Digital Blue: Investigating Sentiment and Bias in Human and AI narratives on Police Officers

Student Name: Mieke Dohmen
Student Number: 742441

Supervisor: João Fernando Ferreira Gonçalves

Master Media, Digitalisation & Social Impact
Erasmus School of History, Culture and Communication
Erasmus University Rotterdam

MA Thesis
June 6th, 2025.

Abstract

The pervasive influence of media on public perceptions of critical societal institutions, particularly law enforcement, forms a significant area of academic research. Existing literature offers a comprehensive understanding of how human-generated media shapes attitudes towards police, evolving from broad observations to detailed analyses of specific formats and audience roles. However, an emerging and crucial dimension – how Large Language Models (LLMs) generate text on police officers – remains unexplored. LLMs, trained on vast textual datasets, inherit and can propagate social biases present in their training data, raising concerns that AI models might capture, reproduce, and even intensify existing stereotypes when generating content about police officers.

Building on cultivation theory, which posits that the long-term impact of consistent media exposure on social reality perceptions, and social location theory, which highlights how an audience's background shapes media interpretation, this thesis aims to study and compare language generated by humans and AI-generated text regarding police officers. Specifically, it investigates how biases from traditional media and digital media translate into both human perception and AI models.

To address this, a mixed-method exploratory design was employed, utilising both qualitative and quantitative data collection and analysis techniques. Human participants completed a survey involving four open-ended prompts designed to elicit perceptions of neighbourhood crime and police, police-civilian communication, trust and authority, and media portrayals of the police. Concurrently, AI models, specifically GPT-4 and Gemini, were prompted with the same scenarios to generate comparable textual data. The collected textual data underwent qualitative thematic analysis to identify patterns and meanings, complemented by quantitative sentiment analysis to compare sentiment distributions. This comparison of human and AI-generated narratives revealed a significant divergence in sentiment concerning police trust and authority, with AI models often articulating a more critical stance than human respondents. This outcome suggests that LLMs may not only reflect but also potentially reinforce existing societal biases embedded within their training data.

KEYWORDS: Large Language Models (LLMs), Human Perception, GPT-4, Gemini, Police Representation.

Preface

This thesis, culminating years of academic endeavours, would not have been possible without the support and encouragement of friends and family members.

First and foremost, I wish to express my gratitude to my parents, whose endless belief in my abilities, constant encouragement, and support have been the bedrock of my academic journey. Their love and care have continuously motivated me, and for that, I am eternally thankful.

This work is dedicated to the memory of my grandfather; André. His wisdom, resilience, and love for the world and family has deeply influenced me. Though he is no longer with us, his spirit continues to inspire me, and I hope that this academic achievement will honour his enduring legacy.

I am also grateful to João Fernando Ferreira Gonçalves, who with his invaluable guidance and insights and of course unwavering positivism throughout this research process.

1. Introduction

The pervasive influence of media on public perceptions of critical societal institutions, particularly law enforcement, has long been an avenue for academic research (Dowler & Zawilski, 2007, p.193). The body of research on media representation of police officers offers a comprehensive and intricate understanding of how the media influences public perceptions of law enforcement, an area of study that has advanced from broad observations to detailed analyses of specific media formats, theories, and audience roles. Mass media serve as a primary source of public knowledge and attitudes concerning the police and the broader criminal justice system (Glascock, 2023, p. 568). However, the portrayal of policing within this media landscape is complex and often contradictory. According to research while entertainment media, such as crime dramas like *NYPD Blue* and *Law and Order*, frequently depict police officers as effective problem solvers who successfully apprehend criminals, often exaggerating arrest rates and crime clearance, previous research also points out that news media, especially network and online news, tend to focus on negative events, including police misconduct, abuse, and discrimination.

Building on these established understandings, this thesis addresses an emerging and crucial dimension: how do LLMs (Large Language Models) generated texts portray police officers? While previous research has thoroughly explored the effects of human-generated media portrayals on attitudes towards the police (Dowler & Zawilski, 2007, p.193; Intravia et al, 2018, p.965), there is a pressing need to understand the influence of artificial intelligence in shaping these narratives. LLMs, trained on immense textual datasets often scraped from the web, inherently inherit and can propagate the social biases present in that data. Consequently, AI models, when generating content about police officers, risk capturing, reproducing, and even intensifying existing biases and stereotypes absorbed from their training data. This research is aimed at studying and comparing language generated by humans and AI-generated text regarding police officers.

This thesis explores the multifaceted ways in which media, encompassing both human and AI-generated content, represent police officers and the significant influence these portrayals exert on public perceptions. Specifically, it aims to investigate how biases present in traditional and digital media translate into human perception as well as into AI models.

To examine these dynamics, this research draws on two theoretical frameworks. First, cultivation theory provides a framework for understanding the long-term impact of consistent exposure to recurring images and narratives on viewers' perceptions of social reality (Gerbner, 1998, p.177). The second framework is social location theory (Intravia et al, 2017, p.159). This framework highlights that the interpretation and impact of media messages are fundamentally shaped by the audience's unique background and social positions.

This study employs a mixed-methods approach, combining both qualitative and quantitative data collection and analysis techniques. An exploratory design was utilised, involving a survey-based approach for human participants and prompting AI models (specifically GPT-4 and Gemini) with the same scenarios to gather comparable textual data. The collected data underwent qualitative thematic analysis to identify key patterns and meanings, complemented by quantitative sentiment analysis to compare statistical sentiment distributions between human and AI-generated responses.

The collected data underwent qualitative thematic analysis to identify key patterns and meanings, complemented by quantitative sentiment analysis to compare statistical sentiment distributions between human and AI-generated responses. This thesis is structured to systematically address these dynamics, first establishing the pervasive influence of media on public perceptions of law enforcement and highlighting the emerging, yet unexplored, role of Large Language Models (LLMs) in shaping these narratives. Next, the theoretical framework will delve into cultivation theory and social location theory, providing foundational concepts for understanding media impact and audience interpretation. The following section details the mixed-methods exploratory design of this study, including the survey administered to human participants and the prompts used to generate comparable text from AI models (GPT-4 and Gemini), along with the qualitative thematic analysis and quantitative sentiment analysis techniques employed. This research is academically relevant as it contributes to the understanding of media effects in a new digital age and explores the biases inherited by AI models. From a societal perspective, this study is relevant because the majority of public knowledge about crime and justice is derived from the media, and the portrayal of law enforcement significantly shapes public attitudes. As LLMs are trained on vast datasets that reflect existing media and societal biases, there is a considerable risk that these models will reproduce and even intensify existing stereotypes about police officers, potentially influencing public opinion and contributing to discriminatory practices.

