 2816 - 4-1 (v) 2916 - 1-2 (w) 3001 - MISSING (x) 3101 - 4-4

Figure 5.6: Groups 28, 29, 30 and 31

Table 5.1: Results T-test gesticulation score leaders versus non-leaders

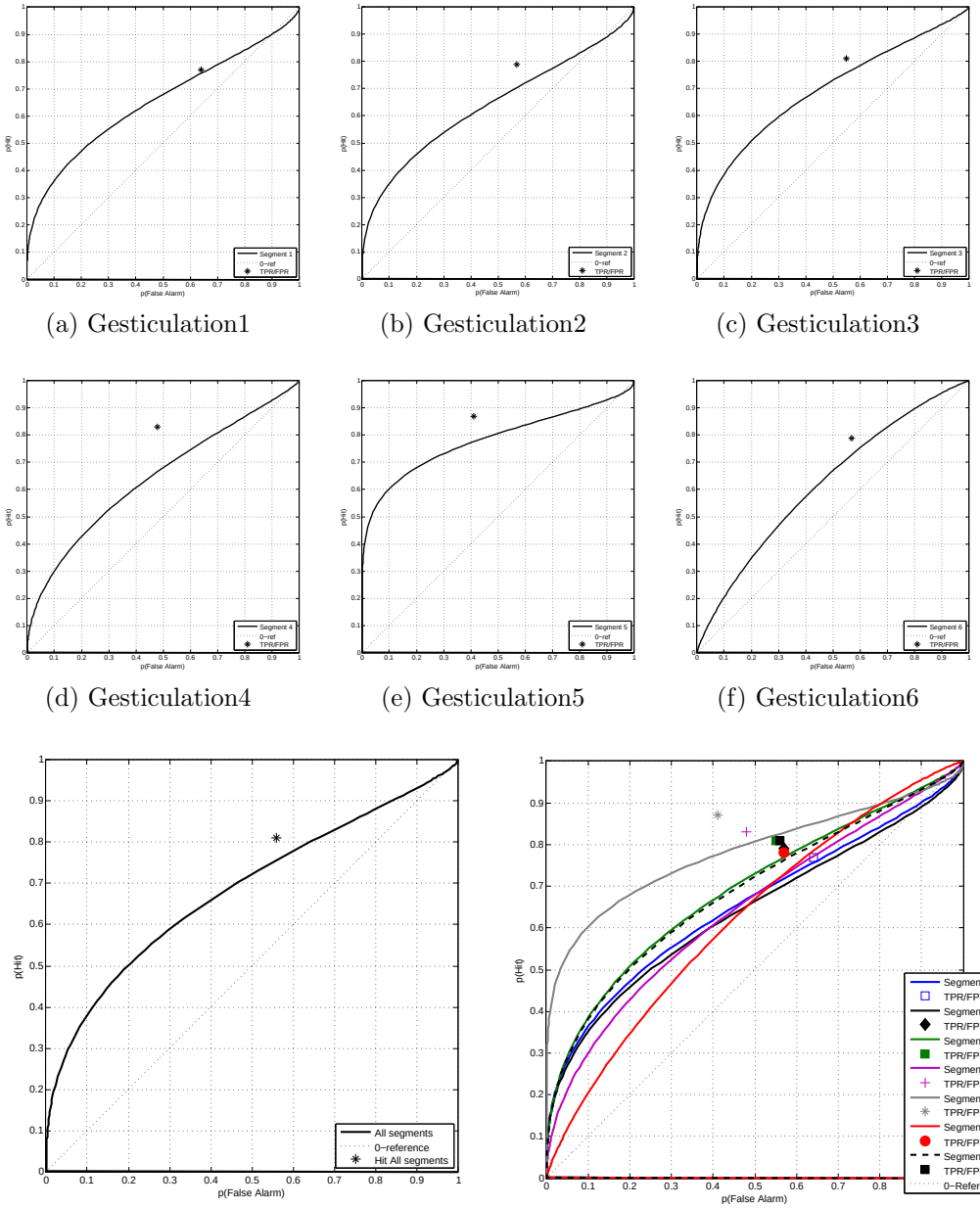
		Leader		Non-Leader		
Segment	P-value	M	SD	M	SD	D'
1	0.0134*	4.24	2.84	2.91	1.82	0.757
2	0.0011**	3.94	2.43	2.38	1.53	0.768
3	0.0024**	3.72	2.43	2.23	1.72	0.707
4	0.0313*	3.42	2.38	2.31	1.84	0.519
5	0.0001***	4.76	2.90	2.16	1.37	1.102
6	0.0806	3.58	2.14	2.65	2.11	0.437
all	0.0001***	3.93	2.52	2.44	1.76	0.685

$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

In order to test the performance of the gesticulation algorithm compared to the manual validation, the TPR and the FPR are calculated. For each segment, the score is based on the ranked score of the gesticulation algorithm and the observed validation rankings. The calculation can easily be done by plugging the data into the table, as described in Section 3.8. A Pearson Correlation test is performed on the same variables, to calculate the correlational strength. These results are shown in Table 5.2. In the table, a somehow proportional relation ship is visible between the TPR/FPR and the R-values. When there is a good TPR/FPR score, the R-score is also good and can even be statistically significant. A visual representation in the form of a ROC is shown in Figure 5.8. The results show that all the segments are close together. The segments four and five have the best score, this can both be seen on the correlation coefficient R and the TPR/FPR score in Table 5.2. This is only a performance measure of the gesticulation algorithm to the visual observation and has nothing to do with the performance on the detection of leadership. Later on in this section, the performance of the gesticulation algorithm is shown on basis of the ground.

