MASTER THESIS

POLARIZATION: IDENTIFYING EXTREME ATTITUDES

Submitted by

S.B. Prins

Department of Psychology of Conflict, Risk and Safety (PCRS)

In partial fulfilment of the requirements

For the Degree of Master of Social Sciences

University of Twente

Enschede, Netherlands

July 2018

Master's Committee:

Advisor: Peter de Vries
Co-Advisor: Renate Geurts

Abstract

This study attempts to scrutinize attitudes that police officers share within their work environment on potentially polarizing discourse. As in every layer of society, polarizing events have the potential to escalate certain conflicts and may lead to increasingly stronger attitudes. Possible dangers include a clash of personal preference versus policy, of which civilians' perception of police partiality may be the result. Written commentary of Dutch police officers were investigated for clues on the presence of polarization through linguistic analysis and administered surveys. Main goal was to detect which possible linguistic variables are optimal in detecting polarization. Surveys that were administrated included part of police commentary to detect which attitudes are perceived as polarized. Potential persuasive effects were explored as participants' attitudes may polarize themselves as well. By combining survey results and linguistic analysis the current study revealed that clues are present that the debate within the Dutch police force may have become polarized. Potent predictors for polarization include words that reflect anger and references to other people. Some officers seem to struggle to reconcile personal opinions with policy. Results are not generalizable to the entire Dutch police force.

Acknowledgement

I would first like to thank my thesis advisor dr.ir. P.W. (Peter) de Vries of the faculty Behavioral, Management and Social Sciences (BMS), Psychology of Conflict, Risk and Safety at the University of Twente. If I had any question I could always e-mail Peter and be ensured of a swift response. His continuing help and support has been invaluable for the progress made. He consistently allowed this paper to be my own work, but steered me in the right the direction whenever I needed it.

I would also like to thank drs. B.S. (Bas) Böing, operational specialist at the Police Netherlands who was involved as the employer and commissioner of this thesis. Without his input and expert knowledge this thesis could not have succeeded.

I would also like to acknowledge dr. R. (Renate) Geurts of the faculty Behavioral, Management and Social Sciences (BMS), Psychology of Conflict, Risk at the University of Twente as the second reader of this thesis, and I am gratefully indebted to her for her valuable comments on this thesis.

Table of contents

1. Introduction

In june 2016, the Dutch newspaper NRC (Kouwenhoven & Rueb, 2016) obtained excerpts of comments Dutch police officers could post within their closed intranet environment, following a discussion on the theme of ethnic profiling. The discussion waged within the police intranet painted a perfect picture of how personal attitudes may clash with professional code of conduct. Different assumptions are made about how policy should be interpreted in the apprehension of certain individuals. Some believe that preliminary attention should be given towards those of a specific ethnic background as well, while others exclaim that they feel ashamed about the behaviour of their colleagues and are accused of openly racist demeanour on the police intranet. It seems then that police officers wrestle with reconciling personal opinions with professional conduct. This can in turn be reflected in the behaviour officer's display in public and the publics' perception of impartiality may be threatened. When officers are seen as being selective in their behaviour and are not treating certain members of the public the same, the legitimacy of the police authority may be threatened and may directly hamper police effectiveness as citizens might refuse to obey orders. It is therefore necessary to detect more extreme attitudes and prevent these from polarizing, by becoming more extreme. Certain (polarizing) events have the potential to escalate conflicts and so lead to an extremization of attitudes. Avoiding polarization or addressing precursor events that may lead to polarization is therefore suggested.

Several events that were considered polarizing in Dutch society are for example allegations of ethnic profiling by the police, the current interpretation of a Dutch holiday called 'Saint Nicholas' (a mythical gift-bearing figure that hands out presents to children and is a part of Dutch tradition[1]) and a discussion whether or not police officers should be allowed to wear a Muslim headscarf. Several news outlets and research have led to accusations of impartiality of the police during events related to these topics, especially when civilians were exercising their right to protest (Laarhoven, 2018). Amnesty International has also criticized the Dutch police before for using disproportionate violence in the ending of a protest (Volkskrant, 2016). Furthermore, there seems to be division among both police officers on the street and

[1] Without going into too much detail, protestors claim that the Dutch tradition has racist characteristics and wish to ban its celebration, which evokes quite some response from counter-protestors as well.

managers within the police institute on whether or not ethnic profiling exists and whether this should be an issue or not (also see Çankaya, 2012; van der Leun et al., 2014).

Incidents like these portray a deeper-rooted debate about the clash of policy interpretation versus practise by Dutch police officers as they struggle with how policy might affect their daily jobs (Kouwenhoven & Rueb, 2016). The events portrayed above are several examples of themes that may have the potential to polarize the Dutch public, police officers themselves may also be a part of that. Amsterdam already had small public disturbances between proponents and opponents on several potentially polarizing subjects like Muslim discrimination, anti-refugee sentiment, the Israel-Palestine conflict, the conflict between Turks and Kurds and a repeating discussion about racism within the celebration of the Dutch tradition of 'Saint Nicholas' (Böing, personal communication, 2016)[2]. Within these conflicts, Dutch police is taking a central role as peacekeepers, making sure conflicts do not spiral out of control. The preferred role of the police within these debates is somewhere in the middle, where police officers attempt to reconcile opposing parties (Böing, personal communication, 2016). In order to do so, police officers need to gather social capital within each of these communities (Böing, personal communication, 2016). One of the pitfalls of taking such a position in times of conflict is that police officers may be viewed as a scapegoat. When the police choose to intervene in only one side of the conflict, they may be viewed as biased – thus becoming a target within these debates. To prevent such things from happening, it is necessary for the Dutch police to act and appear unbiased in order to gain trust and legitimacy in the eyes of the public.

The events portrayed above have shown that remaining impartial in the eyes of the public is a challenge of its own. The symptoms displayed above seem to be systematic and shed a light on how one of the cornerstone principles of the Dutch government is upheld – the principle of neutrality (Lettinga & Saharso, 2016). Questions could be asked for which themes it is challenging for the Dutch police to uphold professional conduct without letting personal opinions interfere with their duty. It is therefore imperative to know what is discussed internally as the attitudes officers hold could potentially negatively impact the behaviour officers should display on the street. An example of escalated polarization is the rise of civil rights movements in the United States due to perceived police injustice by the public (Siff, 2016).

[2] Source derived from the intranet of the police (not publicly available) in the Netherlands.

While the Dutch newspaper 'NRC' seems to suggest that the police work force is internally conflicted (Kouwenhoven & Rueb, 2016), empirical research is needed to ascertain this in an objective manner. This paper will attempt to investigate police officers' written texts for indices on the presence of polarization within threads on the police intranet. The main point of focus here is whether the Dutch police, in their role as civil servant, display evidence for polarization in the form of attitude extremity.

The following chapters will further explain the concept of polarization, its psychological roots and the measurement of its symptoms. Furthermore, a closer look will be taken at social identity processes and conflict escalation as both might be related to the concept of polarization. Persuasive elements found within polarized debates will also be handled, as clues suggest that reading or partaking within a polarized debate might lead to further attitude extremity and therefore further polarization. It is hoped to attain valuable insights in the inner workings of social identity processes online by investigating Dutch police commentary.

2. Theoretical framework

2.1 What is polarization?

The main concept of this paper is the concept of polarization, and is defined as an extremization of attitudes. A potential consequence of polarization is the loss of common middle ground due to partisan-motivated reasoning and dislike of each other (Robison & Mullinix, 2016; Lelkes, 2016). When attitudes diverge this may lead people to instead seek out other people with more similar attitudes – and so contact with those that have a conflicted opinion is minimized (Suhay, Blackwell, Roche, & Bruggeman, 2015). This fostering of bonding capital then comes at the cost of bridging capital as diverging groups are formed that generally do not want to deal with each other. The result is a decrease in trust and interest in compromise with heterogeneous others and may lead to conflict between each of the poles in a polarized debate (Anderson, Brossard, Scheufele, Xenos & Ladwif, 2013; Levendusky, 2013; Lelkes, 2016). This can then further escalate to tension and may even occur on the political level (DiMaggio, Evans, & Bryson, 1996; Noel, 2014). Esteban and Schneider (2008) also related polarization to growing divisions within a society of individuals or between nation states and claims that the concept of polarization was often linked to armed conflict and the use of military force (for examples see McDoom, 2012 or Borge-Holthoefer, Magdy, Darwish, & Weber, 2014). Other effects may be the breakdown of social and political order, political instability, strikes and other ways to hurt a competing group (Esteban & Schneider, 2008).

Several explanations exist on what causes polarization, such as persuasive arguments theory (Vinokur & Burnstein, 1974; Isenberg, 1986) – which suggests that a shift in attitude is dependent on the number and quality of persuasive arguments presented during a group discussion. This theory however leaves out any social factors that may lead to polarization (Pavitt, 1994).

Another approach is social comparison theory (Lamm & Myers, 1978) and directly competes with the former theory. Social comparison states that groups polarize as a result of individuals conforming to group norms. The individual then presents his attitude as similar but slightly more extreme for peer approval and personal pride. Digressing from group norms leads to shame and disapproval – which people generally want to avoid. Peer approval and pride is also generated by differentiating from the norms of (disliked) out-groups (Lamm &

Myers, 1978; McGarty et al., 1992; Suhay, 2015). Both theories do not fully explain the concept of polarization however as one only focuses on psychological aspects and the other exclusively on social factors. Isenberg (1986) suggests both theories are valid explanations and found a relationship between both theories and polarization. It is therefore likely a combination of both theories explain the concept of polarization best. Abrams, Wetherell, Cochrane, Hogg and Turner (1990) furthermore add that self-categorization is a crucial condition for group polarization – their experiments have shown for example that when categorical differences are made between two subgroups within a discussion group, attitudes do not seem to converge. Which group identity is salient is then a determining factor.

Different interpretations and measurements exist on this seemingly "fuzzy" concept of polarization. It is therefore necessary to clearly explain how the concept of polarization is approached within current research. Esteban and Schneider (2008) also remark on the apparent overstretch of this concept and have devised a proper measure for what polarization entails. Loosely speaking, polarization is 'the extent to which the population is clustered around a small group of distant poles' and is most suitable for the analysis of conflict. Esteban and Schneider's (2008) explanation for this is that tensions in society between groups of individuals comes from 1) self-identification with similar others in a group and 2) distancing oneself from other competing groups. To find out how these poles may be detected, a promising direction to take is to focus on cues exhibited by these poles, especially when it is unknown who and how large these poles are. To gain insight on the selection of these cues it is best to look at the processes behind polarization in order to explain the concept of polarization further.

McDoom (2012) summarizes four indicators of group polarization; the first is the framing of an issue in terms of social identity, for example ethnicity. A second indicator is out-group derogation; the third indicator of group polarization is out-group homogenization, which is the unwillingness to distinguish individual members from their group membership. Esteban and Schneider (2008) further add that polarization will increase the more homogenous each group is and the higher the inter-group heterogeneity is. The final indicator is in-group solidarity, which is roughly translated as the loyalty individual members have towards the group they are a member of. The indicators McDoom (2012) mentions are well known concepts within social identity theory (Tajfel, 1970; Tajfel & Turner, 1979) and will be explained further.

2.1.1 Categorization processes.

In order to perceive, understand and make sense of others quickly, people classify others into groups based on age, race, ethnicity etc. This is also called the process of social categorization (Turner, 1991; 1999) and is one of the key components of social identity theory. Once classified, inferences are drawn about characteristics or stereotypes about these groups. For example, Bob is walking on the street; categorization processes quickly identify him as 'man', 30s, 'Dutch' and 'white'. The beliefs people then have about 'Bob' is then influenced by what people belief is typical for a 30's Dutch white man. These are stereotypes and do not necessarily have to be true. These beliefs also contain information about what makes Bob different from other individuals or groups, also dubbed the meta-contrast principle. In the same way Bob is categorized, people also categorize themselves in identity-relevant categories. In a lifetime, individuals often categorize themselves in multiple groups that they feel are relevant for them (Hogg, 2005), such as ones taste in art (Tajfel, 1970).

The attributes people ascribe to themselves also contain information about what groups this person may be a part of and how this differs from other people and other groups. Bob might think of himself as a Democrat, seeing anyone who shares Democrat views as members of the Democrat in-group. People who then do not share similar views are seen as non-inclusive of the in-group and are so called the out-group. In-group identification is then the acceptance of the group as an extension of the self while basing one's identity on the group's qualities and characteristics.

People generally view their own in-group and its members as more favorably than other groups and members, this is also called in-group favoritism. Oakes (1987) and Gaertner et al. (1993) show that the more salient a group affiliation is, the more biased the individuals' beliefs about the in-group and out-group. This saliency can be both context dependent (e.g. characteristics of the environment contain information about identity saliency) or more internalized, such as the loyalty portrayed by an individual group member. For example, during the world championship football we may all root for our own national team, so the shared social identity is then our nationality. When the world championship is over however, individuals may go back to rooting for their own regional football team (e.g. Ajax, Liverpool, Manchester United etc.). A side effect of this ingroup-outgroup bias is that sometimes those who do not belong to the in-group are derogated.

This ‘out-group derogation’ refers to the tendency to view out-groups as more negative. In the detection of polarization, focusing on the extent extreme members view each other as a disliked out-group is a fitting approach in identifying polarization (Iyengar, Sood & Lelkes, 2012). The traditional view by (Allport, 1954) is that in-group favoritism and out-group derogation are reciprocally related, experiments and cross-cultural research however support an alternative view. Brewer (1979) posits that the intergroup bias is motivated by a preferential treatment towards the in-group, rather than hostility towards an out-group.

Both in-group favoritism and out-group derogation are more likely to occur when the in-groups’ worth is questioned, in-group members will then zealously defend the worth of their group and underscore the distinctiveness of their group by increasingly derogating others (Brown & Hewstone, 2005; Dietz-Uhler & Murrell, 1998).[3] Fischer, Haslam, & Smith (2010) also call this ‘retaliation’ and could be considered as a superlative to out-group derogation.

2.1.2 Conflict

As attitudes become more extreme the need to retaliate becomes stronger, sometimes emotions run so high that conflict between two groups is the outcome. Examples of these emotions are feelings of fear, anger, humiliation, embarrassment and frustration are often potent instigators of intergroup conflict and hostility (Brown & Dutton, 1995). The need to retaliate is especially high when a collective identity is threatened (Fischer, Haslam and Smith, 2010), or when two groups are competing over the same scarce resource (Cuddy et al., 2009).

One of the consequences of being involved in conflicts that have spiralled out of control is that in-group members may believe themselves to be morally superior to the out-group, viewing them as even less than human (Bandura, 1999; Reicher, Haslam & Rath, 2008). This dehumanization allows the use of violence to be disinhibited and makes people feel less responsible for any violent action against the out-group as they are seen as a lesser man (Bandura, Underwood, & Fromson, 1975). Referring to each other in animalistic terms is one example of dehumanization (e.g. rat, cockroach, leech, parasite) but other forms of insult that relate to the same processes of dehumanization is calling homosexual men “faggots” (Fasoli et al., 2016).

[3] Some researchers claim that members who are most extreme in the display out-group derogation are often low-status peripheral members of the group who experience a threat to their own self-esteem (Noel, Wann & Branscombe, 1995). This assumption however has not received consistent empirical support (see Brown & Dutton, 1995).

Another example of extreme intergroup conflict is the occurrence of moral exclusion, which allows in-group members to devalue the out-group so completely that the use of violence or other hostile actions is rationalized and the out-group is excluded from any moral concern (Opotow & Weiss, 2000.) Examples of often morally excluded people include the Roma in Romanian society (Tileagă, 2007), but might also happen with homeless people for example.