2. Theoretical framework

The body of research on media representation of police officers offers a comprehensive and intricate understanding of how the media influences public perceptions of law enforcement. This area of study has progressed, moving from initial broad observations to more detailed analyses of specific media formats, underlying theories, and the audience's role. The early works established the fundamental role of mass media as a primary source of public knowledge and attitudes concerning the police and the broader criminal justice system (Intravia et al., 2017, p.963). This literature review delves into the multifaceted ways in which the media represent police officers and the considerable influence these portrayals exert on public perceptions, considering relevant theoretical frameworks and empirical findings. To provide a nuanced understanding of these dynamics, and in alignment with the overarching aim of comparing human and AI-generated narratives, this framework contextualises the research by exploring key areas of perception that guide the empirical investigation. This includes examining how human participants and AI models perceive crime and police presence within a neighbourhood context, considering factors that influence feelings of safety and concerns about police capacity and timeliness. The framework also investigates perceptions regarding police communication with civilians and its influence on public interaction, delving into aspects like tone, attitude, and the perceived effectiveness of police communication strategies. Furthermore, the framework explores the diverse factors that determine trust in the police and their perceived authority from both human and AI perspectives, investigating criteria such as conduct, competence, fairness and accountability. Finally, the framework considers how portrayals of police officers in media resonate with the perspectives articulated by human participants and AI models, acknowledging the influence of various media types and their role in shaping attitudes towards law enforcement. This comprehensive approach allows for an exploration of how biases, absorbed from vast training datasets, might manifest or be amplified in AI-generated narratives, potentially influencing public understanding of police officers in ways that may diverge from human sentiment, especially concerning aspects of trust and perceived authority. This also allows for an investigation into how human perceptions are shaped by their media consumption patterns and demographic factors, offering a more complete picture of how beliefs about police are formed in the current information age.

2.1 Conflicting Images: Positive and Negative Portrayals

The media landscape presents a complex and often contradictory picture of policing. Research indicates a divergence between positive and negative representations across different media types. Entertainment media, particularly crime dramas such as NYPD Blue and Law and Order, frequently feature police officers and prosecutors. These shows often depict police officers as effective problem solvers who successfully apprehend criminals (Dowler & Zawilski, 2007, p.196). This narrative often exaggerates arrest rates and crime clearance, contributing to a perception of police effectiveness that may exceed reality.

Conversely, news media, especially network and online news, tend to focus on negative events, including instances of police misconduct, abuse, and discrimination. Frequent viewers of network news were more likely to report that police misconduct was a common or frequent occurrence (Dowler & Zawilski, 2007, p.196). However, it is noted that theories about television police portrayals need to be fully developed within the context of police dramas, police reality shows, and crime-solving shows (Glascock, 2023, p576-580).

Previous research has explored the impact of consuming these varying media portrayals on attitudes towards police. Dowler & Zawilski (2007, p.198) revealed that heavy viewers of network news are more likely to believe that police misconduct is frequent, particularly among minority respondents, who are also more likely to believe that White individuals receive better treatment from the police, while frequent viewers of crime-solving shows may believe they do not. Another study by Soulierre (2004, p. 222-225) provides a content analysis comparing the depiction of police in two entertainment shows, *Law and Order* and *NYPD Blue*, with one reality-based show, during the 1999-2000 television season. The research aimed to determine the accuracy of these portrayals across different dimensions and compare them to real-world policing statistics. The findings indicate that, with the notable exception of being portrayed as overly successful, television police in entertainment shows were depicted as largely as real-life police in terms of their demographics, organisational structure, and general activities. While male and White officers were overrepresented compared to the general population, it mirrored the actual composition of most US law enforcement agencies at the time. However, police success rates were significantly exaggerated across all three shows. The research concludes that contemporary prime-time entertainment justice programs can offer viewers a basic understanding of police and police work due to their realistic depictions (Soulierre, 2004, p.222-225).

Previous research has indicated that media coverage of police brutality has a strong influence on negative public sentiment expressed on platforms like Twitter (Saccar et al, 2024, p.13-15). While media coverage of local crime might be expected to impact public opinion, this research indicates that it was not identified as a salient driver of overall public sentiment towards the police compared to reports of brutality. The research by Succar and colleagues (2024, p.14) investigates the relationship between media coverage of police brutality and local crime alongside public sentiment on Twitter. Their findings indicate an increase in media reports of police brutality, linked to a rise in negative sentiment towards the police on the social media platform. However, the analysis found no statistically significant evidence to suggest that media coverage of local crime strongly influences public sentiment towards law enforcement. Also, the results indicated that this public response seemed to occur nationally. This emphasis on negative incidents can cultivate a perception of police as ineffective or even incompetent. Moreover, the rise of police reality shows adds another layer. Some studies have examined the content of police-related dramas and reality shows. Police reality shows often present police officers as knowledgeable, sensitive, and competent, reinforcing a "law and order" ideology. Research by Doyle (2018) analyses the reality-TV program COPS and argues that its influence extends far beyond simply shaping audience perceptions of law enforcement (p.3-4). Doyle contends that COPS actively promotes this ideology through various narrative and production techniques while claiming to show "raw reality" (Doyle, 2018, p.14). The research suggests that reality TV shows contribute to a state of "hyperreality" by blurring the lines between fact and fiction, television and real life, and news and entertainment. The analysis highlights how COPS selectively presents events to highlight effective policing and rarely airs footage that reflects negatively on law enforcement.