In a comparison of the rankings between the gesticulation algorithm and the manual validation, the highest rank of the manual validation is compared with the rank based on the algorithm. The score of the rank comparison shows that 52.34% is ranked correct on rank 1. For the second rank this is 18.75 %, for the third rank this is 17.19% and 11.72% is ranked fourth.

The third question to discuss is the predictability of leadership. The dataset from preceding psychological research is reused. After adding the results from the new tests, a correlational overview is calculated. The results are shown in Table 5.3. It can be observed that the different standardized introspective dominance tests correlate significantly with each other. The



(a) Gesticulation1 (b) Gesticulation2 (c) Gesticulation3

(d) Gesticulation4 (e) Gesticulation5 (f) Gesticulation6

(g) Gesticulation All (h) Gesticulation Combined

Figure 5.7: ROC Ground truth - Gesticulation score

internal rankings also correlate high with each other. For the gesticulation scores, all correlate significantly with each other, except for segment 6. The scores of the Team Member Dominance test correlate with the internal rankings.

Table 5.2: True Positive Rate, False Positive Rate and Pearsons R of the rankings on basis of the gesticulation score compared to the observed score of the video segments.

Segment	TPR	FPR	R
1	.77	.64	.134
2	.79	.57	.222*
3	.81	.55	.261*
4	.83	.48	.357**
5	.87	.41	.469**
6	.79	.57	.216
ALL	.81	.56	

Notes: $^{*}p < .05$, $^{**}p < .01$

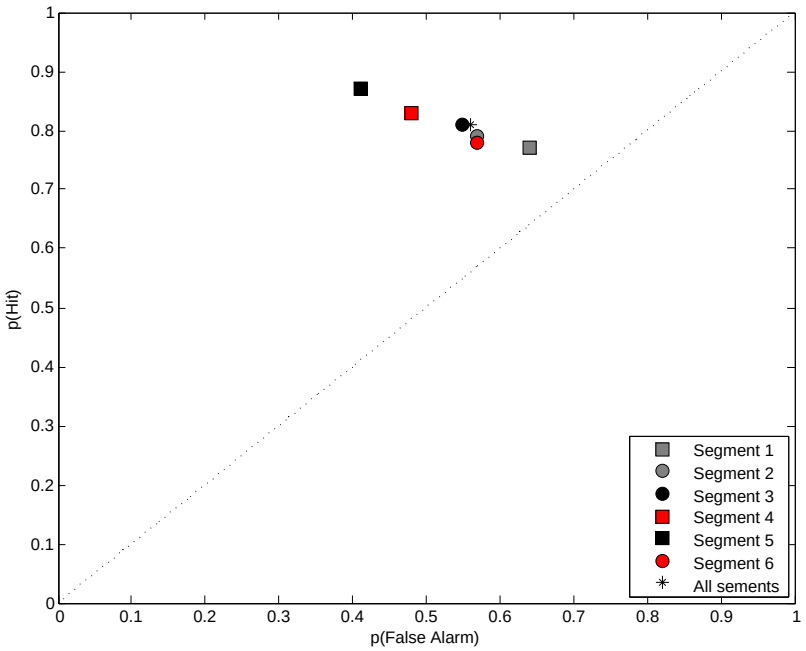


Figure 5.8: ROC performance gesticulation algorithm with observations as basis

As a final comparison, the hit and miss rate are calculated for the video observations and the gesticulation algorithm, compared to the ground-truth. The results from the psychological research are included for a complete overview, see Table 5.4 and Figure 5.9. From Table 5.4 can be observed that a high score is established for the Dominance Rank and Dominance

Points. With the ground-truth as a basis, human observations score better than the personality questionnaires. The gesticulation algorithm scores best of all as a measure for leadership. To interpret the values from the table, TPR is the True Positive Rate, which is the percentage of correctly identified cases. The FPR is the False Positive Rate, which is the percentage of false alarms. The best performance is obviously achieved when the TRP is high (near 1.0) and the FPR is low (near 0.0). The best performance for the gesticulation algorithm is segment 2 and 4, here the hit rate is quite high and the reject rate relatively low. The performance of the gesticulation algorithm is shown in a ROC in Figure 5.7. This means that leadership through the gesticulation measure is best detectable in segment 2 and 4.