Summarized, moral exclusion places outsiders outside the moral realm, while dehumanization places them outside the human realm and both are extreme examples of out-group derogations. Both processes increase the likelihood that hostile or aggressive actions occur. When conflict eventually erupts, groups tend to follow the norm of reciprocity; threats are answered with threats, insults with insults and aggression with aggression (Chen, Chen, & Portnoy, 2009). Conflict intensification often follows this upward spiral, starting with mild annoyances, followed by exclusion, then verbal abuse to discrimination and finally, physical assault (Streufert & Streufert, 1986).

2.1.3 Attitude

While most researchers focus on the division between parties on certain policy issues in order to study polarization, a better alternative may be to assess whether partisans view each other as a disliked out-group (Iyengar, Sood and Lelkes; 2012; Iyengar & Westwood, 2015; Mason, 2015). For example, research in American politics have shown that citizens' initial policy preferences have not changed much, but Democrats and Republicans have become increasingly hostile towards each other (Iyengar, Sood, & Lelkes, 2012). Furthermore, partisan identification was found to be an increasingly strong correlate of policy attitudes and behavior (Abramowitz and Saunders, 1998; Bartels, 2000; Hetherington, 2001; Bafumi and Shapiro, 2009), which is in line with the approach to social identity theory (Tajfel, 1970; Tajfel & Turner, 1979). This partisan identification is then sufficient to evoke negative sentiments about the out-group and is most strong around the "activists", also known as the extreme identifiers (Fiorina, Abrams & Pope, 2008). Results of Iyengar, Sood and Lelkes (2012) confirm that these extreme identifiers are increasingly more negative in terms of in-group and out-group ratings.[4] Examples of identity markers among which polarization

[4] Iyengar, Sood and Lelkes (2012) point out that out-group hostility is lower at the individual level rather than the group level – which is in line with the *person positivity bias* (Sears, 1983) – where individuals are more favorably disposed towards the target at the individual level rather than the abstract level.

can take place include party identification, as is the case in America (Iyengar, Sood & Lelkes, 2012), the politically engaged versus unengaged (Abramowitz, 2010) and the religious versus secular (Baker, 2005). And so a key role in detecting polarization is the attitudes members display in a group setting.

Boninger et al. (1995) found that the most passionately held attitudes concern issues that (1) directly affect self-interest, (2) are related to deeply held philosophical, political, and religious values and (3) are of concern to close friends, family or social in-groups. When people are surrounded by like-minded others, attitudes are held on especially strong and are even more resistant to change (Visser & Mirabile, 2004). People become more confident about their own beliefs when their attitudes are successfully defended against a persuasive message (Tormala & Petty; 2002). Furthermore, van Swol et al. (2016) note that those with an extreme attitude also perceive themselves as more knowledgeable.

Summarized, polarization then is the result of an increase in feelings of "us against them", with an increase in framing in terms of social identities. In-group favoritism and out-group derogation is the result and consequentially leads to stereotyping as positive traits are ascribed to in-group members, while negative ones to out-group members. Feelings of animosity then increase between opposing group members, and the end result is often a conflict where attitudes grow more extreme. Detecting those with an extreme attitude therefore seems like an appropriate path to take to ascertain whether polarization is present within a certain debate, one of the ways to do so is to study the language used within polarized debate.

2.2 Language

Language is the most common and reliable way for people to translate their internal thoughts and emotions into a form others can understand (Tauczik & Pennebaker, 2010) and so serves as an indicator in the detection of polarization. First we explore the persuasive effect of language found within polarized debate, after which the focus will shift towards an exploration of linguistic indicators of polarization.

2.2.1 The persuasive effect of polarized language

After the terrorist attacks of 9/11, the rhetoric of former U.S. president George W. Bush changed – as his language became more aggressive, approval ratings skyrocketed and

Bush was perceived to be more charismatic (Bligh, Kohles, & Meindl, 2004). It may then be possible that the use of more extreme language has a persuasive effect. Extreme language often carries an emotional component as well as a linguistic component. Craig and Blankenship (2011) call the linguistic component 'linguistic extremity' and emotional component 'linguistic intensity'. Examples of 'extreme language' include the use of words like "much more", "extremely", "very" and "wonderful" within a sentence. Research by Craig and Blankenship (2011) corroborate that using extreme language used in a topic associated with valence (positive and/or negative) leads to increased persuasion, the effect is mediated however by message processing. An explanation for this effect is that extreme language increases involvement in the message (Petty & Cacioppo, 1986). Craig and Blankenship (2011) also researched the effect of extreme language in combination with argument quality in two different studies. Their overall results suggest that weak arguments were perceived as weaker when combined with extreme language and strong arguments as stronger when paired with extreme language. Extreme language is then a linguistic indicator that has some persuasive effect.

It is possible however that pre-message attitudes towards the topic presented can moderate some of the persuasive effects of extreme language, especially when the persuasive message is discrepant from the attitude the recipient already holds. Craig and Blankenship (2011) hypothesize that the less favorable an attitude is towards a certain topic, the more diminishing the effect of linguistic extremity on persuasion.

2.2.2 Emotionality

Garcia et al. (2016) found that in terms of emotional content in online discussions, the valence of language that is used changes according to the polarity of a thread. Threads with high emotional content lead to higher arousal than threads with neutral content. Research shows that when an internet user perceives emotional content, arousal increases, leading to higher chances of participating in the discussion. Participation in the discussion then induces an instant decrease in arousal, combined with internal relaxation of arousal – as the human body makes sure arousal levels gravitate towards a homeostatic level – decreases the probability of further participation. Additionally Garcia et al. (2016) found that if other users in the online discussion would post more emotional content (again), arousal would also increase again; eventually leading to what Garcia et al. (2016) suggests collective behavior in online communities. Anti-Social Behaviour for example is

often related to certain emotions that include anger, frustration, arrogance, shame, anxiety, depression, sadness, low levels of fear, and lack of guilt (Cohen, 2005). Following this line of thought, and previously mentioned research on out-group derogation; polarized debate will probably include more emotional content words. Examples of such words are "angry", "mad", "sad", "shame", "anxious" etc.

Kramer, Guillory and Hancock (2014) also found evidence that suggests that reading more negative expressions about other people online, also may evoke more negative emotions outside of the online environment. This assumption about emotional contagion also holds true within a social network environment (Kramer, Guillory, & Hancock, 2014).

2.2.3 Temporal orientation

Godin (1999) argues that people infer on irrational grounds coming from past experience, expectations, status quo biases and beliefs. Evidence then rarely counts as people reject what does not conform to these grounds, only the extremes are remembered which will eventually lead to further polarization. Godin (1999) also notes that debate often takes place about *anticipated unknown* consequences. This may for example mean the fear of a terrorist attack. These are would-be events and are therefore open to *subjective* evaluation. Godin (1999) notes that *facts* are often absent within such discourse, as these are generally absent in the case of future consequences. No previous study exists to the knowledge of this author considering references to temporal orientation in polarized debate. Examples of temporally oriented words are "past", "present", "future", "will", "before", "gonna".

2.2.4 First person and third person plurals

Kaati, Shrestha, Cohen and Lindquist (2016) found that the use of third person plurals ("they", "them", "their" etc.) is indicative of a preoccupation with an opposing group. Furthermore, frequent use of third person plurals serve to strengthen group identity and adherence to a group's cause. Additionally, the use of first person and third person plurals has been found to be a good predictor for extremism (Pennebaker & Chung, 2008), for example American Nazis or animal rights activists. Alternative media also tend to use more "they"-words (Kaati, Shrestha, Cohen, & Lindquist, 2016). An explanation for this can for example be an exposure towards political campaigns. These might heighten the saliency of partisan identity in all party identifiers. According to Baldassarri and Gelman (2008) then polarization

induces alignment along multiple lines of conflict and organizes individuals and groups around exclusive identities, interests are then crystalized into opposing groups. A strong social identity should therefore promote the use of third person plurals and is therefore expected to be present within polarized debate.

2.2.5 Uncivil language

Polarization also leads to increasing incivility between individuals and decreases their support for compromise (Lelkes, 2016). This incivility is defined as language that is clearly disrespectful to another person or group (Mutz & Reeves, 2005; Sobieraj & Berry, 2011). Within the police intranet it is expected that incivility online is especially associated with opinion extremity, which can be a clue for polarization. The goal of uncivil language is to lessen the persuasive appeal of an opponent's message (Herbst, 2010). Furthermore, incivility can be a genuine expression of antipathy towards the opposition and is in extension a by-product of social identification of ideologues or partisans that share the same ideals and views as them and view those that disagree as an out-group (Levendusky, 2013). This out-group is then due to social identity processes met with contempt and loathing. Incivility can take shape in many forms such as an infinite range of swear words, ridicule, ad hominem attacks and more extreme examples may come in the form of previously mentioned dehumanization (Bandura, 1999; Reicher, Haslam, & Rath, 2008). Considering the interaction of social identity and polarization it is expected that contempt in the form of uncivil language will increase when social identity cues of an out-group are present within online Police discourse.

2.2.6 "You"-referencing

The use of "you" pronouns, when used excessively, can indicate the blaming of others. "you"-referencing could also be indicative of distancing and differentiation between speaker and listener as the speaker tries to assert their dominance over the listener (Kane & Rink, 2015; Simmons, Gordon & Chambless, 2005). Additionally, excessive use of "you" pronouns also indicates criticism and confrontation (Simmons et al., 2005). It is likely then that individuals with an extreme attitude use more "you" pronouns compared to those with a more moderate view on things. Pronoun words have been shown to reflect emotional states and features of social relationships (Chung & Pennebaker, 2007). These types of words are also more difficult to control as they are used without much awareness (Pennebaker, 2011),

pronouns also account for about 14% of word usage by people who are speaking and is therefore the most commonly used function word (Pennebaker, Chung, Ireland, Gonzales, & Booth, 2017). Furthermore, they indicate a relation between the speaker and the listener and could be indicative of attentional focus and status (Brown & Levinson, 1987); Kacewicz, Pennebaker, Davis, Jeon & Graesser, 2014; Van Swol & Carlson, 2015). When partisans use of "you" pronouns increase, this can be indicative of a more negative environment – which in turn leads to less information sharing and less complex language within the group (Van Swol et al., 2016), this is also called integrative complexity.

2.2.7 Integrative complexity

Individuals with extreme opinions may have more non-negotiable convictions that are less grounded on reasoned consequences and have more arguments based on morality (van Swol et al. 2016). Within research on extreme attitudes this is called the degree of integrative complexity and refers to the degree to which individuals consider and recognize multiple perspectives. This complexity can either derive from what Conway et al. (2011) calls *Elaborative complexity*, or *Dialectical complexity.*

2.2.7.1 Elaborative complexity. Elaborative complexity is the defence of a singular point of view in a complex manner by relating to different dimensions. Conway et al. (2011) uses the following example: "Peanut butter is delicious; it is creamy and sweet and makes for a healthy meal". The sentence never questions how great peanut butter is, but rather explains why it is great from multiple perspectives (both taste- and health related).

2.2.7.2 Dialectical complexity. The other branch of integrative complexity refers to Dialectical complexity, and derives from viewing a certain subject from multiple points of view. In the previous statement about peanut butter, the perspective is singular. Considering however the following statement: ''Peanut butter is delicious and creamy, but I do not like how it gets stuck on the roof of my mouth." In this sentence, two opposing views are given, one negative, the other positive – viewing a subject from multiple perspectives is therefore also a sign of integrative complexity (Conway et al., 2011). Those with an extreme opinion tend to display a lack of this integrative complexity (Tetlock, 1984, 2005).

2.2.7.3 Linguistic indicators of integrative complexity. When translated to language use, a couple of categories reflect depth and complexity of thought. These are cognitive mechanism words and are indicative of more complex word use and rational argumentation (Gelfand et al., 2015; Pennebaker, Booth, & Francis, 2007; Pennebaker, Booth & Francis, 2007; Pennebaker & King, 1999; Pennebaker, Mehl, & Niederhoffer, 2003; Tausczik & Pennebaker, 2010). Words that reflect integrative complexity are exclusion words ("but", "without", "exclude"), conjunction words ("and", "but", "whereas"), prepositions ("to", "with", "above"), tentative words ("maybe", "perhaps"), negations ("never", "without", "cannot"), discrepancies ("should", "would", "ought" and "wish"), insight words ("think", "know", "consider") cognitive mechanisms ("cause", "know", "ought"), more words per sentence and lower levels of inclusive words ("with, and") (Slatcher, Chung, Pennebaker, & Stone, 2007, p.67; Tausczik & Pennebaker, 2010, p.35).

In closing remark, the author of this paper acknowledges that more linguistic indicators may exist that may help in the detection of polarization, most of these focus on the detection of extremists and/or terrorists. While some overlap exists and arguments can be made that this is a branch of polarization, it is deliberately left outside the scope of this paper. We cannot be sure that more extreme rhetoric follows the same set of rules as more common debate that has been polarized.

2.3 Measuring Language

Analyzing what a person says is no different online than in the real world, however with the help of computers and statistical programs it is possible to do these types of analyses for us. Within linguistics, researchers often choose to classify emotions based on word list-basis, by using resources such as lists of emotion-bearing words, lexicons or affective dictionaries (Elliot, 1992; Ortony, Clore, & Foss, 1987). Measuring language by using Linguistic Inquiry and Word Count (LIWC) is an example of the former, and can be used to analyze certain keywords and the relations between them in order to infer whether clues of polarization are present or not. LIWC works by counting and coding each specific word that is mentioned into a meaningful psychological category and has been successfully applied in a variety of research settings, such as depression (Stirman & Pennebaker, 2001; Rude et al., 2004), deception (Newman, Pennebaker, Berry, & Richards, 2003), quality of relationships

(Simmons, Gordon & Chambless, 2005), thinking styles (Graesser, McNamara, Louwerse, & Cai, 2004) and many more.[5]

To find out whether we can find evidence for polarization within police discourse, using computerized text analysis programs is an ideal tool and should prevent the caveats of traditional manual coding, where raters first need to read transcripts and then need to manually tag phrases or words that represented the dimensions investigators were studying (Tauczik & Pennebaker, 2010). By analysing certain keywords and the relations between them it is possible to infer whether clues of polarization are present or not.

LIWC contains two central features, their processing component and the dictionaries. What LIWC is able to do is analyse a text file and then going through this file word by word, whereby a comparison is made to an attached dictionary file. Tauczik and Pennebaker (2010) exemplify this by referring to the word "it" within a certain sentence. This word is coded within the program as a function word, a pronoun and, more specifically an impersonal pronoun. The same can be done for the word "we" and "us" – third person plural pronouns, which are important concepts within social identity theory. These should tell something about the prevalence of social identity processes by counting the amount of references to such keywords. These outcomes can then be combined with the percentage of positive and negative emotion words within a text and the thinking styles employed by commentators (e.g. self-reflection, causal thinking). This way, important social identity processes like in-group favouritism and out-group derogation could be analysed and may provide clues on whether polarization is present or not.

2.4 Current research

2.4.1. Goals and expectations

The aim of this current paper is to come up with a set of linguistic indicators that will enable an automatic measurement of whether a debate has been polarized, this paper can be seen as a prelude to that and so the main objective is to test the theory mentioned above in practise. The main research question being here then is: has debate within the Dutch Police intranet polarized? If this is the case, the second research question would be: Which linguistic indicators are the most potent in the detection of polarization?

To answer this question, first, LIWC will be supplemented with any specific content

[5] See Tausczik & Pennebaker (2010) for a summary of LIWC application in published research studies.

words that fall under the mentioned linguistic categories. The following hypotheses have been constructed.