These divergent media narratives reveal a critical tension in public perceptions: news coverage frequently highlights police brutality and misconduct, which has been linked to increased negative public sentiment and a perception of ineffectiveness. Conversely, entertainment genres such as police reality shows often present a contrasting image, portraying police officers as competent and effective while reinforcing a 'law and order' ideology through selective portrayal. Consequently, public attitudes towards law enforcement are shaped within this complex and often contradictory framework of positive and negative media representations,

2.2 Cultivation Theory: Shaping Long-Term Perceptions

Cultivation theory provides a framework for understanding the long-term impact of these media portrayals. Gerbner (1998) investigated how the increasing growth of television viewership affected viewers' conceptions of social reality (Intravia et al, 2017, p.961-964). According to this theory, consistent exposure to recurring images and narratives can shape viewers' perceptions of social reality (Intravia et al, 2017, 964). While other researchers argue that television news can create a 'scary world' view (Garcia-Castro and Pérez-Sánchez, 2018, p.13), other factors can influence this relationship. An analysis conducted in Costa Rica found that narrative transportation – the extent to which individuals become immersed in news stories – was a significant predictor of fear of crime. This suggests that simply consuming news may not be the only factor; the psychological experience of being drawn into the narrative also plays a crucial role. The more time spent consuming media, the greater the likelihood that users' perceptions of the real world will align with what is depicted in the media. Heavy viewers of media that consistently portray the police in a certain light may gradually adopt those perspectives on their own. This can manifest in both first-order effects, such as beliefs about the prevalence of crime, and second-order effects, which involve more general attitudes and beliefs about the police, such as their fairness and effectiveness (Glascock, 2023, p.573-574). For example, for individuals who consistently were exposed to news reports of police brutality, cultivation theory would suggest that they may come to believe such events are more common than they are. Conversely, heavy viewing of crime dramas where police are successful might cultivate an unrealistic expectation of police effectiveness.

The early formulation of cultivation theory faced criticism for being too broad and assuming that all viewers would be affected by media consumption in the same way, regardless of their characteristics and backgrounds (Intravia et al, 2017, p 159). This led to the development of reception research or audience effects theory, which explores how factors such as demographics, social context, and personal experiences mediate the relationship between media consumption and attitudes. These perspectives suggest that the media effect might be stronger for certain subgroups of the population based on factors like their social location, resonance with media portrayals, perceived vulnerability, or lack of real-world experience with the issues depicted. While initially focused on television, scholars recognise that the core arguments of cultivation theory can apply to other forms of media, including the Internet and social media (Intravia et al, 2017, p.158-163). For example, the more time young adults spend on social media, the more likely they are to experience an increased fear of crime, although this relationship can vary based on their perceptions of safety. Similarly, consuming news online has been linked to negative attitudes towards police legitimacy. The study by Hermann and Eisand (2020) directly addresses the application of cultivation theory to Facebook as a new media vehicle. They argue that while traditional cultivation focuses on television's ubiquitous and unified message system, Facebook presents a more fragmented and individualised message system shaped by user-generated content and interactions. Through a survey and follow-up experiment conducted with a German student sample, Hermann and Eisand (2020, p 1131) found that higher Facebook use is associated with stronger perceptions of ethnic diversity in society and more positive attitudes towards ethnic minorities. Notably, their mediation analysis suggests that the cultivation of positive attitudes is mediated by users' perceptions of ethnic diversity within their close social environment on Facebook.

A cross-national study by Mustafaj and Van den Bulck (2021, p.274) also offers valuable insights, as their study revisits cultivation theory by examining the relationship between television viewing and fear and mistrust across 27 European countries, using data from the European Social Survey (ESS). The study found that welfare state regimes moderate the relationship between television viewing and outcomes like fear of victimisation and trust. This research provides evidence for mainstreaming effects across different welfare state regimes, suggesting that the societal context influences cultivation processes. Underscoring the need to consider broad contextual factors when employing cultivation theory.

It is also important to note that cultivation theory is being applied by the researcher as motivation to better understand the impact of AI-generated content. As AI models are trained on vast datasets of text and media, there are concerns that they might cultivate representations of social groups, including police officers, potentially influencing public perceptions if individuals are heavily exposed to AI-generated narratives. Cultivation theory provides a valuable lens for understanding how generative language models, trained on vast datasets reflecting existing media and societal content, might inadvertently cultivate representations of police officers. Exposure to AI-generated content that reproduces biases and stereotypes present in its training data could contribute to shaping user perceptions, potentially reinforcing existing attitudes towards law enforcement.

2.3 Critical perspectives: power, bias and stereotypes

From a critical standpoint, media portrayals, including those of the police, are seen as actively contributing to the maintenance of existing power imbalances and societal inequalities, rather than simply reflecting reality. These viewpoints contend that the media often functions to legitimise the authority of dominant groups and the prevailing social order. For instance, law enforcement may be frequently depicted in ways that emphasise a division between order-keepers and those who threaten order, potentially obscuring systemic issues within the justice system. This critical lens also draws attention to the "minority threat hypothesis", where media depictions can associate minority groups with criminality, which in turn can be used to justify increased policing and control within those communities (Dowler & Zawilski, 2007, p.196).

Moreover, critical perspectives highlight how media narratives can sustain damaging stereotypes concerning various social groups concerning crime and justice. This is particularly relevant to generative AI models, which, by being trained on extensive datasets that may contain societal biases, can inadvertently capture and amplify these stereotypical biases in their generated content. This can have tangible consequences by influencing public opinion and potentially contributing to discriminatory practices. The creation of an "us versus them" mentality in media representations is another key concern, where law-abiding citizens are often implicitly aligned with dominant groups, while criminals are frequently racialised or marginalised, potentially hindering a deeper understanding of the underlying causes of the crime. Expressions like "ACAB" (All Cops Are Bastards) and "FTP" (Fuck The Police), prevalent in social movements, are viewed from a critical perspective as significant indicators of a rejection of what is perceived as oppressive police power, by those who directly experience its negative effects (Wall & McClanahan, 2025, p.?). These slogans, despite their bluntness, are seen as naming the police power as fundamentally coercive, often racist and violent, and as an obstacle to a more equitable society. While some might argue about the individualising aspect of "ACAB", others contend that its central critique targets the systemic nature of policing and its impact on marginalised communities.

Importantly, critical perspectives, drawing on theories of ideology, recognise that while dominant worldviews are embedded in media, audiences are not passive and can engage in resistance and negotiation of meanings. Concepts like polysemy suggest that popular media needs to be sufficiently open to allow diverse subcultures to find meanings that resonate with their identities, often about or in opposition to the dominant ideology present. Subordinate groups can even excorporate by repurposing dominant cultural products into resisting discourses (Fiske, 1986, p.400-401). The insights from these critical perspectives are increasingly vital when considering generative AI and its representations of police officers. As AI models are trained on vast amounts of data reflecting existing media and societal biases, there is a considerable risk that they will reproduce and even intensify these biases in their portrayals of law enforcement. This could lead to the further entrenchment of stereotypes and the reinforcement of unequal power dynamics, potentially worsening tensions between police and specific communities.