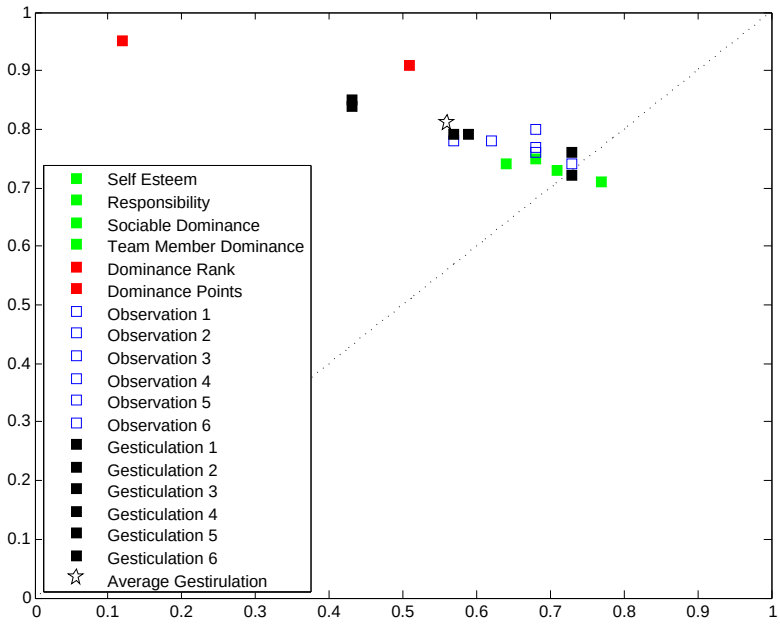


Figure 5.9: Visual representation of True Positive Rate and False Positive Rate of the rankings with the internal chosen leader as basis as shown in Table 5.4

When quantifying the score of the gesticulation algorithm on basis of the ground-truth, the scores are as follows. There is a 36.72% chance of scoring the leader correct, 19.53% to score the second in the peck order correct, 27.34% for rank 3 and 16.41% for rank 4. Effectively, this is an increase of 11% over the random guess. The visualizations in Figure 5.1 - 5.6 show that the curves of rank 1 and 2, in many cases, lie real close together. When combining the first and second rank, more than 31% increase is established.

Table 5.3: Correlations for all measures.

	Resp	SE	SD	TMD	Gest1	Gest2	Gest3	Gest4	Gest5	Gest6	Lpnt	Dord	Dpnt
Responsibility	1.0	.412**	.322**	-.223*	.251*	-.059	.143	.130	.034	.209	.006	-.077	-.102
Self Esteem	.	1.0	-.255**	-.033	.047	-.068	-.117	-.156	-.046	.090	.007	-.051	-.078
Sociable Dominance	.	.	1.0	-.178*	.077	.047	-.042	.067	.132	.168	-.048	.017	-.046
Team Member Dom	.	.	.	1.0	-.018	.221	-.002	.018	.206	-.005	.462**	.432**	.483**
Gesticulation 1	.	.	.	.	1.0	.414**	.554**	.532**	.279*	.142	.075	.015	.015
Gesticulation 2	.	.	.	.	.	1.0	.321**	.431**	.457**	.186	.222*	.157	.157
Gesticulation 3	.	.	.	.	.	.	1.0	.550**	.323**	.138	.076	.076	.015
Gesticulation 4	.	.	.	.	.	.	.	1.0	.357**	.100	.164	.229*	.100
Gesticulation 5	.	.	.	.	.	.	.	.	1.0	.261*	.076	.199	.138
Gesticulation 6	.	.	.	.	.	.	.	.	.	1.0	.015	.015	.076
Leadership Points	.	.	.	.	.	.	.	.	.	.	1.0	.693**	.816**
Dominance Order	.	.	.	.	.	.	.	.	.	.	.	1.0	.794**
Dominance Points	.	.	.	.	.	.	.	.	.	.	.	.	1.0

Notes: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Table 5.4: True Positive Rate and False Positive Rate of the rankings with the internal chosen leader as basis.

Questionnaire	TPR	FPR	R
Self Esteem	.74	.64	.007
Responsibility	.75	.68	.006
Sociable Dominance	.71	.77	-.048
Team Member Dominance	.73	.71	.462**
Dominance Rank	.91	.51	.693**
Dominance Points	.95	.12	.816**
Observation	TPR	FPR	R
Video Observation 1	.80	.68	.147
Video Observation 2	.78	.57	.416**
Video Observation 3	.76	.68	.199
Video Observation 4	.78	.62	.421**
Video Observation 5	.76	.68	.031
Video Observation 6	.74	.73	.216
Video Gesticulation	TPR	FPR	R
Video Gesticulation 1	.73	.73	.075
Video Gesticulation 2	.84	.43	.222*
Video Gesticulation 3	.79	.59	.076
Video Gesticulation 4	.85	.43	.164
Video Gesticulation 5	.76	.73	.076
Video Gesticulation 6	.79	.57	.015

Chapter 6

Conclusion and Recommendation

A journey of a thousand miles begins with a single step

CONFUCIUS

Many groups of at least two people visit public events such as a sport game or a festival. For public safety, CCTV systems are used to monitor the crowd. In case an intervention is required, it is advisable to address (this to) the leader of the group (for a greater effect). To make CCTV systems a more valuable addition to public safety, smart tools are needed. How can the data provided by CCTV systems be used to detect leaders of a group?