Hypothesis 1. Within polarized debate an increase is expected in the following linguistic indicators when compared against non-polarized debate:

a) first person plurals, b) third person plurals, c) you-referencing, d) uncivil discourse, e) negative emotional content words, f) temporal orientation words, of which future orientations will be expected more than past orientations. (Kaati, Shrestha, Cohen and Lindquist, 2016; Pennebaker & Chung, 2008; Garcia et al., 2016; Cohen, 2005; Godin, 1999; Lelkes, 2016; Bandura, 1999; Reicher, Haslam & Rath, 2008; Simmons et al., 2005).

Additionally, a decrease of complex language is expected that reflect a lack of integrative complexity as poles within polarized discourse will often express themselves in more singular points of view. Those with a more extreme attitude will often lack integrative complexity in their discourse (Tetlock, 1984, 2005).

Therefore the second hypothesis is:

Hypothesis 2. Within polarized debate a decrease is expected in the following linguistic indicators when compared against non-polarized debate:

a) cognitive mechanism words, b) insight words, c) tentative words, d) exclusion words, e) discrepancy words, f) conjunction words, g) prepositions h) negations and i) and number of words per sentence.[6] With the exception of j) inclusive words ("with, and") as higher display of these words reflect a lack of integrative complexity instead (Slatcher, Chung, Pennebaker & Stone, 2006, p. 67; Tausczik & Pennebaker, 2010, p.35).

Regarding attitudes deemed to be present within polarized debates it is expected that these are especially negative, due to the assumed presence of out-group derogation, conflict, incivility etc. The third hypothesis therefore is:

Hypothesis 3. Attitudes about and within polarized debates are significantly more negative than non-polarized debate.

[6] The linguistic indicators related to hypothesis 1 and 2 are summarized in table 1.

Regarding the potentially persuasive appeal of polarized debate it may be possible to isolate the words responsible for persuasion. However, this study takes a different approach for that instead. It may be best to let participants read a presumed to be polarized debate and test whether this affects pre-tested attitudes.

Hypothesis 4. It is expected that after reading a polarized discussion pre-tested attitudes grow more extreme. So a negative disposition may grow more negative and a positive disposition more positive.

Table 1

Linguistic categories related to polarization

Predicted effect for each linguistic indicator			
Linguistic category	English example	Dutch example	Expected effect
References to social identity	"Police, immigrants, Dutch"	"Politie, immigranten, Nederlanders"	+
Third person plural	"they, them, their"	"Zij, hen"	+
First person plurals	"we, us, our"	"Wij, on, hen"	+
"you"-referencing	"you, you're"	"Jij, je"	+
References to the past	"before, were, past"	"Voordat, was, vroeger"	+
References to the future	"Later, will, gonna"	"Later, zal, gaan"	+
Swear words	"Bastard, idiot"	"Klootzak, Idioot"	+
Dehumanization	"Rat, cockroach"	"Rat, kakkerlak"	+
Anxiety	"nervous, afraid, tense"	"Nerveus, bang, gespannen"	+
Anger	"hate, pissed, angry"	"Haat, boos"	+
Words per sentence	X	X	-
Prepositions	"to", above"	"Naar, op, boven"	-
Cognitive mechanism words	"cause, know, ought"	"oorzaak, weten, denken"	-
Insight	"think, know consider"	"Denken, weten, overwegen"	-
Discrepancy	"should, would, could"	"Zou, kunnen"	-
Tentative	"maybe, perhaps, guess"	"Misschien, wellicht, raden"	-
Conjunction	"and, also"	"En, ook"	-
Exclusion	"but, without, exclude"	"maar, zonder, exclusief"	-
Inclusion	"with, and"	"en, inbegrepen, ook"	+
Negations	"never, without, cannot"	"Nooit, zonder, niet"	-

2.4.2 Excluded categories

A few of the previously mentioned linguistic indicators have been exempted from further analysis however. These are the use of ridicule and ad hominem attacks. While very specific, these are context-dependent and so not possible to measure with LIWC. It is expected however that both will be reflected by an increase in "you-referencing" and negative

emotion words or incivility. Other categories that have been exempted are moral disengagement words (hate, fear, judge, criticize), references containing a violation of sacred values (attack on a religious belief), social isolation words (confine, abandon and withdraw) and violence (attack, fight, kill). With the exception of the 'violation of sacred values', all of these are in theory detectable with LIWC, these word categories however are specifically used in detecting rhetoric used by extremist individuals such as terrorists and therefore fall outside the scope of the current study. The use of words greater than six letters was also exempted from research as the results based on research using six-letter word indicators is based on English discourse and does not directly translate to Dutch due to the differences in word-length. It can therefore not be assumed that use of words greater than six letters is an indication of, in this case, integrative complexity.

Another category that is excluded is "linguistic extremity". Linguistic extremity is a tricky concept as this cannot be related to a shift in emotion; something called linguistic intensity (Craig & Blankenship, 2011). Separate studies would need to be done to isolate these words in the Dutch language as some words that are considered as linguistic extremity words in English may not be linguistic extremity words in Dutch (Swan & Smith, 2001).

3. Method

The current study was divided in three parts. Study 1 was the preliminary analysis of the subject matter collected from the police intranet. This includes both an investigation of the topic being discussed by police officers as well as what officers have commented regarding this topic. Study 2 was the administration of a survey to detect perceived polarization and persuasive elements which are presumed to be present within polarized debate. Study 3 represented the comparison of LIWC scores from the entire collected debate between conditions, ranked on perceived polarization scores following study 2. Together, these results should be indicative of whether debate within the police intranet has polarized and which indicators are best for detecting polarization.

Using the linguistic indicators in table 1 scores were compared between discourses gathered from the police intranet. It was expected that if debate within the police is polarized, the use of words as reflected in table 1 would increase as well. Specifically, an attempt was made to detect whether debate has polarized around two topics: 1) allowing the Muslim headscarf as part of the police uniform and 2) the celebration of the Dutch tradition of "Saint

Nicholas". Results were compared with discourse that was considered to be a control; these are gathered comments from police officers regarding the purchase of new service vehicles.

Hypothesis 5. It is expected that the control condition will score lower on the prevalence of linguistic indicators, as mentioned in table 1, compared to the other conditions.

In order to reliably detect whether the discussed debates were polarized, a survey was administrated among university students where participants are pretested about their opinion of the topic in question. After this pre-test participants have read a selection of the commentary from the police intranet, a polarization scale has been developed in order to ascertain whether participants perceived the debate as polarized. Additionally, participant opinions were analysed again to see what the impact of reading such discourse is. Following hypothesis 4 it was expected that pre-tested attitudes may polarize after reading a polarized debate.

Results were controlled by the degree of involvement and perceived argument quality. Scores of these surveys were compared against the LIWC scores from the commentary participants got to read. These scores should indicate the degree of polarization regarding the excerpts participants have read.

4. Study 1: Preliminary research

Study 1 is an additional literature review which was done to investigate whether the topics discussed could have polarized in Dutch society. The goal for this study was to better estimate which discussion could serve as a suitable control and which discussion could be polarized. Furthermore, the dictionary of LIWC was supplemented where necessary to reflect the linguistic indicators mentioned in table 1. To do this the source material gathered from the police intranet was investigated for words that were linked to the indicators in table 1 but were not present yet in the LIWC dictionary.

Commentary was collected from a total of three existing police intranet publications and discussions. Of these three, two threads were considered polarized and one was considered a control article. An explorative analysis was done to explain and provide clues why the selected articles were polarized.

Police officers were not aware their commentary was being used for research purposes, so anonymity is safeguarded by coding each text fragment from each police officer

in a non-retraceable manner. Only the researcher has a coding scheme of the authors of text fragments. Text fragments will not be quoted either. Removed commentary resulted in a notification that the user has deleted his/her comment; these are therefore excluded from the current study as well. Each article consisted of a news-like presentation of facts, accompanied by images portraying each situation. The length of these articles was approximately the size of one printed page on A4 format. Police officers were able to leave a comment directly below the article. Comments were placed in a chronological manner, with the most recent comment displayed first. Each comment also included information about the employee including their full name and employee number; these were made anonymous for ethical reasons. Publication date and time were also present.

The following case studies reveal why the selected articles in this study were considered polarized. The main presumption was that polarized debate had a salient social identity present and the control did not.

4.1 Case study – "Saint Nicholas"-protest

A yearly discussion repeats itself within the Netherlands regarding the celebration of 'Saint Nicholas' – a historical Dutch celebration where kids receive presents from 'Saint Nicholas' and his helper 'Black Pete'[7] – which has repeatedly been the target of anti-racism activists. Opponents suggest that the celebration of 'Saint Nicholas' glorifies Dutch slavery and has racist tendencies. Proponents on the other hand do not see the harm in this tradition and are increasingly agonized by the vocal response surrounding this tradition each year.

A reason why the debate surrounding the celebration of 'Saint Nicholas' may be polarized is that a collective identity may be threatened. Which according to Lettinga and Saharso (2014) may be an antecedent to polarization. The celebration of 'Saint Nicholas' (i.e. Santa Clause) is a historical Dutch celebration and could be considered as part of the Dutch national identity. Recently debates surrounding this tradition have flared up because of assumed racial characteristics of the mythical helpers of 'Saint Nicholas'. Opinions surrounding the historical Dutch celebration are often derogatory, racist and in some cases violent on both sides of the spectrum. It is telling that Pegida (Patriotic Europeans Against the Islamisation of the West) have also participated in protests (see NOS, 2017).

[7] 'Black Pete' is a character that is painted entirely black and helps 'Saint Nicholas' – an elderly white bishop – on his yearly naval journey to the Netherlands.

4.2 Case study – Framing of the headscarf

A verdict by the Netherlands Institute for Human Rights declared that a Muslim police officer should not be forbidden to wear a headscarf during work when the tasks are administrative in nature (Netherlands Institute for Human Rights, 2017). This has sparked discussion nationwide as well as within the police apparatus (AD, 2017). This discussion surrounding the headscarf was initiated by a Rotterdam police officer, because she as a Muslim officer was not allowed to wear a headscarf as part of her uniform. The debate surrounding the headscarf hints at deeper rooted Muslim discrimination and could constitute an infringement upon the freedom of religion principle within the Dutch constitution. Opponents for allowing the headscarf as part of the police uniform however state that having a neutral appearance as the government is of great importance. Both parties agree on this subject but have different perspectives on the matter. While one party claims that neutrality lies in the behaviour of the subject, the other party claims that it is the appearance of neutrality on its own that is more important. The debate surrounding the Muslim headscarf is considered controversial because of how the headscarf is framed within current discourse, as either discriminative or oppressive to female Muslims. As a result, the neutrality of the police is being questioned when they allow the Muslim headscarf, especially when police dress code prohibits the wearing of religious symbols.[8]

4.3 Control condition

On November 2017 the police shared on their intranet which service vehicles will be purchased for the next four years. Vehicles were chosen with the needs of the police force in mind. The canine brigade for example got different vehicles than those of the emergency aid. The article itself only consisted of specification of the vehicles, which concrete improvements they had over older service models and when officers can expect the newer vehicles to replace the old one. Considering the article only consisted of factual information and was about improvements made on vehicles and was not about people and/or negative consequences – it is assumed that this article was a suitable control.

4.4. Supplementing the LIWC dictionary – an explorative analysis

Initial explorative analysis is done to supplement the 2007 Dutch LIWC dictionary where necessary within the context of the topics surrounding the current study. All of the

[8] For a more in-depth analysis also see Lettinga & Saharso (2014).

collected narratives were manually checked for words that were considered noteworthy and based on the linguistic indicators mentioned table 1. The most noteworthy of these include Dutch swear words as well as references to religion and nationality (see appendix A for a full list of added words[9]). Names were deleted and coded as '[naam]'. A total of 227 words and three linguistic categories were added, these categories are 1) Dehumanizing words, 2) Social Identity words and 3) Conjunctions. These categories were added in table 1 as well. Inclusion of collected words therefore did not follow the superior procedure as mentioned in Boot, Zijlsta and Geenen (2017). They compiled the Dutch translation of the 2007 LIWC dictionary by correlating between LIWC scores from both English and Dutch versions of the same discourse. Collected discourse within this study however is only available in Dutch, following this procedure is therefore impossible. The resulting transcript is used for analysis within Linguistic Inquiry and Word Count (LIWC) software version 2015 and the Dutch LIWC 2007 dictionary. These are at the time of writing the most recent available versions. Current Dutch LIWC software has 68 linguistic categories for the classification of words for analysis. While the 2015 version (Pennebaker, Booth, Boyd, & Francis, 2015) includes over 70 categories, the most recent version of the Dutch translation stems from 2007. LIWC scores the percentage of words in each category that is present within the transcript.

[9] Appendix A also included words following an analysis of source material that was considered as an extreme narratives condition. These were texts written by convicted individuals. For methodological purposes it was decided that this condition does not reflect an extreme example of polarization and was thus removed from further analysis. The dictionary still includes words from this analysis.

5. Study 2: Perceived polarization and attitude persuasion

5.1 Method

5.1.1 Participants

A total of 108 participants have participated in a survey across three different conditions, a control condition (n=39), the Muslim headscarf condition (n=36) and the "Saint Nicholas" protest condition (n=33). Participants were randomly assigned across each condition. The only criteria for participation in the study were whether participants agreed to informed consent and were able to read Dutch. Participants consisted mostly of university students and have been approached through convenience sampling. A gift voucher worth € 15,00 was raffled among participants to promote participation in the study.

5.1.2 Procedure

Prior to the survey participants were told that there was an on-going research within the police regarding the communication of police officers online. Participants were asked whether they would like to give their opinion after reading one of three collected excerpts. Participants were told that these are actual comments of actual police officers. Participants first read some introductory remarks about one of three topics: the celebration of "Saint Nicholas", the Muslim headscarf and the purchase of new service vehicles for police officers. The latter being a control. Participants were then asked about their attitude and perceived knowledge regarding the presented topic. Secondly, participants were asked to read an excerpt of a police discussion. After reading this excerpt participants were asked to rate the discussion in terms of perceived polarization, argument quality and involvement in the discussion. Lastly, participants filled in the same attitude scale again in order to see if post-test attitudes may have been influenced. After the study participants were thanked, debriefed and told that the excerpts they have read consisted of opinions of only a small number of officers and is therefore not representative of how the Dutch police view any of the approached topics. All constructed scales within the survey exist of 9-point likert-scales in order to more accurately detect a shift in attitude for example.

5.1.2.1 Independent variables.

Written source material. Participants were asked to read one of three selected topics that have been discussed within the Dutch police intranet. A polarization scale has been

constructed to reveal which topic is assumed to be the most polarizing, this score will form the basis for further LIWC analyses.

Perceived objectivity. A 9-point likert-scale has been constructed to test the perceived objectivity of the debate. *"How objective do you think the commentary is?"* (with 1 = *very subjective* and 9 = *very objective*)

Perceived persuasiveness. A 9-point likert-scale has been constructed to test the perceived persuasive effect of used arguments within the debate. *"How persuasive do you think the suggested argument are?"* (with 1 = *not convincing at all*, and 9 = *very convincing*).

Both *perceived objectivity* and *perceived persuasiveness* have been constructed to test the effect these variables had on any change in post-test attitudes if these occurred.