However, another limitation to consider in analysing AI-generated content about police, based on some of the available research on biases in language models and media effects, is that the examples and baseline comparisons often appear to be more US-centric. Additionally, while generative language models are trained on extensive datasets, including potentially vast amounts of internet content, discussions and measurements of bias do not always explicitly focus on how biases from social media are captured. This is a crucial aspect, especially considering that LLMs are trained on social media data and that social media is an increasingly significant source of news and information. Understanding how biases from platforms like Facebook or X, where users actively disseminate news and engage in discussions, are integrated and potentially amplified by AI models is important for a comprehensive assessment of their potential impact on public perceptions of police.

2.4 Media influence on trust, legitimacy and cooperation

Media portrayals wield a significant influence over public trust in the police, perceptions of their legitimacy, and the willingness of citizens to cooperate with law enforcement (Glascock, 2023, p.583). As most of the public's knowledge regarding crime and the criminal justice system is derived from media consumption, the way police are depicted plays a crucial role in shaping public attitudes. Negative media coverage, particularly focusing on police misconduct, abuse and discriminatory practices, could significantly erode public trust in law enforcement institutions. Frequent exposure to such narratives, especially through the network and online news, can lead viewers to believe that these incidents are common occurrences. This erosion of trust can contribute to a "legitimacy crisis", where the authority of the police is challenged and public confidence in their integrity diminishes (Intravia et al., 2020, p.58-59). When the police are not seen as trustworthy and acting legitimately, citizens may be less inclined to accept their decisions or comply with their directives.

The impact of media on trust, legitimacy and cooperation is further complicated by the differential influence of various media types. As mentioned before, research shows that varying impacts influence the complex relationship between media consumption and attitudes towards the police. However, maintaining public perception of the police is fundamental for maintaining trust, ensuring legitimacy, and fostering citizen cooperation. A lack of trust stemming from negative media coverage can directly impact citizens' willingness to cooperate with the police, such as reporting crimes or providing information. When police legitimacy is undermined by negative media narratives, the public may be less willing to comply with the law, potentially threatening public order. Research shows that varying impacts influence the complex relationship between media consumption and attitudes towards the police (Dowler & Zawilski, 2003, p.195).

2.5 The role of audience and social location

The interpretation and impact of media messages, particularly those concerning the police, are fundamentally shaped by the audience consuming them and their social location within society. Rather than being passive recipients of media content, individuals actively interpret what they see and hear through the lens of their unique backgrounds and social positions. Social location theory posits that an individual's place within the social structure, defined by factors such as race, class, gender, age, ethnicity and sexual orientation, significantly influences how they understand and experience the world. These factors shape their perspectives, values and attitudes, including their perceptions of authority figures like the police. Reception research, also known as audience reception theory, builds upon this by exploring how these audience characteristics impact the reception and interpretation of media. It moves away from earlier models that viewed audiences as homogenous and easily influenced, recognising the active role of the viewer in making meaning. Fiske (1986) argues that popular television needs to be an "open text" that allows diverse subcultures, defined by their relation to centres of domination, to generate meanings that resonate with their own socially located identities (p.392). Therefore, when analysing the interpretation of police related media texts, consider how different social locations might lead to diverse readings and interpretations. For example, a media representation of a particular ethnic group might be interpreted very differently by someone within that group compared to someone outside of it, based on their lived experiences and social positioning.

The interplay between social location and media consumption leads to the concept of polysemy, the idea that media can have multiple meanings. Different subcultures and social groups, defined by their relation to dominant ideologies, will interpret texts in ways that resonate with their own identities and experiences (Fiske, 1986, p. 392-393). Previous research has demonstrated that race and socioeconomic status are significant factors influencing attitudes towards the police and the impact of the media. The concept of the minority threat hypothesis, for example, suggests that media portrayals can contribute to associating minority groups with criminality, potentially justifying increased police surveillance and control in these communities.

Social location theory is fundamental to understanding the multifaceted nature of social inequalities. The work by Anthia (2012) critiques the class analysis and suggests that solely focusing on economic factors overlooks the crucial roles of gender, ethnicity, and other social divisions in shaping hierarchies (124). Additionally, Anthia discusses intersectionality extensively, arguing that these social divisions are mutually constitutive, meaning they interrelate and create unique experiences of oppression and privilege based on an individual's intersecting social locations. When analysing how social location influences how people interpret media representations of police officers, the perspective of how various forms of inequality are interconnected and experienced differently depending on social location is considered. More recently, academics are focusing on digital media, for example, mimetic wherein social location still provides context (Katz & Shifman, 2017, p.838). The concept of affective meaning and the formation of "phatic, image-oriented, communities" around even nonsensical memes might be influenced by shared aspects of social location or experiences.

Individuals from similar social backgrounds might find resonance and connections through these digital forms of communication, even if the referential meaning is unclear. Therefore, considering how online communities and the meanings generated within them might reflect or be shaped by the social locations of their members. Other academic work demonstrates the connection between socio-demographic aspects of social location (like age, gender, education, and income) and broader human value priorities. Arguing that one's position in the social structure affects the values one prioritises (Keil & Gabriel, 2012, p.2-4). Crucial for understanding how individuals from different social locations might evaluate media portrayals of police officers based on their differing values. The study by Keil & Gabriel (2012) conducts a European-wide comparison investigating the relationship between an individual's social location and their prioritisation of human values. The authors note that while some of the relationships between social location and values are cross-culturally consistent, others are context-dependent; this adds crucial nuance to the analysis of media reception of police officers across various social locations (Keil & Gabriel, 2012, p.26-27).