Small groups of four people are formed and given a brainstorm task with limited resources. During the execution of this task, recordings were made by five cameras. After finishing the brainstorm session, a questionnaire was conducted about leadership and dominance at personal and group level. The results of the questionnaire are used as a ground-truth for leadership. The recordings are preprocessed, manually validated, and analyzed for the availability and detectability of leadership characteristics. Movement of the arms and body, defined as gesticulation, is used as predictor for leadership in the group.

It is found that group leaders move more, and thus have a higher score on gesticulation, than non-leaders. The results of the questionnaire showed little predictive power for leadership. While the results from the manual validation showed better predictive power. The gesticulation measure showed an even better result as a predictive measure for leadership.

The results show that emerged leaders have a significant higher gesticu-

lation score and thus move more than non-leaders in small informal groups. This is both shown on basis of observations by humans and with the use of the gesticulation algorithm. From the results from the psychological questionnaire can be said that the emerged leadership is context dependent. The group members with a high score on questionnaire score do not automatically acquire the leadership position. The emerging character can for example rise from the familiarity with the subject or the importance to complete the task.

Previous work showed that in a small group with assigned team roles, thus also the leader role compared to emerging leadership, the leader showed more movement than non-leaders (Jayagopi et al., 2008). Another difference was the focus of movements, in their study it was on subtile movements of the head, in this study it is on macro movements.

Although a practical application is not feasible yet, the results of the current research, could be used to analyze CCTV footage from public events to increase the effectiveness of an intervention. Note that before this could be used, there are other problems that need to be tackled, for example finding a subgroup in a crowd. The results only show the person who is more likely to be the leader in the group, this leadership can change by at a different context or environment. Another issue is the increase of complexity when the problem size increases. In the experimental setting, the groups only consisted of four people. Although the efficiency of the software was totally neglected, the results could be computed in real time. This is an important issue that needs to be further investigated.

In the result section is shown that based on the ground-truth, the gesticulation algorithm performs well in some cases. The recorded segments that perform well are segment 2 and 4. These are the segments that occur just before minute 1 and minute 2 in the brainstorm session. It could be possible that leadership is not always visible or expressed. In the experiment, each third segment is analyzed and leadership is shown in two of these segments. Leadership is strongly shown in recordings 2 and 4. All participants from the experiment are university students, who should in general be able to perform a task individually. When the participants read or heard what the assignment was, they could start brainstorming with little guidance. At a certain point, decisions need to be made or need to be moved on to the next step. At theses moments, a strong sense of leadership could be shown and this could be in segments 2 and 4.

The results show a better performance than random guess in all of the recordings that were analyzed. The gesticulation algorithm was revised two times, so it is questioned whether the optimal algorithm is found and used. The gesticulation algorithm takes the length of a person into account by using normalized values. The algorithm can also handle different patterns in

clothing (solid color or block patterns). In order to further reduce the weight of the length parameter, the denominator in the equation could be square rooted. Due to the quadratic relation between between the surface and the perimeter, taller people get a higher score. To remove the small irrelevant movements, the number of frames is reduced to filter them out. It is not clear if there are some unseen factors that can increase the performance. This could for example be the person its own shadow or the impact of crowdedness, where people are standing close together. In the lab setting, the people are standing around a table and not close together. This makes it easy to observe the movements of a single person. When applying the same technique to a crowded environment, the observation of movements of a person becomes a more complex task.

It could also be questioned if the groups are formed properly. In the design of the experiment, it is assumed that the members within the groups don't know each other and all have the same sex, nationality and mother tongue. Due to availability and time constraints, this was not feasible and concessions had to be made. Finally, the groups consisted of a mix of the assumed properties. This could cause a bias where certain persons have a higher chance of becoming a leader. It is shown that, in general, taller people are often dominant over smaller people, male over females and that age has an influence (Van Vugt, 2006; Van Vugt et al., 2008).

From previous research it is known that leaders move more than non-leaders. This was known for subtile movements on the micro level where a close-up of the head is analyzed. This study changed the micro movements to macro movements, since these are better applicable in a real world setting, and showed that leaders move more than non-leaders. It was reported that leaders have more micro movements than non-leaders, when they are assigned to a certain role. The results showed that the findings form previous research can be broadened to emerging leadership.

In the context of a public event, where most of the visitors come in groups who can be classified as small informal groups, this can help to increase the effectiveness of the security officers. When the security officer can effectively steer people, the threat can be reduced quickly and the atmosphere and public safety can be back to a comfortable level. This reduces the chance of further escalation.