5.1.2.2 Dependent variables.

Perceived polarization. To check whether participants perceived the debate as polarizing five 9-point likert-scales have been constructed that asked them to rate the degree of agreement between police officers with each other, group formation in two parties, hard contradicting statements within the discussion, and homogeneity of officers. "*To what degree do participants within this discussion agree with each other?"* (with 1 = *not at all* and 9 = *very*). Additionally two 9-point likert-scaled were constructed to rate the degree of friendliness/hostility within the debate. *"How friendly/hostile do you think this discussion is?"* (with 1 = *very hostile* and 9 = *very friendly*). A scale for the emotionality of the thread was also constructed. "*How emotional do you find the language that is used?"* (with 1 = *very businesslike* and 9 = very emotional). Finally participants were asked whether they thought the discussion was polarized. "*I find this discussion polarized"* (with 1 = *Strongly disagree* and 9 = *Strongly agree*).

Perceived friendliness/hostility of the discussion and perceived agreement among officers were first recoded as both were inversely scaled within the survey. Scale validation has not been done prior to this study as no previous polarization scale existed. Cronbach's alpha was .718 and Guttman's Lambda-2 was .733, these scores are deemed sufficient. The item related to perceived homogeneity of officers has been deleted from the scale, resulting in a total of 6 items. Summated scores reflect the degree of perceived polarization of the selected article.

Attitudes. After reading the introductory message, participants were asked to rate their attitude towards the three mentioned topics. After the survey attitudes were measured again to detect whether a shift has occurred. Attitudes were measured via four 9-point semantic differential scales. "*What is your opinion about allowing the headscarf as part of the police uniform?"* (with 1 = *harmful, foolish, unfavourable* and *undesirable* and 9 = *beneficial, wise, favourable* and *desirable*. Furthermore, participants were asked to rate how much they agreed with their given topic on a 9-point likert-scale. "*I think the headscarf should be allowed as part of the police uniform"* (1 = *Strongly disagree,* 9 = *Strongly agree).* Cronbach's alpha of pre-test attitude scores was .954 and Guttman's Lambda-2 was .955. For post-test attitude scores alpha was .956 and Guttman's Lambda-2 was .957 Difference scores have been calculated by subtracting the pre-test score from the post-test score, using this new variable the shift in attitude can be measured after reading the source material participants got to read, this way conclusions were made about persuasive influences.

Perceived knowledge. Participants needed to rate their perceived knowledge regarding the presented topic to see whether this had an effect on their attitude regarding both pre- and posttest attitude scores. Perceived knowledge was rated on a 9-point likert scale. "*How much knowledge do you think you have on this topic"* (with 1 = *Very little knowledge* and 9 = *A lot of knowledge*).

Involvement

Involvement should lead to increased message processing during the administration of this study and polarized individuals should also feel more involved in the topic they have read about. Participant involvement is measured with four 9-point likert-scales that asked participants to rate the degree of involvement, importance and interest in the topic. *"I feel involved in the discussion around the role of the headscarf within the police uniform"* (with 1 = *Not at all* and 9 = *very much*). Additionally participants were asked whether discussing this topic with another individual would have a lot of influence on them. Cronbach's alpha was .630 and Guttman's Lambda-2 was .649. See appendix B for all survey items.

5.2 Results

5.2.1 Perceived polarization

Descriptive statistics were first calculated for the *perceived polarization scale* for survey A: the headscarf debate (m = 5.9, sd = 1.2), for survey B: the 'Saint Nicholas' protest debate (m = 6.51, sd = 0.9) and for survey C, the control condition (m = 4.7, sd = 1.15). These results showed that participants rated the *protest* discussion to be the most polarizing, and the control to be the least polarizing.

Shapiro-Wilkinson test was used to test for normality as these are best for sample sizes smaller than 50, the results can be seen in table 3. It seemed that survey C (control) failed to reject the null hypothesis. For further analyses we do not assume normality; in the discussion section a short reflection will be given.

Table 2

Test of normality per condition

	Survey type	n	Shapiro-Wilk's W	*df*	*p*
Perceived polarization score	Headscarf	36	.979	35	.743
	Protest	33	.947	32	.119
	Control	39	.941	39	.041

As the design is unbalanced and mean polarization score is not normally distributed in one of the conditions, a non-parametric alternative to the ANOVA was performed. In this case the Kruskal-Wallis H test was used as differences were compared in three groups. For the dependent variable the *perceived polarization score* was inserted, grouping variable was *survey type.* Kruskal-Wallis first orders data from smallest to largest score by ranking these, it then compares across groups whether the mean of these ranks are the same. If a group has a lower mean rank it then scored lower on perceived polarization, as well as the other way around.

Kruskal-Wallis revealed that the degree of perceived polarization differs significantly between groups with $p < .001$. Further post-hoc pairwise comparisons showed that there is a significant difference between *control* and *headscarf* ($p < .001$), with *control* having a mean rank of 28.7 and *headscarf* a mean rank of 60.4. C*ontrol* and *protest* also differed significantly ($p < .001$), with *protest* having a mean rank of 76.08. No significant differences were found between *headscarf* and *protest* ($p = .112$).

It is therefore assumed that the debate within the control condition is not polarized; in contrast participants did view the headscarf and 'Saint Nicholas' protest debate as polarized. This confirms hypothesis 5 and makes it possible to rank conditions across their degree of perceived polarization.

5.2.2 Attitudes

First a correlational analysis was done to see whether pre-test attitude scores had a relationship with the perceived polarization score. See appendix B for the items used. Correlations revealed that pre-test attitude scores had no relationship with the previously constructed perceived polarization score. Then the normality of both pre-test and post-test attitudes was tested with a Shapiro-Wilkinson test to determine which tests were needed for further analyses. Pre-test attitude scores were not normally distributed ($p < .01$) so non-parametric tests were employed for further analyses instead.

Secondly, between-group differences for each condition were measured to detect whether attitudes changed significantly across surveys. Kruskal-Wallis was employed with pre-test and post-test attitude as dependent variables and survey type as grouping variable (see table 3). Having a negative attitude in this case means being against the headscarf as part of the police uniform (*headscarf* condition), being against change in the 'Saint Nicholas' tradition (*protest* condition) and against the purchase of new service vehicles (*control*). Having a positive attitude then means the opposite.

Table 3

Summary of differences on pre- and post-test attitudes between each condition.

	Control (n=37)	Headscarf (n=36)	Protest (n=32)			
Variables	Mean rank	Mean rank	Mean rank	Kruskal-Wallis H	*df*	*p*
Pre-test	70.7	48.2	37.9	21.4	2	.000
Post-test	70.7	43.4	44.5	18.6	2	.000

Additional post-hoc comparison with bonferroni correction revealed within-group differences across each condition.

Pre-test attitudes were significantly more negative in both *headscarf* ($H = -22.5, p < .001$) and *protest* ($H = -32.9, p < .001$) when compared against *control*. No differences were found between *headscarf* and *protest. Headscarf* had a mean rank of 48.2, *protest* a mean rank of 37.9 and control a mean rank of 70.7. Post-test attitudes were significantly more negative in

both *headscarf* (*H = -27.4, p <.001)* and *protest (H = -26.3, p <.01)* when compared against *control*. No differences were found between *headscarf* and *protest. Headscarf* had a mean rank of 43.4, *protest* a mean rank of 44.5 and control a mean rank of 70.7.

With attitudes being significantly lower in non-control conditions this confirms the third hypothesis.

With between-group differences known on attitude for each condition, the next step was to calculate whether pre- and post-test scores also changed significantly to measure whether reading polarized debate may have some persuasive influences. Wilcoxon signed-rank test is used as an alternative to the paired samples t-test with pre-test and post-test attitude scores as related samples. The datafile was split by survey type to reveal significant differences within each condition. To control for pre-test knowledge a correlation analysis was run with pre-test attitude scores. No significant relation was found (*p > .05).*

Wilcoxon signed-rank test revealed that attitude was significantly more negative in *headscarf* (*W* = 108,500; *p < .05)*. Pre-test attitude had a mean rank of 49.2 and post-test attitude had a mean rank of 43.4. Attitude was significantly more positive in *protest* (*W* = 209,500; *p < .05).* Pre-test attitude had a mean rank of 37.9 and post-test attitude had a mean rank of 44.5. Attitudes did not differ change in the control condition. These findings confirm the expectation that reading polarized discourse may have some persuasive influence. The next step was to calculate whether attitudes grew more extreme or were instead moderated. To do this, both pre-test and post-test attitude variables were split at the median of the neutral value of the 9-point likert-scale (5). Scores ranging from 5.1 to 9 were coded as positive with numerical value 1, while scores ranging from 1 to 4.9 were coded as negative with the numerical value -1. See figure 1 for a summary of the shift in attitudes for both pre-test and post-test attitudes.

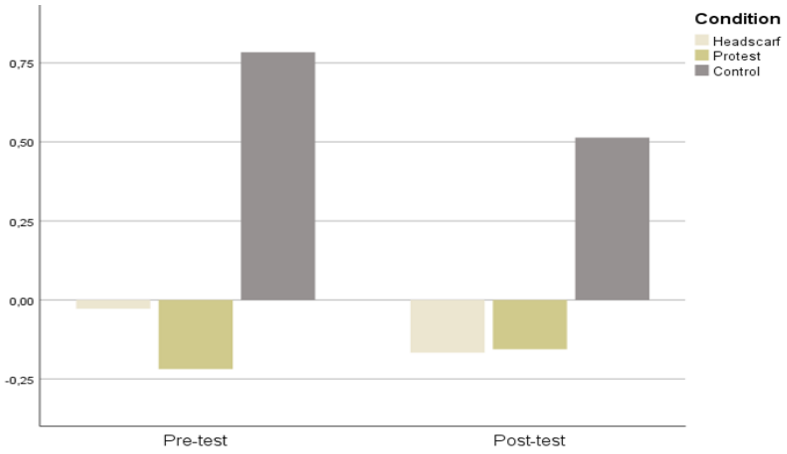

Figure 1 . Pre- and posttest attitude shift scores where values above 0 denote a positive attitude and values below zero denote a negative attitude.

To determine whether negative attitudes grew more negative, cases were filtered for both pre-test and post-test attitudes below the median attitude level. Cases for negative attitudes were selected if the condition for pre-test and post-test attitude scores <= 4.9 was fulfilled. The same procedure was employed for positive attitudes, except cases were selected above the 5.1 attitude score for both pre-test and post-test attitudes. All other cases were unselected. Data was then split across survey type condition. Then a Wilcoxon signed rank test was employed on the selected cases, filtered by survey type which gave the following results:
For *headscarf*, negative attitudes did not grow significantly more negative ($W = 29.0$, $p = .428$). For *protest*, negative attitudes grew more positive ($W = 69.5$, $p < .05$). No negative attitudes were found in the control condition so no cases were selected here.

A Wilcoxon signed rank test was then employed again on the selected cases, filtered by survey type which revealed that none of the positive attitudes grew more positive ($p > .05$)

These results lead to the rejection of hypothesis 4: attitudes do not grow more extreme after reading polarized debate. Instead it seems to have a (partially) moderating effect.

5.2.3 Predictor variables

For further analyses a new variable called *absolute attitude shift* (*AbsAttShift*) has been computed, which is the value of the post-test attitude score minus the pre-test and then the derived absolute value of this score. This variable reflects the true change of post-test attitude scores (both positive as negative). The more ones opinion changed, the higher the score. This variable is computed to calculate whether an opinion change is the result of either: *argument*

objectivity, *argument persuasiveness*, *involvement* or *perceived polarization score* (which reflects the degree of polarization of the thread). *Post-test attitude* was inserted as well to see which role this variable plays in relation with the other variables.

Normality was tested for each of these variables using Shapiro-Wilkingson. Both *involvement* and *perceived polarization* were not normally distributed (*p >.05)*. A correlational analysis was done to reveal if a relation was present between these variables using Spearman's R (see table 4).

Table 4

Summary of correlations for variables suspected to have a relation with either attitude shift or polarization scores

Variables		*1*	*2*	*3*	*4*	*5*	6	*7*
1.	Absolute Attitude shift	-						
2.	Argument objectivity	.076	-					
3.	Argument persuasiveness	.129	.491**	-				
4.	Involvement	.637	.324**	.258**	-			
5.	perceived polarization	.050	-.301**	-.322**	-.103	-		
6.	post-test attitudes	-.014	.010	-.089	.117	-.101	-	

Correlation is significant at the 0.01 level (2-tailed).**

Correlation is significant at the 0.05 level (2-tailed).*

Table 4 shows a significant negative relation between *perceived polarization* and *argument objectivity* and *argument persuasiveness.* There was also a positive relation between *involvement* and *argument objectivity* and *argument persuasiveness*. This might indicate that the arguments used by police officers moderate the perception of perceived polarization, especially when participants feel involved about the presented topic. No significant correlation was found on both the *post-test attitudes* and the *absolute attitude shift*.

Predictor variables for *perceived polarization* were further explored by calculating a linear regression, with *Absolute attitude shift* as dependent variable and all other variables from table 4 as independent variables. Cases were excluded pairwise to account better for missing data, otherwise data analysis discards data if any of the values within any of the variables are missing. Linearity assumption is relaxed by using Spearman's R instead of Pearson correlation.

A number of assumptions need to be met before interpreting linear regression. There first needs to be a correlation higher than .3, which for *perceived polarization* there is with *argument objectivity* and *argument persuasiveness*. As a general rule of thumb these

correlations should not correlate more than .7 between the independent variables to avoid multicollinearity. A Durbin-Watson test was employed to measure multicollinearity and gave a d = 1.559, while close it is still between the critical values of 1.5 and 2.5. It is assumed that no multicollinearity is present.

Furthermore, residuals should ideally range between -3 and 3, ensuring a low number of outliers. Figure 2 shows that no outliers were present for the linear regression on *perceived polarization*. Figure 2 also revealed that residuals do not fan out following the line. This shows the data also followed the assumption for linear regression of homoscedasticity.

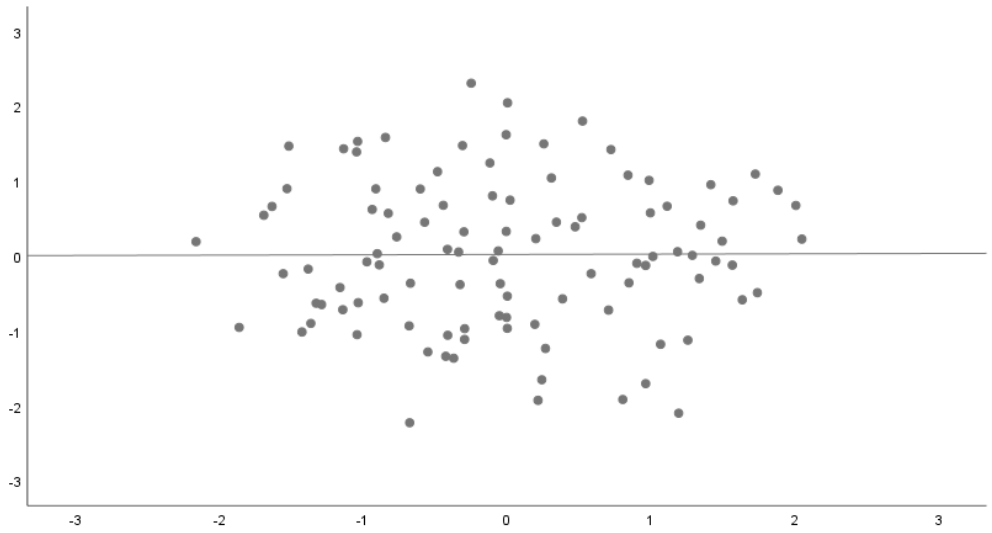

Figure 2. Scatterplot of residuals following a linear regression on *perceived polarization*

Residuals should also follow an approximate normal distribution, which is shown by a normal P-P plot in figure 3 as data points closely following the straight line.