Anthia’s work also emphasises the transnational frame, highlighting the importance of considering how social locations are shaped and experienced across national borders. Migrants, for example, occupy social locations that are influenced by their country of origin, their experiences in the new country, and their positioning within transnational social structures. This lens focuses on social locations rather than fixed group categories, recognising that individuals are embedded in hierarchical relations within multiple specific situational and conjectural spheres (Anthia, 2012, p.130). Highlighting that lives are located across multiple but also fractured and interrelated social spaces of different types, both locally and transnationality, which can produce complex and even contradictory social locations. This perspective is particularly relevant in understanding how individuals positioned across nation-state borders or within complex multicultural spaces interpret media representations. Anthia's work highlights the importance of intersectionality, understanding how different social categories such as gender, ethnicity, and class intersect and create unique experiences of inequality and hierarchy (126). Understandings of class, for instance, are often gendered and ethnicized, influencing how individuals from different social locations perceive authority figures like the police. Intersectionality moves beyond viewing these categories in isolation to understand their interconnectedness in shaping social experience and interpretation.

2.6 Sources of Bias in LLM Training Data

Large Language Models (LLMs) are sophisticated computational systems designed to generate and process human language, trained on enormous textual datasets (Navigli et al, 2023, p.2-4). Models like GPT-2 and GPT-3, built upon architectures like the Transformer (Vaswani et al, 2017, p2-4), have demonstrated impressive performance across a variety of Natural Language Processing (NLP) tasks, including generating text, translation, summarization, and sentiment analysis (Wen and Younes, 2024, p.2). Their capability stems from being trained on massive text corpora, often scraped from the web. However, this reliance on vast, real-world data means that LLMs inevitably inherit and can propagate the social biases present in that data (Navigli et al, 2023, p.13).

Bias in LLMs can arise fundamentally from data selection bias, which is introduced by the choices made about which texts are included in the training corpus. Even widely used sources like Wikipedia, often considered high-quality data for NLP, exhibit significant imbalances. For instance, Wikipedia's content is heavily skewed towards domains such as Sports, Music, Places, Media, and Politics, with much less representation in areas like literature or economy

(Navigli et al, 2023, p.3). Beyond domain imbalance, the demographics of the data creators can introduce bias; for example, Wikipedia editors are disproportionately English speakers and do not reflect the global population or country-specific demographics. This means the worldview presented in the data is often predominantly reflective of the Global North (Floridi and Chiriatti, 2020, p.5). Every document selected for training contains its own information and potential social biases, making selection an inherently influential process. These AI models are trained on vast datasets, which inevitably include existing text and media content that reflects societal biases and stereotypical representations, including those found in traditional and digital media portrayals of the police. When these models generate content, there is a risk that they will capture, reproduce, and even intensify these existing biases and stereotypes in their portrayals of law enforcement.

These biases manifest in LLMs through stereotypical associations linked to various demographic and social categories such as gender, race, religion, and profession. For instance, bias can be seen when a model disproportionately links certain attributes or occupations to specific groups. Previous research has explored intersectional bias, examining how biases interact when related to multiple characteristics like gender and ethnicity (Kirk et al, 2021, p.3). Buolamwini & Gebru (2018) also revealed that word analogy (e.g., "man is to computer programmer as woman is to homemaker") demonstrates how societal gender biases are encoded within word embedding models, leading to stereotypical associations and completions in word analogy tasks (p.1-3). The work by Buolamwini & Gebru reveals a series of mechanisms that amplify discriminative behaviours which are representative of already existing and deep-rooted social inequalities in AI systems. Researchers have identified and measured biases across various protected attributes and through different model behaviours. Prominent types of social biases include those related to gender, where models may stereotypically associate genders with specific occupations or attributes, such as disproportionately linking female individuals with roles like ‘waitress’, ‘nurse’, or 'maid’, or showing bias in pronoun resolution (Kurita et al, 2019, p.4-5). Race and ethnicity biases are also well-documented, manifesting as associations between racial or ethnic groups and positive or negative sentiment or ‘regard’ (Sheng et al, 2019, p.4). Religion bias appears as stereotypical associations with attributes or professions, like predicting Buddhist men to be monks or associating religious titles with specific faiths. Sexuality bias has been examined, particularly in how it intersects with gender to influence occupational associations (Kirk et al, 2021, p.3-5). Nationality bias can lead to stereotypical associations, such as the example of a model completing a sentence with “They are Chinese, so... they are not very good at English”. Biases often intersect, meaning that the combination of attributes like gender, ethnicity, religion, or sexuality can create unique biases that are not apparent when considering attributes in isolation, influencing outcomes like the likelihood of predicting certain outcomes.

Just as traditional media exposure cultivates certain views over time, exposure to content created or influenced by AI models that have absorbed media biases about policing could similarly impact how police perceive police effectiveness, the prevalence of misconduct or discrimination, police legitimacy, trust, and bias. For instance, if AI models trained on biased media data produce content that exaggerates police effectiveness, normalizes misconduct, or perpetuates stereotypes about which groups are more likely to be involved with crime or treated unfairly by police, this content could further shape public perceptions as AI-generated text becomes more prevalent and integrated into the information ecosystem. The ability of AI to generate text closely resembling human language makes the potential for such influence particularly pertinent, as users may not always differentiate between human-generated and AI-generated content.

2.7 Measuring Bias in LLMs

Measuring bias in LLMs involves different approaches. One common method is to probe the model directly using template sentences. For example, templates like “T are A” or “T likes A” are used, where ‘T’ represents target words (like demographic groups or professions) and ‘A’ represents attribute words (characteristics or associated concepts) (Kurita et al., 2019, p.3). This allows researchers to quantify how likely a model is to associate specific targets with specific attributes. The Word Embedding Association Tests (WEAT) is an example of such a method, originally used to study biases in word embeddings but adapted for models like BERT (Kirk et al, 2021, p.3-7).

Another measurement involves Context Association Tests (CATs), notably implemented in the StereoSet dataset. StereoSet is designed to measure stereotypical bias in pretrained language models across domains, including gender, profession, race, and religion. CATs present a context about a target group and require the model to evaluate the likelihood of association with stereotypical, anti-stereotypical, or unrelated instances. For intersentence CATs, the task is like BERT's next sentence prediction, where the model must choose a follow-up attribute sentence given a target sentence (Nadeem et al., 2020, p.2-3). Models like BERT, ROBERTA, XLNET, and GPT-2 have been evaluated using StereoSet, showing varying stereotypical scores (Nadeem et al., 2020, p.7). Metrics such as the Language Model Score (LMS), Stereotype Score (SS), and Idealised CAT (ICAT) score are used to evaluate models on this task. StereoSet revealed that current pretrained language models exhibit strong stereotypical biases. The size of the model can influence its language modelling ability and correspondingly its stereotypical score, but surprisingly, large models do not always achieve better idealised CAT scores (Nadeem et al., 2020, p.9).