For future research, some changes can be made to increase the performance. In this research gesticulation is used as a measure to determine the rank and leadership status in a group. This measure could be replaced with, for example, the number of short movements. Next to the recordings from the ceiling, a huge dataset of recordings is available. There are recordings from the side that could be used, from here other information can be ex-

tracted. These allow better extraction of kinetic information about type of movements. Another possibility is to switch modality. Acoustic information is available from all sessions. From previous research is known that the amount of speech is a good predictor for leadership (Jayagopi et al., 2008).

This study was conducted in a lab with students who were put together and performed a task. This gives a nice introduction to the topic, but the step to a real life application is far away. One of the steps in between could be the use of existing small groups. Colleagues from work who gather after work, or members from a football team that go for a run outside training hours. How would these 'real' small informal groups perform on the same task as the students did? Or how do they behave when they are together as a group.

Enjoying a life performance of your favorite band at a festival, watching your favorite sport club win the cup, this are two examples of events that you want to enjoy in peace. Disturbance of these events reduces the pleasure and gives it a bitter taste. With the ubiquity of CCTV cameras, big brother is watching you, big brother is protecting you. Gesticulation shows some promising results and can be a useful addition to ensure our beloved public safety.

Chapter 7

Acknowledgments

Leadership is unlocking peoples potential to become better.

Bill Bradley

Here I wish to thank various people for their contribution to this project. Without them, this thesis could not be established;
Raymond Veldhuis I would like to express my appreciations to Raymond Veldhuis, my first and daily supervisor, for the supervision, suggestions and sparring sessions. It was always a joy for me, working under your supervision on the various subjects. Over the past years I have learned a lot from you. The first course in Biometrics where we identified people by their hearts, the meta analysis on the prevalence of PII in biometrics and this final project.
Peter de Vries Special thanks go to my second supervisor Peter de Vries. It was always a joy to come to our meetings and discuss the various obstacles we met along the way.
Staff Support The technical support from Henny Kuipers and Geert Jan Laanstra was excellent, more than welcome and deeply appreciated.
Alfred de Vries Support for recording the sessions in the lab.
Sparring partners I hope you guys enjoined the sparring sessions as much as I did, they were extremely useful for this study.
Participants Thanks for participation, without you there was nothing to analyze.
Jessica Heijmans Special thanks for all your support, questions and help.
My family Thank you so much for all those years of unconditional support, love and confidence in succeeding. It is questionable if the results were the same without you.

References

Anderson, C., & Kilduff, G. J. (2009, February). Why do dominant personalities attain influence in face-to-face groups? The competence-signaling effects of trait dominance. *Journal of Personality and Social Psychology*, *96*(2), 491–503. doi: 10.1037/a0014201

Andrews, P. H. (1984). Performance self-steem, and perceptions of leadership emergence: A comparative study of men and women. *Western Journal of Speech Communication*, *48*(1), 1-13. doi: 10.1080/10570318409374137

Ashby, S., Bourban, S., Carletta, J., Flynn, M., Guillemot, M., Hain, T., ... Wellner, P. (2005). The ami meeting corpus: A pre-announcement. In *Workshop on multimodal interaction and related machine learning algorithms (mlmi)* (p. 4).

Bales, R. F. (1953). *The equilibrium problem in small groups* (R. F. Bales & A. E. Shils, Eds.). T. Parsons.

Bass, B. M. (1954). The leaderless group discussion. *Psychological Bulletin*, *51*(5), 465 - 492.

Boom, B. J. (2010). *Face recognition's grand challenge: uncontrolled conditions under control.* Unpublished doctoral dissertation, University of Twente, Enschede. http://eprints.eemcs.utwente.nl/19351/.

Bovik, A. (2005). *Handbook of image and video processing.* Academy Press.

Burghouts, G., de Penning, L., Kruithof, M., Hanckmann, P., ten Hove, J., Landsmeer, S., ... others (2013). A search engine for retrieval and inspection of events with 48 human actions in realistic videos.

Buss, D. M., & Craik, K. H. (1980). The frequency concept of disposition: dominance and prototypically dominant acts1. *Journal of Personality*, *48*(3), 379–392. doi: 10.1111/j.1467-6494.1980.tb00840.x

Carney, D., Hall, J., & LeBeau, L. (2005). Beliefs about the nonverbal expression of social power. *Journal of Nonverbal Behavior*, *29*(2), 105-123. doi: 10.1007/s10919-005-2743-z

Cartwright, D., & Zander, A. (1968). *Group dynamics: research and theory.* Harper & Row.

CNN Sports, I. (2001, may). *Cnnsi.com - soccer - major stadium disasters.* Retrieved from `http://sportsillustrated.cnn.com/soccer/world/news/2000/07/09/stadium_disasters_ap/`

Cohen, J. (1988). *Statistical power analysis for the behavioral sciencies.* Lawrence Erlbaum Associates, Incorporated.

Costanzo, F. S., Markel, N. N., & Costanzo, P. R. (1969). Voice quality profile and perceived emotion. *Journal of Counseling Psychology*, *16*(3), 267 - 270.