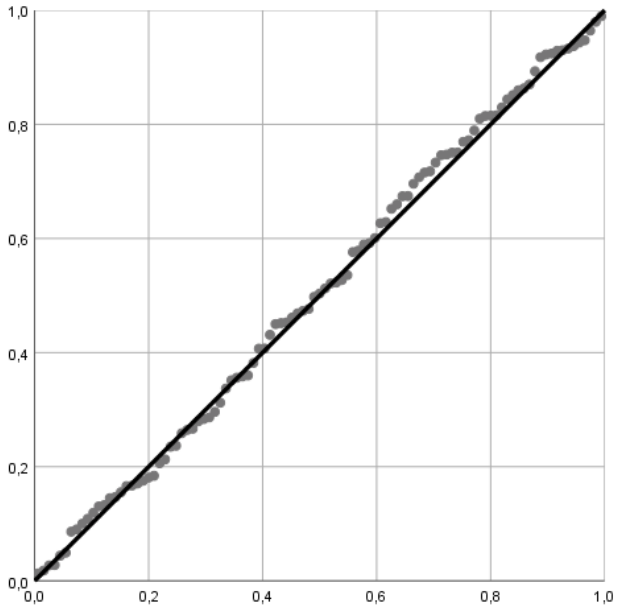

Figure 3. Normal P-P plot following a linear regression on *perceived polarization.*

As all assumptions were met, the actual linear regression can be done (see table 5).

Table 5

Linear regression model for *perceived polarization* score

Variables	*B*	*SE B*	β	*t*	*p*
Perceived polarization[a]	7.38	.523	-	14.1	.000
Involvement	0.04	.082	.049	0.5	.622
Argument objectivity	-.13	.076	-.184	-1.7	.091
Argument persuasiveness	-.21	.072	-.304	-2.8	.006
Post-test attitude	-.11	.062	-.171	-1.8	.070
Absolute attitude shift	.37	.205	-.167	1.8	.075

[a]Constant dependent variable

A simple linear regression was calculated to predict *perceived polarization* based on *involvement, argument objectivity, argument quality, post-test attitude scores* and *absolute attitude shift.* A significant regression equation was found (F (5,97) = 4.485, $p < .01$), with an R^2 of .188. Which means the sum of independent variables entered in the linear regression explain 18.8% of change in *perceived polarization* score, the dependent variable. However, table 5 reveals that only *Argument persuasiveness* contributed significantly in the linear regression model. So, participants predicted *perceived polarization* score is equal to 7.4 - 0.21 (*argument persuasiveness)*. This means participants' perception of the degree of polarization

on the given subjects decreases 0.21 for each unit increase on the 9-point likert-scale of *argument persuasiveness*. All other variables do not contribute significantly to the model.

5.3 Discussion

The goal of this study was to gain insight whether the discourse could be deemed as polarizing. It was expected that both *headscarf* and the *'Saint Nicholas' protest* were rated as more polarized than the *control*. Results are in line with hypothesis 5 with mean *perceived polarization score* significantly higher when compared against *control*. The *protest* condition scored highest on the polarization scale, which may indicate this debate polarized the most, however this score did not differ significantly from the *headscarf* debate. It is assumed both are polarized. A predictor variable for the polarization score was the degree of argument persuasiveness of police officers' arguments as the variable *argument persuasiveness* had a negative linear relationship with the *perceived polarization score.* This indicates that when participants rate the arguments of police officers as very persuasive, the perception of polarization decreases. A secondary analysis focused on the attitude participants had for each subject and whether these shifted after the survey. It could be possible that reading commentary about a polarized debate may lead to more extreme attitudes of the participant as well. Pre-test attitudes were more negative for both polarized debates when compared against control and remained negative after the survey. These findings confirm hypothesis 3. However, for *headscarf* these attitudes grew significantly more negative and for *protest* attitudes grew significantly less negative; indicating participants were less opposed to reform of the '*Saint Nicholas'* tradition. These effects hint at persuasive elements being present, further zooming on the direction of attitude change revealed that participants who had a negative pre-test attitude in the 'Saint Nicholas' *protest* condition grew significantly more positive. Hypothesis 4 is therefore rejected.

Overall it seems attitudes were moderated, instead of polarized. This moderating effect could not be explained by any present variables, including participant involvement. Craig and Blankenship (2011) suggested that pre-message attitudes could be responsible for the moderation; this study could not confirm that. The quality of police officers' arguments was also likely not responsible as no significant relation was found. A possible theory is that participants grew aware of polarization processes within the thread and therefore adjusted their own attitudes accordingly to not be seen as polarized as well – this theory however discounts the effect found on the persuasiveness of arguments used by police officers.

6. Study 3: Linguistic Analysis

To assess whether extreme attitudes were present within Dutch police commentary linguistic cues within the narratives of police officers were analysed. Analysis was done within discourse assumed to be more polarizing and compared against a control article. By comparing LIWC scores across each condition it may be possible to gain meaningful insights which linguistic indicators reveal more about the degree of polarization found within Dutch Police discourse.

6.1 Method

6.1.1 Participants

Commentary was provided by a Dutch police representative from a closed intranet environment from the Dutch national police in which every officer could partake. A total number of 398 comments were collected.

6.1.2 Procedure

Collected commentary was first ranked by the constructed *mean polarization score*, which was calculated in study 1. The control article (*control*) had a mean score of 4.65, the headscarf debate (*headscarf*) had a mean score of 5.90 and the protest debate (*protest*) had a mean score of 6.51.

Text fragments of police commentary were taken "as is", editing is done to correct for spelling mistakes. Correcting these mistakes was important as the goal of this study was to test the ideas mentioned previously. Excerpts were then analysed with LIWC. At present, LIWC only works with words that are spelled correctly. Practical implications are then a secondary goal and may be a next step for future research. Codes are used when names are mentioned. Output of LIWC is given as the percentage of total words within a linguistic category except for word count and words per sentence. All utterances are analysed via LIWC (2007, Dutch version with supplemented words from appendix A) and SPSS (version 25).

6.2 Results

To determine which tests are appropriate, normal distribution was first measured. Checking for skewness and kurtosis levels revealed that each linguistic indicator within collected discourse followed a non-normal distribution. Some words were barely used, while others were mentioned very frequently. Raw LIWC output was used in combination with non-parametric tests to analyse the dataset further.

6.2.1 Emotionality

Correlational analysis was done to reveal if emotionality had a relation with the polarity of a thread (see table 6). All emotional word variables were included in the correlational analysis, including positive emotion words and relevant subcategories. Means and standard deviations were excluded due to the presence of outliers.

Table 6

Summary of correlations for linguistic indicators on the emotionality of text compared against the mean perceived polarization score of police commentary

Variables	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
1. Mean polarization score[a]	-								
2. Affect	.189**	-							
3. Positive emotions	.120*	.742**	-						
4. Positive feelings	.061	.198**	.309**	-					
5. Optimism	.013	.399**	.506**	.121*	-				
6. Negative emotions	.126*	.548**	.006	.008	.137**	-			
7. Anxiety	.080	.142**	-.012	.001	.065	.260*	-		
8. Anger	.147**	.238**	.065	-.039	.090	.417**	.050	-	
9. Sadness	.043	.329**	-.015	-.015	.062	.606**	.115*	.121*	-

Correlation is significant at the 0.01 level (2-tailed).**

Correlation is significant at the 0.05 level (2-tailed).*

[a]Mean polarization score: *control* = 4.65, *headscarf* = 5.90 and *protest* = 6.51.

Mean polarization score correlated significantly with the following linguistic indicators: *Affect*, *Positive emotions*, *Negative emotions* and *Anger*. *Positive feelings* and *Optimism* were subcategories of *Positive emotions. Anxiety, Anger* and *Sadness* were subcategories of *Negative emotions* respectively. The highest correlation for a specific emotion was *Anger* (r_s = .147; $p < .001$). The more polarized the debate, the angrier participants got. An effect was also found on *positive emotion words*, further subcategories of *positive emotion words* revealed no further correlations however. A Kruskal-Wallis with emotional words as dependent variables and *mean polarization* score as grouping variable was done to reveal between-group significant differences (see table 7).

Table 7

Summary of between-group differences within police commentary ranked by mean polarization score on emotionality.

	Control (n=77)	Headscarf (n=249)	Protest (n=72)			
Variables	Mean rank	Mean rank	Mean rank	Kruskal-Wallis H	*df*	*p*
Affect	170	197	240	14.5	2	.001
Posemo	188	194	231	7.5	2	.023
Posfeel	197	198	208	2.1	2	.344
Optim	192	203	195	1.5	2	.476
Negemo	167	206	210	8.9	2	.012
Anx	192	201	203	3.0	2	.228
Anger	186	198	218	8.8	2	.012
Sad	185	205	196	3.5	2	.171

A post-hoc pairwise comparison with Bonferroni correction was done to reveal the within-group differences (see table 8).

Table 8

Summary of within-group differences of emotionality words between each condition ranked on mean polarization score.

Pairwise comparison [a]	*Kruskal-Wallis H*	*Std. error*	*p*
Affect			
1 – 2	-27.556	14.788	.187
1 – 3	-70.212	18.592	.000
2 – 3	-42.656	15.175	.015
Positive emotions			
1 – 2	-6.695	14.076	1.000
1 – 3	-43.071	17.697	.045
2 – 3	-36.375	14.445	.035
Negative emotions			
1 – 2	-39.027	13.712	.013
1 – 3	-42.170	17.240	.043
2 – 3	-3.143	14.071	1.000
Anger			
1 – 2	-12.203	8.485	.451
1 – 3	-31.347	10.667	.010
2 – 3	-19.144	8.707	.084

Significance values have been adjusted by the Bonferroni correction for multiple tests.

[a]Pairwise comparison ranked by mean polarization score: 1 = control, 2 = headscarf, 3 = protest.

Following table 7 and 8 it seemed that the higher the *mean polarization score* is, the more affective language was used: with a mean rank of 169.56 for the control, 197.11 for *headscarf* and 239.77 for *protest*. No significant difference was found between *control* & *headscarf* (*H = -27.556, p =.187).* As well as, the higher *mean polarization score,* the more negative language was used: with a mean rank of 167.45 for the control, 206.48 for *headscarf* and 209.62 for *protest*. No significant differences were found between *headscarf* and *protest*.

But also more words related to *anger* were used when *mean polarization score* was higher. with a mean rank of 186.19 for the control, 198.40 for *headscarf* and 217.54 for *protest*. No significant differences were found when comparing against the *headscarf* condition.

These findings are in line with hypothesis 1e, polarized debate include more negative emotional content words.

Positive emotion words stood out as well with *protest* having 230.59 as a mean rank score, while control has a mean rank score of 187.52 and *headscarf* has a mean rank score of 194.21. Effects found on *protest* here were likely the result of context of the analyzed debate and will be reflected upon in the discussion section. *Affec*t was excluded because this variable represents both positive as negative emotion words. *Positive emotions* was excluded as well and will also be reflected upon in the discussion.

6.2.2. Incivility

To test the degree of incivility in a thread, the number of swear words and dehumanisations were analysed and compared against the control. Analysing police commentary revealed no significant differences between groups for both *swear words* ($H = .408$, $p = .816$) and *dehumanizations* ($H = .598$, $p = .741$). Table 9 summarizes the descriptive statistics of incivility across each condition for *swear words* and *dehumanizations*.

Table 9

Descriptive statistics percentage scores of dehumanization and swear words per condition

Source		*N*	*Mean*	*Std. Deviation*
Control	swear words	77	0.02	.22
	dehumanizations	77	-	-
Headscarf	swear words	249	0.03	.23
	dehumanizations	249	0.01	.17
Protest	swear words	72	0.02	.12
	dehumanizations	72	-	-

No dehumanization took place in the *control* and *protest* condition. The mean percentage of swear words used within sentences found within police commentary was also very low.

These findings reject hypothesis 1d, uncivil discourse was not higher when compared against a control.

6.2.3 Social identity

To check whether social-identity processes were more salient in debates that scored higher on polarization, significant differences were checked on each linguistic variable related to that concept, those being *I, We, Self, You, Other, Social identity words, Occupation (Occup), Job* and *references to others (Othref).* A Kruskal-Wallis was run to reveal any between-group differences with social identity words as dependent variables and *mean polarization score* as grouping variable (table 10).

Table 10

Pairwise comparisons of social identity words between each condition

	Control (n=77)	Headscarf (n=249)	Protest (n=72)			
Variables	Mean rank	Mean rank	Mean rank	Kruskal-Wallis H	*df*	*p*
I	170	208	202	6.5	2	.039
We	183	202	208	4.1	2	.129
Self	166	210	201	9.0	2	.011
You	202	198	203	0.2	2	.891
Other	172	211	189	11.1	2	.004
Othref	160	206	218	12.0	2	.002
Occup	149	225	164	38.4	2	.000
Job	149	228	157	45.5	2	.000
SocID	151	218	187	29.1	2	.000

Significant differences were found on all of the variables except on the degree of we-referencing and you-referencing. A post-hoc pairwise comparison with Bonferroni correction was done to reveal within-group differences (table 10). Job was removed from table 10 due to a very high correlation with Occupation ($r_s = .968$; $p < .001$). It is likely the two measured the same construct. The same argument holds true for I and Self ($r_s = .895$, $p < .001$).

Table 11

Summary of within-group differences of social identity words between each condition ranked on mean polarization score.

Pairwise comparison [a]	*Kruskal-Wallis H*	*Std. error*	*p*
Self			
1 – 2	-43.9	-2.9	.008
1 – 3	-35.7	-1.9	.159
2 – 3	8.3	0.6	1.000
Other			
1 – 2	-39.6	12.5	.005
1 – 3	-16.9	15.7	.842
2 – 3	22.6	12.9	.234
Othref			
1 – 2	-46.1	14.8	.006
1 – 3	-57.7	18.6	.006
2 – 3	-11.5	15.2	1.000
Occup			
1 – 2	-76.4	14.1	.000
1 – 3	-15.4	17.8	1.000
2 – 3	61.0	14.5	.000
SocID			
1 – 2	-67.2	12.8	.000
1 – 3	-36.2	16.0	.071
2 – 3	30.9	13.1	.055

Significance values have been adjusted by the Bonferroni correction for multiple tests.

[a]Pairwise comparison ranked by mean polarization score: 1 = control, 2 = headscarf, 3 = protest.

Results of table 10 and 11 showed that when the mean polarization score increases, there is less self-referencing. Mean rank of *Self* in the control condition was 166 which was significantly lower than the other conditions ($p < .05$).

When mean polarization score increases the higher the amount of references to other people. Both *other* and *othref* giving higher scores in non-control conditions. However, *Other* did not produce a significant effect when comparing between *control* and *protest* (*p = .842).*

The findings from table 10 and 11 reject hypotheses 1a and 1c. No effects were found on we-referencing and you-referencing. Hypothesis 1b was partially confirmed, references to other persons were significantly higher in the *headscarf* condition.

6.2.4 Temporal orientation

To find out whether temporal orientation could be an indicator of polarization, significant differences were checked using a Kruskal-Wallis test across each condition with *time, past, present* and *future* as dependent variables and *mean polarization score* as grouping variable (see table 12).