Just as traditional media exposure cultivates certain views of police over time, AI-generated content, potentially infused with learned media biases about policing effectiveness, misconduct, legitimacy, trust, and bias, presents a new channel through which similar cultivation could occur. If individuals are heavily exposed to AI-generated narratives that perpetuate or amplify stereotypes and biases absorbed from media training data, these portrayals could further shape public understanding and attitudes towards police, potentially entrenching existing views based on skewed representations or contributing to a “legitimacy crisis” (Intravia et al., 2020, p.58-59). The understanding and acknowledge that public perceptions potentially could be shaped not only by human-produced media but also AI-systems that have learned from, and may even amplify, the biases present in that existing media environment, offering a more comprehensive picture of how beliefs about police effectiveness, misconduct, legitimacy, trust and bias are formed in the current information age.

3. Methodology

This research was designed to explore and compare the language generated by humans and AI models in response to prompts concerning perceptions of police officers. To achieve this, a mixed-methods approach was employed, incorporating both qualitative and quantitative data collection and analysis techniques.

This research utilises an exploratory design to investigate how humans and AI language models frame their understanding of police trust, authority, and related factors. By employing a survey-based approach for human participants and prompting AI models with the same scenarios, the study gathered comparable textual data for qualitative analysis and structured responses for quantitative analysis. The qualitative component, including sentiment analysis and structured questions, provided measurable data for comparison between groups and correlation with other variables like media consumption and demographic factors

3.1 Sample

A convenience (snowball) sampling method was used by sharing a link to the questionnaire on various social media platforms to recruit participants. A link to the questionnaire was shared by multiple acquaintances, and a small group of friends distributed the link through their online networks. Answers from participants who did not fill in the questionnaire completely were removed from the dataset, which resulted in the dataset containing responses from 45 participants. Participation was voluntary. Completing the questionnaire took 5 to 10 minutes.

The AI sample consists of text generated by the GPT-4 and Gemini language models. A total of 4 prompts were asked 4 times per model, resulting in 32 completed prompts. A new e-mail account was created to access Gemini, and each prompt was integrated after the other to maintain consistency and minimise bias in data collection. Combining the human survey completions and the GPT-4 and Gemini completions resulted in a total sample size of 77 unique completions for analysis.

The sample consisted of 45 participants (11 men; 33 women; 1 preferred not to say) (*M* = 1.80, *SD* = .59) with an age range between 20 and 70 years (*M* = 45.33, *SD* = 14.84). Participants were asked to indicate the highest level that they had reached: 20% had obtained a master's degree, 22.1% had obtained an bachelor's degree, 14.3% had obtained a secondary degree, and the final 17.8% had obtained an high school degree (M = 3.60, *SD* = 1.01). Participants were also asked to indicate their nationality; 43 participants indicated that to were from the Netherlands, and the other 2 participants were from Belgium (*M* = 116.38, *SD* = 21.67). Additionally, participants were asked to indicate their employment status, 37.8% of the participants are full-time employed, 35.6% of participants are employed part-time, 6.7% of participants are self-employed, unable to work or retired, and the remaining 4.4% of participants are students (*M* = 2.51, *SD* = 2.02). During the data cleaning process, the median duration for checking the duration of the respondents was found to be *Mdn* = 556.00. After establishing this median duration, there were no additional excluded responses.

3.2 Measures and procedure

Before the start of the survey, participants were informed that the questionnaire focused on opinions about police officers and the media representation of police officers. It was explained that the data were collected anonymously, that participation was voluntary, and that the data were used for academic purposes only.

The main part of the survey consisted of open-ended scenarios or statements that participants were asked to complete by adding sentences or words. These completions were intended to be a consistent continuation of the original scenario or statement. Participants were encouraged to respond truthfully based on their values and ideologies, with no right or wrong answers. The prompts were designed to elicit responses related to specific themes and identity categories, like those used for the AI text generation.

The first open question (Q5): "Thinking about police and crime in my neighbourhood, what often comes to mind is...." This prompt taps into perceptions of crime and police in a specific geographical location (Intravia et al, 2017, p 964). Previous research highlights that an individual's perceptions of crime rates and criminal activity in their neighbourhood can be influenced by ecological factors. The second open question (Q6): "When I hear people talk about police communication with civilians, I think of...". This scenario directly relates to the concept of procedural justice, which is a significant determinant of police legitimacy and public trust (Glascock, 2023, p.594). This procedural justice involves the perceived fairness of how police treat citizens, including aspects like whether police officers treat people fairly, neutrally, and respectfully during interactions. The third prompt (Q7): "Whether I trust the police and see them as authority, comes down to things like...". This prompt directly addresses the core concepts of trust, authority, and police legitimacy (Intravia et al, 2017, p964). Previous research has emphasised that public attitudes towards the police, including willingness to cooperate, are strongly tied to whether the police are viewed as legitimate (Glascock, 2023, p.594-603). The final open question (Q8): "The portrayal of police officers in media, which is ..., resonates with... ". This prompt specifically focuses on the influence of media portrayals on public perceptions of police. Research shows that the media is a fundamental source of knowledge about crime and justice for most of the public, and media consumption can shape attitudes towards law enforcement (Dowler & Zawilski, 2007, p.194-198).

The media consumption patterns were assessed on a Likert-type scale ranging from 1 (Never) to 5 (Every day or every other day). The self-reported frequency of respondents' consumption of various media types (e.g. local news, national news, online news, newspaper, entertainment). Additionally, how often participants see content about specific topics like police misconduct, discrimination, effectiveness, and fair treatment is also assessed on a similar Likert-type scale. Participants were also asked whether they had been in direct contact with a police officer in the past 5 years and how satisfied they were with this interaction. In the final part of the questionnaire, demographic information from participants, including age, gender, education level, nationality, religion, and employment status, was collected. This demographic data serves to contextualise the human responses and potential variations within the participant group.

This study also includes AI-generated text, such as GPT-4 and Gemini. These models are representative of advanced systems trained on extensive datasets, capable of generating human-like texts. For each of the four open-ended questions used in the human survey, four responses were generated from both ChatGPT and Gemini. The same prompts used for the survey were input for the models to elicit their continuations of the provided scenarios. To ensure consistency and replicability in the AI-generated outputs, Gemini was used with a new email account. This is to ensure that the model produces the most probable response each time, and this is not connected to the previous conversation.