Côté, S., Lopes, P. N., Salovey, P., & Miners, C. T. (2010). Emotional intelligence and leadership emergence in small groups. *The Leadership Quarterly*, *21*(3), 496 - 508. doi: 10.1016/j.leaqua.2010.03.012

Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, *28*(2), 117-139.

Derksen, J., & de Mey, H. (1997). *Mmpi-2 - minnesota multiphasic personality inventory -2*. Pen Tests Publishers.

Dovidio, J. F., & Ellyson, S. L. (1982). Decoding visual dominance: Attributions of power based on relative percentages of looking while speaking and looking while listening. *Social Psychology Quarterly*, *45*(2), pp. 106-113.

Dunbar, N. E., & Burgoon, J. K. (2005). Perceptions of power and interactional dominance in interpersonal relationships. *Journal of Social and Personal Relationships*, *22*(2), 207-233.

Exline, R. V., Ellyson, S. L., & Long, B. (1975). Visual behavior as an aspect of power role relationships. In Pliner (Ed.), *Nonverbal communication of aggression* (pp. 21–52). New York: Plenum.

Ge, W., Collins, R., & Ruback, B. (2009, dec.). Automatically detecting the small group structure of a crowd. In *Applications of computer vision (wacv), 2009 workshop on* (p. 1 -8). doi: 10.1109/WACV.2009.5403123

Hall, J., Coats, E., & LeBeau, L. (2005, 11). Nonverbal behavior and the vertical dimension of social relations: a meta-analysis. *Psychological bulletin*, *131*(6).

Haslam, S. A., Reicher, S., & Platow, M. (2011). *The new psychology of leadership : identity, influence, and power*. Hove, East Sussex [England]; New York: Psychology Press.

Hathaway, S., McKinley, J., & Committee, M. R. (1989). *Mmpi-2: Minnesota multiphasic personality inventory-2 : manual for administration and scoring*. University of Minnesota Press.

Helbing, D., & Johansson, A. (2009). Pedestrian, crowd and evacuation dynamics. In *Encyclopedia of complexity and systems science* (p. 6476-6495).

Hijum, v., H. (2011). *Handreiking evenementenveiligheid delen i en ii, de processen geordend.*

Hogg, M. A., Fielding, K. S., Johnson, D., Masser, B., Russell, E., & Svensson, A. (2006). Demographic category membership and leadership in small groups: A social identity analysis. *The Leadership Quarterly*, *17*(4), 335 - 350. doi: 10.1016/j.leaqua.2006.04.007

Huang, T., Yang, G., & Tang, G. (1979, feb). A fast two-dimensional me-

dian filtering algorithm. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, *27*(1), 13 - 18. doi: 10.1109/TASSP.1979.1163188

Hughes, R. L. (2003). The flow of human crowds. *Annual Review of Fluid Mechanics*, *35*(1), 169-182. doi: 10.1146/annurev.fluid.35.101101.161136

Hung, H., & Gatica-Perez, D. (2010, oct.). Estimating cohesion in small groups using audio-visual nonverbal behavior. *Multimedia, IEEE Transactions on*, *12*(6), 563 -575. doi: 10.1109/TMM.2010.2055233

Hung, H., Jayagopi, D. B., Ba, S., Odobez, J.-M., & Gatica-Perez, D. (2008). Investigating automatic dominance estimation in groups from visual attention and speaking activity. In *Proceedings of the 10th international conference on multimodal interfaces* (pp. 233–236). New York, NY, USA: ACM.

Hunt, R. (2005). *The reproduction of colour*. Wiley.

Jayagopi, D. B., Ba, S., Odobez, J.-M., & Gatica-Perez, D. (2008). Predicting two facets of social verticality in meetings from five-minute time slices and nonverbal cues. In *Proceedings of the 10th international conference on multimodal interfaces* (pp. 45–52). New York, NY, USA: ACM. doi: 10.1145/1452392.1452403

Johnson, N. R. (1987). *Panic at the who concert stampede: An empirical assessment*. National Emergency Training Center.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002, August). Personality and leadership: a qualitative and quantitative review. *The Journal of applied psychology*, *87*(4), 765–780.

Kalma, A. P., Visser, L., & Peeters, A. (1993). Sociable and aggressive dominance: Personality differences in leadership style? *The Leadership Quarterly*, *4*(1), 45 - 64. doi: 10.1016/1048-9843(93)90003-C

Kimble, C. E., Forte, R. A., & Yoshikawa, J. C. (1981). Nonverbal concomitants of enacted emotional intensity and positivity: Visual and vocal behavior1. *Journal of Personality*, *49*(3), 271–283.

Kimble, C. E., & Musgrove, J. I. (1988). Dominance in arguing mixed-sex dyads: Visual dominance patterns, talking time, and speech loudness. *Journal of Research in Personality*, *22*(1), 1 - 16. doi: 10.1016/0092-6566(88)90021-9

Kline, P. (2000). *The handbook of psychological testing*. Taylor & Francis Group.

Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior*, *20*(3), 165-182. doi: 10.1007/BF02281954

Lance, B., & Marsella, S. (2007). Emotionally expressive head and body movement during gaze shifts. In C. Pelachaud, J.-C. Martin, E. Andre, G. Chollet, K. Karpouzis, & D. Pele (Eds.), *Intelligent virtual agents*

(Vol. 4722, p. 72-85). Springer Berlin Heidelberg.

Langston, P. A., Masling, R., & Asmar, B. N. (2006). Crowd dynamics discrete element multi-circle model. *Safety Science*, *44*(5), 395 - 417. doi: 10.1016/j.ssci.2005.11.007

Lee, R. S., & Hughes, R. L. (2006). Prediction of human crowd pressures. *Accident Analysis & Prevention*, *38*(4), 712 - 722. doi: 10.1016/j.aap .2006.01.001

Lefter, I., Burghouts, G., & Rothkrantz, L. J. M. (2012). Automatic audio-visual fusion for aggression detection using meta-information. In *Advanced video and signal-based surveillance (avss), 2012 ieee ninth international conference on* (p. 19-24).

Lefter, I., Rothkrantz, L., Burghouts, G., Yang, Z., & Wiggers, P. (2011). Addressing multimodality in overt aggression detection. In *Text, speech and dialogue* (Vol. 6836, p. 25-32). Springer Berlin Heidelberg.

Luteijn, F., Starren, J., & van Dijk, H. (2000). *Tweede heziene npv handleiding*. Swets en Zeitlinger.

Manusov, V. (2005). *The sourcebook of nonverbal measures: Going beyond words*. Lawrence Erlbaum. Retrieved from `http://books.google.nl/books?id=1I8Y\_AvrjvEC`

Mast, M. S. (2002). Dominance as expressed and inferred through speaking time. *Human Communication Research*, *28*(3), 420–450.

Mazur, A., & Mueller, U. (1996). *Facial dominance.* London, JAI Press.

McEllistrem, J. E. (2004). Affective and predatory violence: A bimodal classification system of human aggression and violence. *Aggression and Violent Behavior*, *10*(1), 1 - 30.

Mignault, A., & Chaudhuri, A. (2003, June 1). The Many Faces of a Neutral Face: Head Tilt and Perception of Dominance and Emotion. *Journal of Nonverbal Behavior*, *27*(2), 111–132.

Mullen, B., Salas, E., & Driskell, J. E. (1989). Salience, motivation, and artifact as contributions to the relation between participation rate and leadership. *Journal of Experimental Social Psychology*, *25*(6), 545 - 559. doi: 10.1016/0022-1031(89)90005-X

Northouse, P. (2009). *Leadership: Theory and practice*. SAGE Publications.

Ridgeway, C. L. (1987). Nonverbal behavior, dominance, and the basis of status in task groups. *American Sociological Review*, *52*(5), pp. 683-694.

Sanchez-Cortes, D., Aran, O., Mast, M. S., & Gatica-Perez, D. (2010). Identifying emergent leadership in small groups using nonverbal communicative cues. In *International conference on multimodal interfaces and the workshop on machine learning for multimodal interaction* (pp. 39:1–39:4). New York, NY, USA: ACM. doi: 10.1145/1891903.1891953

Shang, J., Liu, Y., & Fu, X. (2008). Dominance modulates the effects of eye gaze on the perception of threatening facial expressions. In *Automatic face gesture recognition, 2008. fg '08. 8th ieee international conference on* (p. 1-6).

Sherif, M., Sherif, C., & Murphy, G. (1956). *An outline of social psychology.* Harper.

Simpson, A. J., & Fitter, M. J. (1973). What is the best index of detectability?. *Psychological Bulletin*, *80*(6), 481 - 488.

Smith, A., James, C., Jones, R., Langston, P., Lester, E., & Drury, J. (2009). Modelling contra-flow in crowd dynamics dem simulation. *Safety Science*, *47*(3), 395 - 404. doi: 10.1016/j.ssci.2008.05.006

Smith-Lovin, L., & Brody, C. (1989). Interruptions in group discussions: The effects of gender and group composition. *American Sociological Review*, *54*(3), pp. 424-435.

Swets, J. A. (1986a). Form of empirical rocs in discrimination and diagnostic tasks: Implications for theory and measurement of performance. *Psychological Bulletin*, *99*(2), 181 - 198.

Swets, J. A. (1986b). Indices of discrimination or diagnostic accuracy: Their rocs and implied models. *Psychological Bulletin*, *99*(1), 100 - 117.

Van Vugt, M. (2006). Evolutionary origins of leadership and followership. *Personality and Social Psychology Review*, *10*(4), 354-371.

Van Vugt, M., Hogan, R., & Kaiser, R. B. (2008). Leadership, followership, and evolution: Some lessons from the past. *American Psychologist*, *63*(3), 182 - 196. doi: 10.1037/0003-066X.63.3.182

Vroom, V., & Jago, A. (2007).
Am Psychol, *62*(1), 17-24; discussion 43–7.