Table 12

Pairwise comparisons of temporal orientation words between each condition

	Control (n=77)	Headscarf (n=249)	Protest (n=72)			
Variables	Mean rank	Mean rank	Mean rank	Kruskal-Wallis H	*df*	*p*
Time	213	198	189	2.6	3	.463
Past	187	199	214	3.2	3	.362
Present	168	215	179	31.5	3	.000
Future	183	210	182	11.6	3	.009

No significant differences were found for *Time* and *Past. Present* was significantly lower in the control condition with a mean rank score of 168. A post-hoc pairwise comparison with Bonferroni correction was done to reveal within-group differences. Post-hoc comparison showed that *present* was only significant when comparing against the *headscarf* condition (*H = -46.9, p <.01). Future* did not test significantly with a Bonferroni correction. These findings reject hypothesis 1f, temporal orientation words related to the past and future were not significantly higher in polarized debate.

6.2.5 Integrative complexity

A reliability analysis was done to reveal whether variables measured the same construct, this gave a Cronbach's alpha of .388 and Lambda-2 of .494. Deleting items would not ensure a greater alpha score. Exploratory factor analysis also revealed little as loading on one factor only explained 23 % of the variance. Significant differences were then analyzed between each condition using Kruskal-Wallis test with the following variables: *Words per sentence (WPS), Negations (negate), prepositions (preps), cognitive mechanism words (Cogmech), insight words (insight), discrepancies (discrep), tentative words (tentat), inclusion words (incl), exclusion words (excl)* and *conjunctions (conj).* See table 13 for a summary.

Table 13

Pairwise comparisons of integrative complexity words between each condition

	Control (n=77)	Headscarf (n=249)	Protest (n=72)			
Variables	Mean rank	Mean rank	Mean rank	Kruskal-Wallis H	*df*	*p*
Wps	193	205	186	1.8	2	.409
Negate	174	210	190	6.5	2	.038
Preps	213	200	184	2.3	2	.313
Cogmech	163	207	211	9.9	2	.007
Insight	150	213	205	18.6	2	.000
Discrep	196	206	182	2.7	2	.257
Tentat	173	211	189	7.6	2	.022
Incl	214	199	185	2.4	2	.296
Excl	179	207	195	3.6	2	.163
Conj	191	200	205	0.6	2	.722

Table 13 revealed significant differences between each condition on the following variables: *negations, cognitive mechanism words, insight words* and *tentative words*. A post-hoc pairwise comparison with Bonferroni correction was done to reveal the within-group differences (see table 14).

Table 14

Summary of within-group differences of integrative complexity words between each condition ranked on mean polarization score.

Pairwise comparison [a]	*Kruskal-Wallis H*	*Std. error*	*p*
Negations			
1 – 2	-35.9	14.7	.044
1 – 3	-16.1	18.5	1.000
2 – 3	19.8	15.1	.570
Cognitive mechanism words			
1 – 2	-44.9	14.9	.008
1 – 3	-48.7	18.8	.029
2 – 3	-3.8	15.4	1.000
Insight words			
1 – 2	-62.6	14.6	.009
1 – 3	-54.7	18.4	.000
2 – 3	7.9	14.9	1.000
Tentative words			
1 – 2	-37.4	14.4	.028
1 – 3	15.4	18.1	1.000
2 – 3	22.1	14.8	.405

Significance values have been adjusted by the Bonferroni correction for multiple tests.

[a]Pairwise comparison ranked by mean polarization score: 1 = control, 2 = headscarf, 3 = protest.

Results of table 13 and 14 showed that when *mean polarization score* increases more *negations* were used. The control condition scored significantly lower compared against the *headscarf* condition with a mean rank score of 174 ($p < .05$). There was no significant difference between *control* and *protest.*

When *mean polarization score* increases more *cognitive mechanism words* were also used. The control condition scored significantly lower compared against both other conditions with a mean rank score of 163 ($p < .05$). More *insight words* were used as well. The control condition scored significantly lower compared against both other conditions with a mean rank score of 150 ($p < .05$). Finally, more *tentative words* were used. The control condition scored significantly lower compared against the *headscarf* condition with a mean rank score of 173 ($p < .05$).

The findings reject hypothesis 2a, 2b and 2c and were opposite of what was expected. More cognitive mechanism words, insight words and tentative words were used in the

polarized debate. Hypothesis 2h was partially rejected, negations were higher in just one of the polarized conditions. Hypotheses 2d, 2e, 2f, 2g, 2i and 2j were also rejected as no significant differences were found between conditions on exclusion words, discrepancy words, conjunctions, prepositions, inclusions and words per sentence.

6.3 Discussion

Results showed that emotionality increases with the degree of polarization, for both positive and negative emotions, of which anger seemed to be most potent. Contrary to expectations an effect was found on positive emotion words for the 'Saint Nicholas' *protest* condition. This means words related to positive emotions were significantly higher in the *protest* condition. An explanation may be that LIWC detected and interpreted words related to the "celebration" of this Dutch tradition as a positive emotion word – resulting in a false-positive.

Further inspection of the valence of emotions found within police debate suggests that police officers used significantly more negative language and were significantly angrier when compared against the control condition. This suggests that especially negative emotions increases when polarization rises as well. Testing for incivility revealed little, as no effects were found on significant differences and correlational analyses.

Contrary to expectations about social-identity words no effect was found on we-referencing, which literature has suggested may be present within polarized debate (Pennebaker & Chung, 2008). Pennebaker and Chung (2008) however study extremism in this case, which current discourse may not be. Even so, Baldassarri and Gelman (2008) stated that polarization induces alignment along multiple lines of conflict and organizes individuals and groups around exclusive identities, interests should then crystalize into opposing groups. This would suggest an effect on we-referencing. It may be that the presence of "police-officer" as a more inclusive and binding social identity hampers this crystallization of subgroups with opposing points of view. This study did find an effect on third person plurals, which are the *Other* and *Othref* variables. According to research from Kaati, Shrestha, Cohen and Lindquist (2016) this could indicate a preoccupation with an opposing group.

The variables *Job* and *Occupation* were included because an individual can refer to his occupation as their social identity. These variables were therefore also included as part of the social identity construct and revealed significant effects as well in this *headscarf* condition. Content of the headscarf debate focused mostly on the police-uniform however, it may be that the context here led to false-positives instead. The same argument holds true for the newly

constructed *Social identity words (SocID)* variable. This variable included words which according to this researcher were related to social-identity salience and were manually added in the LIWC library following preliminary research from study 1. The *headscarf* debate contained the most words and so most words that were supplemented in LIWC also originate from this condition, this may have skewed the results for the *Social identity words* variable.

Regarding temporal orientation, Godin (1999) mentions that debate often takes place about anticipated unknown consequences, finding an effect on future references therefore was expected. An effect on past references was also expected as people often infer on irrational grounds from past experiences. Neither of the expected effects was found. An effect on present focus was found, as the control condition scored significantly lower than other conditions. At first sight this may then seem like a potential candidate for the detection of polarization, this author however suspects the variable is not necessarily related to polarization – as this variable has not been mentioned before in literature to the knowledge of this researcher.

Finally, for integrative complexity: effects were found on negations, cognitive mechanism words, insight words and tentative words. The ideal result would be that these variables scored significantly lower in non-control conditions. However, for all of these variables the opposite seemed to be true. An explanation for these unexpected results may be related to the complexity of the topic being discussed. The control condition for example was about the purchase of new service vehicles, while other conditions were about the discussion whether the headscarf should be admitted as part of the police uniform and how the police should handle the *protests* surrounding the 'Saint Nicholas' celebrations – thereby including their own personal opinions. The latter two subjects include many more dimensions than the control, including but not limited to religion, race, ethnicity, policy etc. This may have contributed to the increase in more complex arguments and language.

7. Overall conclusion and discussion

7.1 Research question and goal of the current study

For this study we were interested in two research questions: has debate within the police polarized, and if so, which linguistic variables could be a potent predictor for detecting polarization? To answer this question, this study had a two-pronged approach. Commentary from the police intranet was first collected to test whether polarization was present within

these studies when offset against a control. A survey was administered among non-police force participants, containing excerpts of comments from police officers on the collected articles. The most important item here was the polarization scale and was constructed to help estimate a ranking on the degree of polarization.

7.2 Main findings

Results of study 2 confirm that police commentary was perceived to be significantly more polarized when compared against a control, this partly answers our first research question. Police debate is perceived to be polarized. Study 2 also tested the attitudes participants had – which were significantly more negative in the polarized conditions, thereby confirming hypothesis 3. Participants' attitudes did however not polarize as a result of reading polarized debate, instead attitudes moderated. Hypothesis 4 was therefore rejected.
The perception of polarization was lower when police officers seemed to have convincing arguments, what the conditions are for an argument to be rated as convincing in this case remains elusive – though it is likely that persuasive elements of officers' commentary influenced the perception of polarization within the survey. Effects were small however.

Study 3 attempted to measure polarization through linguistic analysis, for this a number of linguistic indicators were chosen which research has suggested may be connected to the presence of polarization (see table 1). Summarized then an ideal finding would be that the following hypotheses were confirmed:

Hypothesis 1. Within polarized debate an increase is expected in the following linguistic indicators when compared against non-polarized debate: a) first person plurals, b) third person plurals, c) you-referencing, d) uncivil discourse, e) negative emotional content words, f) temporal orientation words.

Hypothesis 2. Within polarized debate a decrease is expected in the following linguistic indicators when compared against non-polarized debate: a) cognitive mechanism words, b) insight words, c) tentative words, d) exclusion words, e) discrepancy words, f) conjunction words, g) prepositions h) negations and i) and number of words per sentence.[10] With the exception of j) inclusive words ("with, and") as higher display of these words reflect a lack of integrative complexity instead.

[10] The linguistic indicators related to hypothesis 1 and 2 are summarized in table 1.

Findings from study 3 partially confirm hypothesis 1b, more references to others were made in the *headscarf* condition. Hypothesis 1e was also confirmed, polarized debate contained more negative affective language (thereby supporting hypothesis 3). Hypothesis 1a, 1c, 1d, 1f were rejected. No effects were found there.

Furthermore, results from study 3 led to the rejection of hypothesis 2d to 2j. No significant differences were found between conditions on exclusion words, discrepancy words, conjunctions, prepositions, inclusions and words per sentence. Hypothesis 2a to 2c did test significantly, just not what was expected and were rejected as well. Cognitive mechanism words, insight words and tentative words were contrary to expectations more prevalent in polarized debate. Hypothesis 2h was partially rejected, negations were higher in one of the polarized conditions. An explanation may be the complexity of the topic being discussed, where more complex topics in an online thread lead to more complex comments.

Summarized then a lot of the linguistic indicators chosen for study 3 were not appropriate for the detection of polarization. Focusing on references to others, in combination with negative affective language seems to be the most potent and reliable indicator to include when attempting to detect polarization.

7.3 Limitations

A caveat of using a survey was that participants have only read parts of the complete discussion; this was done to decrease survey fatigue. This researcher tried to include commentary which most accurately reflected the complete discussion but some form of experimenter expectancy bias might have been present in the choice of which comments to include for each condition. It would have been best to let participants read the debate in its entirety, but the danger for survey fatigue becomes very likely then. The current study tried to maintain a balance on this. A linguistic analysis (study 3) was believed to help compensate for this bias. Furthermore, one could argue that not the actual degree of polarization is measured but a perception of this concept. To control for this participants were first instructed what polarization is: "the exhibition of extreme attitudes where groups are clearly dichotomized around two extreme poles. Discussions are characterized by a lot of emotion and mutual conflicts." In essence participants were instructed to rate attitudes.

A problem with rating text-based attitudes however is whether the written word accurately reflects the subjective experience of the person writing it and if this was inferred correctly by

the participant in study 2. The expression of attitudes is therefore never a direct unbiased reflection of the inner self. Calvo and D'Mello (2010) however point out that other measures of detecting affective language suffer from the same problem – that attitudes can be faked or are open to interpretation. It would be wrong to assume however that the "true" display of attitudes is something that only happens inside our minds. On the contrary, Harber and Cohen (2005) point out that human emotion is not only intrapersonal but also deeply interpersonal and that humans share the need to display these emotions as well. In conclusion then it is just as valid to rate attitudes on what is displayed between individuals. The same argument holds true for the validity of text-based affective measurements for LIWC.

Reflecting on the results of the perceived polarization scale some outliers were found on the control condition on the higher end of the scale. It may have been possible that the phrasing of the survey title and sections of the survey could have had some influence on that. Participants might have been led to believe the debate of the control condition they were about to read was polarizing in nature. Nonetheless significant differences were found between the control condition and the polarized debates, with participants rating both the *headscarf* as the *'Saint Nicholas'* protest debate significantly high.

Post-test debriefings also suggested the presence of extreme arousal among some participants as they wanted to continue the debate with the researcher. According to Garcia et al. (2016) arousal is likely the cause of the need to participate. No effects were found on items of involvement however which means that they did not differ between groups and conditions or were a significant predictor for other variables. The involvement scale had a questionable alpha however; it may be possible that this contributed to the lack of significance as items did not accurately convey participant involvement. Also note that surveys were administered physically and in close proximity of the researcher, it is assumed that has kept participants sufficiently motivated across all conditions – explaining why no difference was found between conditions, even while some participants remarked that the control condition was "boring" too read. Finding no significant effects on involvement could therefore be desirable as well.

Reflecting back on the second approach of text-based measurements using LIWC, many variables did not follow the expectations as mentioned in table 1 of the theoretical framework. One could argue however that the current study is explorative in nature, as an attempt was made to test all previously mentioned linguistic variables related to polarization on real-world

examples of text. Text which was related to debate that is considered polarizing in nature. Previous attempts of this kind of study were not made to the knowledge of this author. Therefore, the absence of certain effects was probable.

Promising predictors for detecting polarization seem to be related to the emotionality of the thread – with anger being the most potent variable in combinations with references to other people in the form of third person plurals. This seems like a logical direction to take as polarized debates often characterize themselves as a highly negative preoccupation with other groups of individuals. This approach however does not take into account the accurate divergence of attitudes that is so characteristic of polarization and is something future research could focus on.

7.4 Future research

Following research should quantify the valence of emotion words found within discourse across valence scores. Imagine for example the difference between "dislike" and "hate" or even "detest". Valence scores could quantify these as an increasingly negative emotion word. Currently LIWC would code these both as negative emotion words and anger – making it improbable to detect extremization of attitudes. This methodology has been employed by Schweighoffer (2018, p. 154) and would require a lexicon of Dutch emotion words that each had their valence rated.

7.5 Concluding remark

As polarization can have very adverse effects in society, matters should be taken to prevent and de-escalate any detected forms of it. Especially if present within the police force as this may affect their daily jobs – in turn leading to perceptions of partiality. Detecting polarizing events is therefore crucial to continue defending the legitimacy of the police and in extension the legitimacy of the national government. As we spend a lot of our daily lives online a promising direction to take is to see what people share and write online. Judgments on safety versus privacy need to be carefully made however. Using text-based measures one could automatically analyze texts for clues on polarization. Steps need to be taken to fine-tune the automatic detection of Dutch emotion words across valence ratings. Constructing a comprehensive Dutch affective word list is thus the next order of business.

8. References

Abramowitz, A.I. (2010). The disappearing center: Engaged citizens, polarization, and American democracy. New Haven, CT: Yale University Press.

Abramowitz, A.I. & Saunders, K.L. (1998). Ideological Realignment in the US Electorate. *Journal of Politics*, 60, 634–652.

Abrams, D., Wetherell, M., Cochrane, S., Hogg, M. A., & Turner, J. C. (1990). Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization. *British Journal of Social Psychology*, 29, 97–119.

Winterman, P. (2017, may 26). 'Korpschef: Geen hoofddoek bij uniform'. *Algemeen Dagblad.* Retrieved from: https://www.ad.nl/binnenland/korpschef-geen-hoofddoek-bij-uniform~a5351e64/

Allport, G. W. (1954). *The nature of prejudice*. Cambridge Mass.: Addison-Wesley Publishing Company.

Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2013). The "nasty effect": Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication*, 19, 373-387.

Bafumi, J., & Shapiro, R.Y., (2009). A New Partisan Voter. *Journal of Politics*, 71, 1–24.

Baker, W. E., (2005). America's Crisis of Values: Reality and Perception. Princeton, NJ: Princeton University Press.

Baldassarri, D., & Gelman, A. (2008). Partisans without Constraint: Political Polarization and Trends in American Public Opinion. *American Journal of Sociology*, 114, 408–446.

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. Personality and Social Psychology Review. [Special Issue on Evil and Violence], 3, 193-209.

Bandura, A., Underwood, B., & Fromson, M. E. (1975). Disinhibition of aggression through diffusion of responsibility and dehumanization of victims. *Journal of Research in Personality, 9*(4), 253–269. https://doi.org/10.1016/0092-6566(75)90001-X

Bartels, L. M., (2000). Partisanship and Voting Behavior, 1952–1996. *American Journal of Political Science*, 44, 35–50.

Bligh, M. C., Kohles, J. C., & Meindl, J. R. (2004). Charting the Language of Leadership: A Methodological Investigation of President Bush and the Crisis of 9/11. *Journal of Applied Psychology, 89*(3), 562–574. https://doi.org/10.1037/0021-9010.89.3.562

Boninger, D. S., Krosnick, J. A., Berent, M. K., & Fabrigar, L. R. (1995). The causes and consequences of attitude importance. In R. E. Petty & J. A. Krosnick (Eds.), Attitude Strength: Antecedents and Consequences (pp. 159–189). Hillsdale, NJ: Lawrence

Erlbaum Associates.

Boot, P., Zijlstra, H. & Geenen, R. (2017). The Dutch translation of the Linguistic Inguiry and Word Count (LIWC) 2007 dictionary. *Dutch Journal of Applied Linguistics*, 6(1).

Borge-Holthoefer, J., Magdy, W., Darwish, K., & Weber, I. (2014). Content and Network Dynamics Behind Egyptian Political Polarization on Twitter. https://doi.org/10.1145/2675133.2675163

Brewer, M. B. (1979). Ingroup bias in the minimal intergroup situation: A cognitive motivational analysis. *Psychological Bulletin*, 86, 307-324.

Brown, J. D., & Dutton, K. A. (1995). The Thrill of Victory, the Complexity of Defeat: Self-Esteem and People's Emotional Reactions to Success and Failure. *Journal of Personality and Social Psychology*, 68, 712-722. http://dx.doi.org/10.1037/0022-3514.68.4.712

Brown, P., & Levinson, S. C. (1987). *Politeness : some universals in language usage*. Cambridge University Press. Retrieved from https://books.google.nl/books/about/Politeness.html?id=OG7W8yA2XjcC&redir_esc=y

Brown, R., & Hewstone, M. (2005). An integrative theory of intergroup contact. *Advances in Experimental Social Psychology*, *37*, 255–343. https://doi.org/10.1016/S0065-2601(05)37005-5

Calvo, R. A., & D'Mello, S. (2010). Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *IEEE Transactions on Affective Computing*, *1*(1), 18–37. https://doi.org/10.1109/T-AFFC.2010.1

Çankaya, S. (2012). *De controle van marsmannetjes en ander schorriemorrie: het beslissingsproces tijdens proactief politiewerk*. Boom Lemma uitgevers.

Çankaya, S. (2017). 'Zorgvuldig omgaan met grondrechten blijkt ook in Nederland een probleem te zijn'. *De Volkskrant.* Retrieved from: https://www.volkskrant.nl/columns-opinie/zorgvuldig-omgaan-met-grondrechten-blijkt-ook-in-nederland-een-probleem-te-zijn~b3b3779e/

Chen, Y. R., Chen, X. P., & Portnoy, R. (2009). To whom do positive norm and negative norm of reciprocity apply? Effects of inequitable offer, relationship, and relational-self orientation. *Journal of Experimental Social Psychology*, *45*(1), 24–34. https://doi.org/10.1016/j.jesp.2008.07.024

Chung, C., & Pennebaker, J. (2007). The Psychological Functions of Function Words. In K. Fiedler (Ed.), Frontiers of social psychology. *Social communication*, 343-359. New York, NY, US: Psychology Press.

Cohen, L. J. (2005). Neurobiology of Antisociality. In C. Stough, *Neurobiology of Exceptionality*, 107-124. New York, USA: Kluver Academic/Plenum Publishers

Conway, L. G., Gornick, L. J., Houck, S., Towgood, K. H., & Conway, K. R. (2011). The hidden implications of radical group rhetoric: Integrative complexity and terrorism. *Dynamics of Asymmetric Conflict: Pathways toward Terrorism and Genocide*, *4*(2), 155–165. https://doi.org/10.1080/17467586.2011.627938

Craig, T. & Blankenship, K.L. (2011). Language and Persuasion: Linguistic Extremity Influences Message Processing and Behavioral Intentions. *Journal of Language and Social Psychology*, 30, 290-310. DOI: 10.1177/0261927X11407167.

Dietz-Uhler, B., & Murrell, A. (1998). Effects of social identity and threat on self-esteem and group attributions. *Group Dynamics: Theory, Research, and Practice*, *2*(1), 24–35. https://doi.org/10.1037/1089-2699.2.1.24

DiMaggio, P., Evans, J., & Bryson, B. (1996). Have American's Social Attitudes Become More Polarized? *American Journal of Sociology*, *102*(3), 690–755. https://doi.org/10.1086/230995

Elliot, C. D. (1992). The Affective Reasoner: a process model of emotions in a multi-agent system. Northwestern University doctoral thesis. Evanston, Illinois.

Esteban, J., & Schneider, G. (2008). Polarization and conflict: Theoretical and empirical issues. *Journal of Peace Research*, *45*(2), 131–141. https://doi.org/10.1177/0022343307087168

Fasoli, F., Paladino, M. P., Carnaghi, A., Jetten, J., Bastian, B., & Bain, P. G. (2016). Not "just words": Exposure to homophobic epithets leads to dehumanizing and physical distancing from gay men. *European Journal of Social Psychology*, *46*(2), 237–248. https://doi.org/10.1002/ejsp.2148

Fiorina, M. P., Abrams, S. A., & Pope, J. C. (2008). Polarization in the American public: Misconceptions and misreadings. *The Journal of Politics*, 70(2), 556–560. doi: 10.1017/S002238160808050X

Fischer, P., Haslam, S. A., & Smith, L. (2010). "If you wrong us, shall we not revenge?" Social identity salience moderates support for retaliation in response to collective threat. *Group Dynamics: Theory, Research, and Practice*, *14*(2), 143–150. https://doi.org/10.1037/a0017970

Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A., & Rust, M. C. (1993). The Common Ingroup Identity Model: Recategorization and the Reduction of Intergroup Bias. *European Review of Social Psychology*, *4*(1), 1–26. https://doi.org/10.1080/14792779343000004

Garcia, D., Kappas, A., Küster, D., Schweitzer, F., & Garcia, D. (2016). The dynamics of emotions in online interaction. *Royal Society Open Science*, 3. https://doi.org/10.1098/rsos.160059.

Gelfand, M. J., Severance, L., Lee, T., Bruss, C. B., Lun, J., Abdel-Latif, A.H. Moustafa Ahmed, S. (2015). Culture and getting to yes : The linguistic signature of creative agreements in the United States and Egypt. *Journal of Organizational Behavior*, *36*(7), 967–989. https://doi.org/10.1002/job.2026

Godin, B. (1999). Argument from Consequences and the Urge to Polarize. *Argumentation*, *13*(4), 347–365. https://doi.org/10.1023/A:1007815429173

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-metrix: analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36, 193-202.

Harber, K. D., & Cohen, D. J. (2005). The Emotional Broadcaster Theory of Social Sharing. *Journal of Language and Social Psychology*, *24*(4), 382–400. https://doi.org/10.1177/0261927X05281426

Herbst, S. (2010). *Rude democracy : civility and incivility in American politics*. Temple University Press. Retrieved from https://www.jstor.org/stable/j.ctt14bt9vz

Hetherington, M. J. (2001). Resurgent mass partisanship: The role of elite polarization. American Political Science Review, 95(03), 619–631. doi: 10.1017/S0003055401003045

Hogg, M.A. (2005), Uncertainty, Social Identity, and Ideology. *Advances in Group Processes*, 22, 203-229. https://doi.org/10.1016/S0882-6145(05)22008-8.

Isenberg, D. J. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology*, *50*(6), 1141–1151. https://doi.org/10.1037/0022-3514.50.6.1141

Iyengar, S. & Westwood, S. J. (2015). Fear and Loathing across Party Lines: New Evidence on Group Polarization. *American Journal of Political Science*, 59, 690–707.

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, *76*(3), 405–431. https://doi.org/10.1093/poq/nfs038

Iyengar, S., & Westwood, S. J. (2015). Fear and Loathing across Party Lines: New Evidence on Group Polarization. https://doi.org/10.1111/ajps.12152

Kaati, L., Shrestha, A., Cohen, K., & Lindquist, S. (2016). Automatic detection of xenophobic narratives: A case study on Swedish alternative media. *IEEE International Conference on Intelligence and Security Informatics: Cybersecurity and Big Data, ISI 2016*, 121–126. https://doi.org/10.1109/ISI.2016.7745454

Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014). Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology*, 33, 125-143.

Kane, A. A., & Rink, F. (2015). How newcomers influence group utilization of their knowledge: Integrating versus differentiating strategies. *Group Dynamics*, 19, 91-105.

Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Editorial Expression of Concern: Experimental evidence of massivescale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, *111*(29), 10779–10779. https://doi.org/10.1073/pnas.1412469111

Lamm. H., & Myers, D. G. (1978). Group-induced polarization of attitudes and behavior. In L. Berkowitz (Ed.). *Advances in experimental social psychology*, 2, 147-195. New York: Academic Press.

Lelkes, Y. (2016). The polls-review: Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*, *80*(1), 392–410. https://doi.org/10.1093/poq/nfw005

Lettinga, D., & Saharso, S. (2014). Outsiders Within: Framing and Regulation of Headscarves in France, Germany and The Netherlands. *Social Inclusion*, *2*(3), 029. https://doi.org/10.17645/si.v2i3.46

Levendusky, M. S. (2013). How partisan media polarizes America. Chicago, IL: University of Chicago Press.

Mason, L. (2015). "I disrespectfully agree": The differential effects of partisan sorting on social and issue polarization. *American Journal of Political Science*, 59(1), 128–145. doi: 10.1111/ajps.2015.59.issue-1

McDoom, O. S. (2012). The Psychology of Threat in Intergroup Conflict: Emotions, Rationality, and Opportunity in the Rwandan Genocide. *International Security*, *37*(2), 119–155. https://doi.org/10.1162/ISEC_a_00100

McGarty, C., Turner, J. C., Hogg, M. A., David, B., & Wetherell, W. S. (1992). Group polarization as conformity to the prototypical group member. *British Journal of Social Psychology*, 31, 1–20.

Mutz, D., & Reeves, B. (2005). The new videomalaise: Effects of televised incivility on political trust. *American Political Science Review*, 99(1), 1–15. doi: 10.1017/S0003055405051452

Netherlands Institute for Human Rights (2017, November 11). Verdict 2017-135. Retrieved from: https://www.mensenrechten.nl/nl/oordeel/5b3cda618de5272746ab2424

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying Words: Predicting Deception from Linguistic Styles. *Personality and Social Psychology Bulletin, 29*(5), 665–675. https://doi.org/10.1177/0146167203029005010

Noel, H. 2014. Polarization Is About More Than Just Sorting, But Sorting Is Polarization Anyway" Available at http://www.mischiefsoffaction.com/2014/06/polarization-is-about-more-than-just.html/.

'Hebben uiterst links en uiterst rechts de Pietendiscussie gekaapt?' (2017, december 2). *Nederlandse Omroep Stichting.* Retrieved from: https://nos.nl/nieuwsuur/artikel/2205683-hebben-uiterst-links-en-uiterst-rechts-de-pietendiscussie-gekaapt.html

Kouwenhoven, A., & Rueb, T. (2016, june 11). 'Zo praat de politie over Typhoon (onder elkaar)'. *NRC Handelsblad*. Retrieved from: https://www.nrc.nl/nieuws/2016/06/11/zo-praat-de-politie-over-typhoon-onder-elkaar-1627729-a1507052

Laarhoven, K. van., (2018, april 29). 'In Dokkum had politie alleen oog voor Zwarte Piet-tegenstanders. *NRC Handelsblad.* Retrieved from: https://www.nrc.nl/nieuws/2018/04/29/de-hooligans-konden-gewoon-doorlopen-a1601268

Oakes, P. J. (1987). The salience of social categories, In J. C. Turner et al., Rediscovering the social Group: A Self-Categorization theory. Oxford/New York: Basil Blackwell.

Opotow, S., & Leah W. (2000). Denial and the Process of Moral Exclusion in Environmental Conflict. *Journal of Social Issues*, *56*(3), 475–490.

Ortony, A., Clore, G. L., & Foss, M. A. (1987). The Referential Structure of the Affective Lexicon. *Cognitive Science*, *11*(3), 341–364. https://doi.org/10.1207/s15516709cog1103_4

Pavitt, C. (1994). Another View of Group Polarizing: The "Reasons for" One-Sided Oral Argumentation. *Communication Research*, *21*(5), 625–642. https://doi.org/10.1177/009365094021005004

Pennebaker, J., Booth, R., & Francis, M. (2007). Linguistic Inquiry and Word Count (LIWC2007). Retrieved from https://researchspace.auckland.ac.nz/handle/2292/18109

Pennebaker, J. W. (2011). *The secret life of pronouns : what our words say about us*. Bloomsbury Press.

Pennebaker, J. W. & Chung, C. K. "Computerized text analysis of al- Qaeda transcripts," in The Content Analysis Reader, K. Krippendorf and M. A. Bock, Eds. Sage, 2008.

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, 77, 1296-1312.

Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (2015). Linguistic Inquiry and Word Count: LIWC2015 Operator's Manual. Retrieved from https://s3-us-west-2.amazonaws.com/downloads.liwc.net/LIWC2015_OperatorManual.pdf

Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2017). The Development and Psychometric Properties of LIWC2007. Retrieved from http://www.liwc.net/LIWC2007LanguageManual.pdf

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological Aspects of Natural Language Use: Our Words, Our Selves. *Annual Review of Psychology*, *54*(1), 547–577. https://doi.org/10.1146/annurev.psych.54.101601.145041

Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of Persuasion. *Communication and Persuasion*, 1–24. New York, NY: Springer New York. https://doi.org/10.1007/978-1-4612-4964-1_1

Reicher, S., Haslam, S. A., & Rath, R. (2008). Making a Virtue of Evil: A Five-Step Social Identity Model of the Development of Collective Hate. *Social and Personality Psychology Compass*, *2*(3), 1313–1344. https://doi.org/10.1111/j.1751-9004.2008.00113.x

Robison, J., & Mullinix, K. J. (2016). Elite Polarization and Public Opinion: How Polarization Is Communicated and Its Effects. *Political Communication*, *33*(2), 261–282. https://doi.org/10.1080/10584609.2015.1055526

Rude, S., Gortner, E. M., & Pennebaker, J. (2004). Language use of depressed and depression- vulnerable college students. *Cognition & Emotion*, 18, 1121-1133.