3.3 Data analysis

The primary analytical approach involved characterising the sentiment of each collected text response. Responses from both human participants and the AI models were categorised into one of three valence categories: Negative, Neutral, or Positive. These categories were assigned manually by the researcher; an inter-coder reliability test was done by a fellow student to ensure the consistency and uniformity of the coding. By providing the second coder with 10 responses, previously coded, for independent review. There was an initial 90% agreement on the coding of these 10 items. This single discrepancy was discussed, leading to a refinement of the coding rules and adjustment of the original code. While the focus of the text completions is Thematic Analysis, the study also involved quantitative analysis. The quantitative analysis was conducted using the statistical software package SPSS. This facilitated the examination of numerical data collected through the survey and potentially numerical aspects of the AI text. The quantitative analysis contributed to the study by providing descriptive information about the sample, allowing for comparisons of numerical variables across different groups, and exploring relationships between variables. Descriptive statistics were utilised to summarise the characteristics of the survey participants and key variables like media consumption. These statistics provide a foundational understanding of the data being analysed. A Chi-Square test was employed to assess the association between categorical variables, such as examining if there was a statistically significant relationship between participant demographic categories, and other categorical survey responses. An ANOVA was run to compare the meaning of the continuous variable (valence) across two or more independent groups. A linear regression was applied to explore the relationship between a dependent continuous variable and one independent predictor variable. This method allows

for assessing the extent to which participant characteristics are predicted or associated with numerical ratings given to the text's completions.
The quantitative analysis conducted in SPSS provides crucial numerical context and allows for statistical testing of relationships and differences, complementing the in-depth qualitative insights gained from TA. It helped to systematically analyse participants' background data and assess numerical outcomes related to the text completions, such as average sentiment scores, which are difficult to capture through qualitative analysis alone.

Thematic analysis (TA) is also used for identifying, analysing, and interpreting patterns of meaning - known as 'themes' - within qualitative data (Braun & Clarke, 2014, p.2). It's described as a versatile analytical tool, or rather a technique, rather than a rigid methodology bound to a specific theoretical framework.
The core process of TA involves generating codes, which are considered the smallest units of analysis that capture interesting features potentially relevant to the research question. These codes then serve as the building blocks for developing themes, which are larger patterns of meaning underpinned by a central organising concept or shared core idea. Themes provide a framework for organising and reporting analytical observations. The primary aim of a TA is not simply to summarise the content data, but rather to identify and interpret key features relevant to the research question. TA is well suited for analysing the textual data collected in this research because it applies to any type of qualitative data, including qualitative surveys and story completion, which in this case is analogous to the prompt completion method used in the survey and for the LLMs. It is also suitable for both large and small datasets, and homogenous or heterogeneous samples, allowing analysis of the collected human and AI-generated text. TA can be employed for both inductive (data-driven) and deductive (theory-driven) analyses, offering flexibility in exploring emerging patterns in the data or investigating themes related to pre-existing concepts like bias regarding police officers. To maintain rigour in TA, systematic procedures such as reviewing themes against coded data and the entire dataset are important. Analysing text from AI models using methods like TA allows for comparison of discursive features between human and algorithmically generated text. This systematic generation of codes and themes facilitates a direct comparison of discursive features, revealing not only what aspects of attitudes towards the police are

highlighted by humans versus AI, but also how these perceptions are framed and what underlying assumptions or biases might be embedded in the language used by each. This is particularly vital for understanding how LLMs, trained on vast datasets that reflect existing societal biases, might inadvertently capture and amplify stereotypes in their generated content. This qualitative depth provides context for the quantitative findings.

3.4 Ethical considerations

Participants are fully informed about the research purpose, procedures, time commitment, and data handling before agreeing to participate. Participants are informed, and it is stated that participation is voluntary, with the right to withdraw at any time. Responses are guaranteed to be anonymous, and the collected data is limited to research purposes and securely stored, not shared with third parties. Data anonymisation for publication is also mentioned. Collected data is stated to be securely stored.

4. Results

4.1 Quantitative Analysis of Attitudes of Police Perceptions and Media Influence

This section presents the quantitative findings of the study, derived from analyses conducted using SPSS. The primary analytical step involved manually characterising the sentiment of each collected text response, categorising them into Negative (0), Neutral (1), and Positive (2) valence. Quantitative analysis serves to provide crucial numerical context and allows for statistical testing of relationships and differences, which is essential for quantitative data. The results are structured first, describing participants of the survey, presenting the frequency and percentage distributions of sentiment, correlations between key variables and whether media consumption influences sentiment.

4.1.1 Sentiment Distribution Across Human and AI Responses

For the quantitative comparison, crosstabulation tables were constructed to display the frequency and percentage distribution of responses across the three valence categories (Negative, Neutral, Positive) for each source (Human, ChatGPT, Gemini). To determine if there was a statistically significant association between the source of the response (Human, ChatGPT, Gemini) and the assigned overall valence, a chi-squared test for independence was conducted. This cross-tabulation and Chi-Square analysis is conducted in a separate dataset. In Table 1, the aggregate sentiment distribution across all found open-ended questions (Q-5-Q8) showed some differences between respondents and AI models, although this overall difference was not statistically significant.

Table 1: Aggregate valence across all questions and responses (Q5-Q8, N=212)

Valence	Human	ChatGPT	Gemini	Total
Negative	26.1 %	50 %	37.5 %	28.8 %
Neutral	49.9 %	43.8 %	56.3 %	49.5 %

Positive	24.4 %	6.3 %	6.3 %	21.7 %
Total	100 %	100 %	100 %	
Chi-square	7.72	*p* .100		

Additionally, this analysis was conducted for each question (Q5, Q6, Q7, Q8). This provides a clear visual of how sentiment varies by source and question. To determine if there was a statistically significant association between the source of the response (Human, ChatGPT, Gemini) and the assigned valence, a Chi-Square Test, specifically the Pearson Chi-Square test, was conducted. The Chi-Square test is appropriate for examining the relationship between categorical variables such as source and sentiment category. A p-value of less than 0.05 was used as the threshold for statistical significance, indicating that the observed association between source and valence was unlikely to have occurred by random chance. Additionally, the Linear-by-Linear Association test was examined to identify the potential linear trends in the distribution of sentiment categories across different sources.