Wei, Y., Guo, Y., Dong, D., & Li, D. (2012). Public places safety management evaluation of railway stations. *Procedia Engineering*, *45*(0), 240 - 247. (2012 International Symposium on Safety Science and Technology) doi: 10.1016/j.proeng.2012.08.151

Weisfeld, G., & Beresford, J. (1982). Erectness of posture as an indicator of dominance or success in humans. *Motivation and Emotion*, *6*(2), 113-131. doi: 10.1007/BF00992459

Wiggins, J. S. (1979, March). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology*, *37*(3), 395–412.

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (roc) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, *39*(4), 561-77.

Appendix A

Documents

Contents

A.1 Informed Consent

Informed Consent

Bij dezen geef ik toestemming voor het maken van video en audio opnames van mijn gedrag tijdens dit experiment. De gemaakte opnames zullen uitsluitend gebruikt worden voor wetenschappenlijkonderzoek en zullen niet buiten de Universiteit Twente gedistribueerd worden.

Alle gegevens zullen anoniem verwerkt worden, samengevoegd met die van andere proefpersonen, aan statistische analyses onderworpen en eventueel gebruikt in wetenschappelijke publicaties. Na het ondertekenen van het consent zal ik tijdens het gehele onderzoek mijn volledige medewerking verlenen en mij volledig inzetten ten behoeve van het onderzoek.

Het is voor het huidige onderzoek noodzakelijk dat je de opzet en inhoud van het onderzoek niet bekend maakt aan medestudenten die nog willen deelnemen. We gaan er vanuit dat je je hieraan houdt.

Alvast bedankt voor je medewerking.

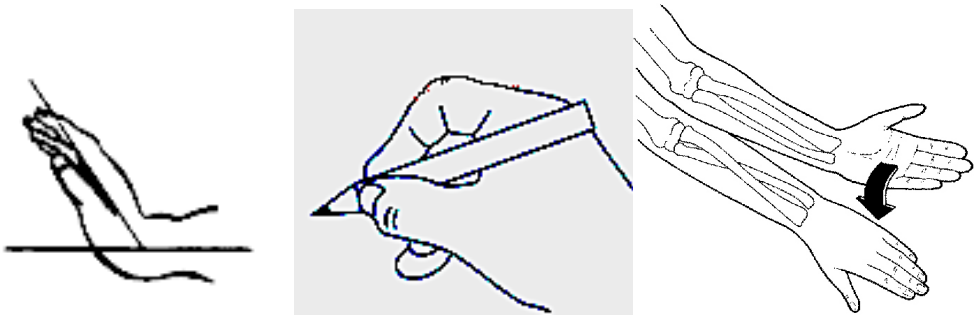
A.2 Experiment Desription

Experiment Description

Cerebral Paresis is an umbrella term for a group of disorders which result in a disorder of posture and or motor function. This is due to a permanent abnormality in the brain, caused in the first year of life. The prevalence of Cerebral Paresis is two till three occurrences per 1000 children, twice as much compared to Down Syndrome. Cerebral Paresis is the most common motor deficiency in early childhood. One of the symptoms of Cerebral Paresis is an above average muscle tension in the hand and arm (Cans, 2000).

Unfortunately, Cerebral Paresis can not be healed, but surgery and therapy can help to improve the abilities of the patients. The so-called modified Constraint Induced Movement Therapy (mCIMT) makes uses of the training joints in the arms and hands (Gordon, Charles, & Wolf, 2005). This therapy focusses on three types of movements:

- Dorsal flexion in the wrist (move the back of the hand to the arm).
- Pincer grasp (pick things up between you thumb and index finger).
- Supination and Pronation (rotate the the hand).



(a) Dorsal flexion (b) Pincer grasp (c) Supination

Figure 1: Training movements

Although this therapy is effective, a lot of repetitions in an exercise is not stimulating for children and a negative attitude can be developed against the treatment. For motivational purposes, the use of games can be used to make the treatment more interesting and pleasant for children.

The objective is develop at least two games that are suitable for children in the age category 7-10 years. These games should train the movements that fit the treatment of Cerebral Paresis, as show in Figure 1. The duration of this session is 30 minutes. Five minutes before the end I'll notice you with a buzzer. At the end of this session one of you has to give a short presentation of the games/findings. This presentation, should per game at least include why this game fits the treatment, why this game is preferred over the normal treatment, the materials that are needed to play this game. You can use the big paper to make some kind of a presentation and visualise your ideas. The duration of the presentation should not take longer than 2 or 3 minutes.

Good luck

References

Cans, C. (2000). Surveillance of cerebral palsy in europe: a collaboration of cerebral palsy surveys and registers. *Developmental Medicine & Child Neurology*, *42*(12), 816–824.

Gordon, A. M., Charles, J., & Wolf, S. L. (2005). Methods of constraint-induced movement therapy for children with hemiplegic cerebral palsy: Development of a child-friendly intervention for improving upper-extremity function. *Archives of Physical Medicine and Rehabilitation*, *86*(4), 837 - 844.