Schweighofer, S. (2018). Affective, Cognitive, and Social Identity Related Factors of Political Polarization. Doctoral thesis, Federal Institute of Technology, Zürich. Retrieved from: https://doi.org/10.3929/ethz-b-000256680

Siff, S. B. (2016). Policing the police: A civil rights story. Retrieved from http://origins.osu.edu/article/policing-police-civil-rights-story/page/0/1

Simmons, R., Gordon, P. C., & Chambless, D. (2005). Pronoun use in marital interaction: What do "you" and "I" say about marital health? *Psychological Science*, 16, 932-936.

Slatcher, R. B., Chung, C. K., Pennebaker, J. W., & Stone, L. D. (2007). Winning words: Individual differences in linguistic style among U.S. presidential and vice presidential candidates. *Journal of Research in Personality*, *41*(1), 63–75. https://doi.org/10.1016/j.jrp.2006.01.006

Sobieraj, S., & Berry, J. M. (2011). From incivility to outrage: Political discourse in blogs, talk radio, and cable news. *Political Communication*, 28, 19-41.

Stirman, S. W., & Pennebaker, J. W. (2001). Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic Medicine*, *63*(4), 517–522.

Streufert, S., & Streufert, S.C. (1986). The development of international conflict. In S. Worchel & W.G. Austin (Eds.), *Psychology of intergroup relations*, 2, 134-52. Chicago: Nelson-Hall.

Suhay, E. (2015). Explaining Group Influence: The Role of Identity and Emotion in Political Conformity and Polarization. *Political Behavior*, *37*(1), 221–251.

https://doi.org/10.1007/s11109-014-9269-1

Suhay, E., Blackwell, A., Roche, C., & Bruggeman, L. (2015). Forging Bonds and Burning Bridges: Polarization and Incivility in Blog Discussions About Occupy Wall Street. *American Politics Research*, *43*(4), 643–679. https://doi.org/10.1177/1532673X14553834

Swan, M., & Smith, B. (2001). *Learner English : a teacher's guide to interference and other problems*. Cambridge University Press.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223, 96-102.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In Austin, W. G. & Worchel, S. (Eds.). *The social psychology of intergroup relations*, 33–47. Monterey, CA: Brooks-Cole.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, *29*(1), 24–54. https://doi.org/10.1177/0261927X09351676

Tetlock, P. E. (1984). Cognitive style and political belief systems in the British House of Commons. *Journal of Personality and Social Psychology*, 46, 365-375.

Tetlock, P. E. (2005). Expert political judgment. Princeton, NJ: Princeton University Press.

Tileagă, C. (2007). Ideologies of moral exclusion: A critical discursive reframing of depersonalization, delegitimization and dehumanization. *British Journal of Social Psychology*, *46*(4), 717–737. https://doi.org/10.1348/014466607X186894

Tormala, Z. L., & Petty, R. E. (2002). What doesn't kill me makes me stronger: The effects of resisting persuasion on attitude certainty. *Journal of Personality and Social Psychology*, 83, 1298–1313.

Turner, J. C. (1991). Social influence. Milton-Keynes, UK: Open University Press.

Turner, J. C. (1999). Some current issues in research on social identity and self-categorization theories. In Ellemers, N. Spears, R. (Eds). Social identity: Context, commitment, content, 6–34. Oxford, UK: Blackwell.

Van der Leun, J. P., van der Woude, M. A. H., Vijverberg, R. D., Vrijhoef, R. P. M., & Leupen, A. J. (2014). Etnisch profileren in Den Haag? Een verkennend onderzoek naar beslissingen en opvattingen op straat. Retrieved from https://www.politieacademie.nl/kennisenonderzoek/kennis/mediatheek/PDF/89640.PDF

Van Swol, L. M., & Carlson, C. L. (2015). Language use and influence among minority, majority, and homogenous group members. *Communication Research*, 44, 51-529. doi:10.1177/0093650215570658

Van Swol, L. M., Prahl, A., Kolb, M. R., Lewis, E. A., & Carlson, C. (2016). The Language of Extremity: The Language of Extreme Members and How the Presence of Extremity

Affects Group Discussion. *Journal of Language and Social Psychology*, *35*(6), 603–627. https://doi.org/10.1177/0261927X16629788

Vinokur. A.. & Bumstein. E. (1974). Effects of partially-shared persuasive arguments on group-induced shifts: A group-problem-solving approach. *Journal of Personality and Social Psychology*, 29, 305-315.

Visser, P. S., & Mirabile, R. R. (2004). Attitudes in the Social Context: The Impact of Social Network Composition on Individual-Level Attitude Strength. *Journal of Personality and Social Psychology*, *87*(6), 779–795. https://doi.org/10.1037/0022-3514.87.6.779

'Amnesty: beëindigen anti-Zwarte Pietprotest in Rotterdam disproportioneel' (2016, november 15). *De Volkskrant.* Retrieved from: https://www.volkskrant.nl/nieuws-achtergrond/amnesty-beeindigen-anti-zwarte-pietprotest-in-rotterdam-disproportioneel~b04a6f59/

Appendix A – Supplemented words in the LIWC 2007 dictionary

Table 15

Summary of LIWC dictionary word categories

	English	*Dutch*		*English*	*Dutch*
1	Pronoun	Voornaamwoord	37	Time	Tijd
2	I	1e persoon enkelvoud	38	Past	werkwoorden vt
3	We	1e persoon meervoud	39	Present	werkwoorden tt
4	Self	Totaal eerste persoon	40	Future	werkwoorden tkt
5	You	Totaal tweede persoon	41	Space	Ruimte
6	Other	Totaal derde persoon	42	Up	omhoog
7	Negate	Ontkenningen	43	Down	beneden / neer
8	Assent	Toestemming/instemming/goedkeuring	44	Incl	Inclusief
9	Article	Lidwoorden	45	Excl	Exclusief
10	Preps	Voorzetsels	46	Motion	Beweging
11	Number	Getallen	47	Occup	Beroep/bezigheden
12	Affect	Emotionele processes	48	School	School
13	Posemo	Positieve emoties	49	Job	Werk
14	Posfeel	Positieve gevoelens	50	Achieve	Prestatie
15	Optim	Optimisme en energie	51	Leisure	Vrijetijdsbesteding
16	Negemo	negatieve emoties	52	Home	In en om het huis
17	Anx	angst en vrees	53	Sports	Sporten
18	Anger	boosheid	54	TV	Televisie en films
19	Sad	treurigheid of depressie	55	Music	Muziek
20	Cogmech	cognitieve processen	56	Money	Geld en Fin diensten
21	Cause	veroorzaken	57	Metaph	Godsdienst / zingeving
22	Insight	Inzicht	58	Relig	Religie
23	Discrep	Tegenstrijdigheid	59	Death	Dood en sterven
24	Inhib	Remming/geremdheid	60	Physcal	Lichamelijke toestand
25	Tentat	Tentatief/voorzichtig/aarzelend	61	Body	Lichamelijk functioneren
26	Certain	Zekerheid	62	Sexual	Seks en seksualiteit
27	Senses	Zintuigen en Perceptuele processen	63	Eating	Eten, drinken en dieet
28	See	Zien	64	Sleep	Slapen en dromen
29	Hear	Horen	65	Groom	Lichamelijke verzorging
30	Feel	Voelen	66	Swear	Vloekwoorden
31	Social	Sociale processen	67	Nonfl	Uhm / hmm / ehm
32	Comm	Communicatie	68	Fillers	Opvulwoorden / blabla / weetjewel
33	Othref	Verwijzingen naar andere personen	69	Dehum	Dehumaniseringen
34	Friends	Vrienden	70	SocID	Sociale Identiteit
35	Family	Familie	71	Conj	Conjuncties
36	Humans	Mensen			

Table 16

Added words to 2007 LIWC dictionary

bloedzuiger*	69	66			
parasiet*	69	66			
hond	69	66			
honden	69	66			
want	71				
doordat	71				
voordat	71				
als	71				
mits	71				
zodat	71				
opdat	71				
teneinde	71				
alsof	71				
zoals	71				
dan	71				
Zwijn	66	69			
Zwijnen	66	69			
Zwijnen*	66	69			
Zwarten	66	70			
zodra	37	71			
Zielig	12	16	19		
Zeiken	12	16	32	18	39
Xenofobe	12	16	17		
Xenofobie	12	16	17		
Xenofoob	12	16	17		
Witte	70				
Westers	57	70			
Westerse	57	70			
Wereldbeeld	57				
wanneer	37	71			
Vooringenomenheid	12	16			
Vliegend Spaghettimonster	58				
Verzuiling	57	58	31		
Vervelende	12	16	18		
Vergiet	57	58			
Verdraagzame	12	13			
Verdeeldheid	31	32			
Verbaasd	12	17			
verbaast	12	17			
Ventileren	12	16	32	18	39

Veilige	12	13		
varken	69	66		
varkens	69	66		
Turk	70			
Turken	70			
Turkije	70			
Turks	70			
Turkse	70			
Tulband	58			
totdat	37	71		
toen	37	71		
terwijl	37	71		
tenzij	45	71		
Staat en kerk	57	59		
slet	62	66		
sletje	62	66		
sinds	10	37	71	
Seculier	58	57		
Seculiere	58	57		
roetmop	66			
Respectloos	12	16	31	32
Respectloze	12	16	31	32
Ratten	69	66		
Ratten*	69	66		
Rat	69	66		
Provocatie	12	16		
Progressief	70			
Progressieve	70			
Professional	47	32	36	49
Professionals	47	32	36	49
Positieve	12	13		
Politie	70	47	49	
Politie-agent	70	47	49	
Politieapparaat	70			
Politiecultuur	70	31	32	
Politiehoofddoek	57	58		
Politiek	70			
Politiek correct	31	32		
Politiek correcte	31	32		
Politieke	70			
Politieleiding	70	47	49	
Politiemedewerker	70	47	49	
Pasta	57	58		
Pastafarian	57	58		
palestijn	70			

palestijnen	70
Onzin	12 16 18
Onverdraagzaam	12 16
Onverdraagzaamheid	12 16
Ongelovige	57 58 70
Ongelovigen	57 58 70
om	10 71
omdat	20 21 71
Negatief	12 16 7
Negatieve	12 16 7
Nederlander	70
Nederlanders	70
Nederlands	70
Nederlandse	70
nazi	70
nazi's	70
Moskee	58
Moslim	58 70
Moslima	58 70
Moslims	58 70
Minderheden	70
Minderheid	70
Minderheidsgroep	70
meid	31 36
Marokkaan	70
Marokkaans	70
Marokkaanse	70
Marokkanen	70
Marokko	70
maar	20 23 45 71
Levensbeschouwelijke	57
Levensbeschouwing	57
Levensovertuiging	57
Levensstijl	57
Leidinggevenden	47 49
Kwakzalver	66
Kwakzalvers	66
Kruisjes	57 57
Kortzichtig	12 16 27 28
Kortzichtigheid	12 16 27 28
Koran	57 58
Korps	70
Korpsleiding	70
Koekoek	66
Kinderfeest	31 36 51

Kerkvanhetvliegendspaghettimonster	57	58	
Kerk en Staat	57	58	
Keppel	57	58	
Keppeltje	57	58	
Katholieke	57	58	
Kakkerlak	69	66	
Kakkerlakken	69	66	
Jood	70		
Joods	70		
Joodse	70		
joden	70		
Israel	70		
Israël	70		
Issue	12	16	
Issues	12	16	
Islam	57	58	
Islamitisch	57	58	
indien	20	25	71
Inclusieve	44		
Identiteit	70		
Ideologie	57	70	
Ideologische	57	70	
Hoofddeksel	58	57	
Hoofddeksels	58	57	
Hoofddoek	58	57	
Hoofddoeken	58	57	
Hoofddoekje	58	57	
homo	60	62	66
homo's	60	62	66
hoewel	25	45	71
hoeren	60	62	66
Hijab	58		
Hijabs	58		
Hetero	60	62	
Gezeik	12	16	18
Getriggerd	12	16	18
Geschreeuw	12	16	
Gereformeerde	57	58	70
Gereformeerden	57	58	70
Gelovige	57	58	70
Geloofsgemeenschap	57	58	70
Geloofsoplegging	57	58	
Gedoe	16		
Gediscrimineerd	16		
Gay	60	62	70

Geaardheid	60	62				
Fundamentalisten	57	58	70			
Fries	70					
Friese	70					
Friezen	70					
Flying Spaghetti Monster	57	58				
Fatsoen	57					
Fatsoenlijk	57					
Extremisme	70	57	58			
Extremisten	70	57	58			
Ezel	69					
ezel	69					
ezels	69					
Ezels	69					
Europees	70					
Europese	70					
Ethiek	57					
Ethisch	57					
Etnisch	70					
Etnische	70					
en	44	71				
doodschieten	12	16	18	57	59	39
Domme	12	16	66			
Djellabah	58					
Democratische	70					
dat	45	71				
Conservatief	70					
Conservatieve	70					
Burka	58					
Buitenlander	70					
Buitenlands	70					
Buitenlandse	70					
blanke	70					
Bi	62	60				
Atheïst	70	58	57			
Atheistische	70	58	57			
aso	70	12	16	31	32	
asociaal	12	16	31	32		
Ape*	66	69				
apen	66	69				
Apen*	66	69				
Anti	7					
Anti-*	7					
Angstcultuur	12	16	17			
Allah	58					

allah akbar	58
allah akhbar	58
allahu akbar	58
allahu akhbar	58
Aap	69
[naam]	33

Appendix B – Survey items

Attitude scale

1.a Wat is je mening ten opzichte van het invoeren van de hoofddoek als onderdeel van het politieuniform?

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Schadelijk — Wenselijk

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Dom — Wijs

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Ongunstig — Gunstig

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Ongewenst — Gewenst

1.b Ik vind dat de hoofddoek onderdeel moet uitmaken van het politieuniform

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Zeer oneens — Sterk mee eens

Polarization scale

2.a In hoeverre zijn participanten binnen deze discussie het eens met elkaar?

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Helemaal niet — Heel erg

2.b Ik vind dat er sprake is van duidelijke groepsvorming in twee partijen.

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Helemaal niet — Heel erg

2.c Ik vind dat er sprake is van harde tegenstellingen tussen participanten binnen deze discussie.

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Helemaal niet — Heel erg

2.d Ik vind de participanten erg op elkaar lijken.

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Zeer oneens — Helemaal mee eens

2.e Hoe vriendelijk/vijandig vind je deze discussie?

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Zeer vijandig Zeer vriendelijk

2.f Hoe emotioneel vind je de taal die gebruikt is?

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Zeer zakelijk Zeer emotioneel

2.g Ik vind deze discussie gepolariseerd.

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Zeer oneens Helemaal mee eens

Argument quality items

3.a Hoe objectief vind je het commentaar?

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Zeer subjectief Zeer objectief

3.b Hoe overtuigend vind je de aangedragen argumenten?

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Helemaal niet overtuigend Zeer overtuigend

Involvement scale

4.a Ik voel mij betrokken tot de discussie die gaande is rond de rol van de hoofddoek binnen het politieuniform.

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Helemaal niet eens — Heel erg eens

4.b Ik vind het belangrijk dat er een discussie gevoerd wordt over de rol van de hoofddoek binnen het politieuniform.

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Helemaal niet eens — Heel erg eens

4.c Ik ben geïnteresseerd in de rol van de hoofddoek binnen het politieuniform.

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Helemaal niet eens — Heel erg eens

4.d Contact hebben met een agent die een hoofddoek draagt zal veel invloed op mij hebben.

☐ 1 ☐ 2 ☐3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9

Helemaal niet eens — Heel erg eens