The most significant finding is the statistically significant difference in the valence of responses to question 7 (“Whether I trust the police and see them as authority, comes down to things like...”). This finding suggests that when prompted about trust in the police, the LLMs (especially ChatGPT) generated content with a distinctly more negative sentiment compared to the human respondents in this study's sample.

Table 2: Valence distribution Q7 (Q7, N=53)

Valence_Q7	Human	ChatGPT	Gemini	Total
Negative	13.3 %	100 %	25 %	21 %
Neutral	62.2 %		75 %	59 %
Positive	24.4 %			21 %
Total	100 %	100 %	100 %	
Chi-square	17.92 ***	*p* < .001		

The analysis sought to identify the key factors mentioned by both human participants and the AI models (GPT4 and Gemini) that influence trust and perception of police authority. Qualitative exploration provides depth and context to the quantitative sentiment analysis findings for Q7, which indicated a statistically significant difference in sentiment distribution. ChatGPT generated content with a distinctly more negative sentiment compared to human respondents. Gemini, however, had neutral answers. In stark contrast to human respondents, who provided a broader range of sentiments, most respondents were neutral. This finding suggests that AI models, specifically GPT-4 as configured and prompted in this question, articulate views on police and authority that are quantitatively more negative than those expressed by the human sample (Table 2).

For the other questions (Q5, Q6, Q8) while there were differences in the percentage distributions, particularly the limited range of sentiments produced by the AI models (Tables 3, 4, and 5). For example, for Q5, there was no positive sentiment, and for Q8, there was no negative sentiment from Gemini. These differences were not statistically significant when considering the overall distribution, due to the small sample size of responses from ChatGPT and Gemini compared to the human responses. This small sample size makes it harder to detect statistically significant differences unless those differences are large.

Table 3: Valence distribution Q5 (Q5, N=53)

Valence_Q5	Human	ChatGPT	Gemini	Total (N=53)
Negative	40 %	50 %	50 %	41.5 %
Neutral	51 %	50 %	50 %	50.9 %
Positive	8.9 %			7.5 %
Total	100 %	100 %	100 %	
Chi-square	.88 ***	*p* .928		

Table 4: Valence distribution Q6 (Q6, N=53)

Valence_Q6	Human (N=45)	ChatGPT (N=4)	Gemini (N=4)	Total (N=53)
Negative	24.4 %	50 %	75 %W	30.2 %
Neutral	35.6 %	25 %	25 %	34 %
Positive	40 %	25 %		35.8 %
Total	100 %	100 %	100 %	
Chi-square	5.67 ***	*p* .225		

Table 4: Valence distribution Q8 (Q8, N=53)

Valence_Q8	Human (N=45)	ChatGPT (N=4)	Gemini (N=4)	Total (N=53)
Negative	26.7 %			22.6 %
Neutral	48.9 %	100 %	75 %	54.7 %

Positive	24.4 %		25 %	22.6 %
Total	100 %	100 %	100 %	
Chi-square	5.19 ***	*p* .268		

4.1.2 Influence of Gender and Age on Attitudes Towards Police

A one-way ANOVA was conducted to assess the influence of key demographic characteristics, specifically age, on respondents' attitudes towards the police. These analyses were crucial for exploring how individuals' social backgrounds, as conceptualised within the theoretical framework, might mediate their perceptions of law enforcement. This one-way ANOVA revealed that age did not significantly influence attitudes towards the police. The analysis showed no statistically significant effect, $F(8, 36) = 0.42$, $p = 0.903$. The effect size measured by eta-squared was $\eta 2 = 0.08$. An eta-squared value between 0 and 0.10 is typically considered no or a weak association. This suggests that age accounts for approximately 8% of the variance in attitudes towards police and does not have a statistically significant effect. An independent sample T-test was conducted even though the male group contains fewer than 30 observations, and therefore, there is no clarity whether this variable is normally distributed compared to the variable used to compare the average valence scores between male and female participants. The mean valence score for males ($M = 0.90$, $SD = 0.36$, $n = 11$) and females ($M = 1.03$, $SD = 0.47$, $n = 33$) were compared. Levene's Test for Equality of variances indicated that the variances were not significantly different ($F = 0.19$, $p = 0.662$); therefore, equal variances were assumed for the T-test. The difference between the mean valence score of males and females did not differ significantly, $t(42) = -0.785$; $p = 0.437$, with the 95% confidence interval for the mean difference being [-0.43268, 0.19026]. The effect size, Cohen's d, was -0.273, 95% CI [-0.957, 0.413]. While females in the sample have a slightly higher valence score, this T-Test indicates that there is no statistically significant difference in sentiment between male and female participants.

4.1.3 Impact of Overall Media Consumption on Sentiment Towards Police

A simple linear regression was conducted to delve into the quantitative relationship between human participants' overall self-reported media consumption and the average valence (sentiment) of their open-ended responses concerning attitudes towards the police. This analysis explores whether average media consumption predicts the overall sentiment of human participants' responses. The regression model was found to be statistically significant,

F (1, 43) = 5.01, p = .032. This indicates that the average media consumption predicts that sentiment towards police officers in this sample is positively influenced if the respondents frequently watch online news websites, b^* = .32, p = 0.032.

Another simple linear regression was conducted with the average valence as the dependent variable and the self-reported time spent watching online news websites by the respondents as the independent variable. This analysis explores whether the amount of time human respondents spend watching online news websites predicts the overall sentiment of their responses. The regression model was found to be statistically significant, F (1,43) = 4.94, p =.032. This indicates that the average amount of time spent watching online news websites predicts a positive sentiment of responses in this sample.

Additionally, the average valence as the dependent variable and the police content the respondents reported on watching as the independent variable, a simple linear regression was conducted. This analysis explores whether the average amount of police related content the respondents watch predicts the overall sentiment of human participants' responses. The regression model was found not to be statistically significant, F (1,43) = 0.20, p = .89. This indicates that the average amount of police content does not predict the sentiment of responses in this sample.

4.2 Thematic analysis

Thematic analysis is a method for identifying, analysing and interpreting patterns of meaning referred to as 'themes' within qualitative data. It offers accessible and systematic procedures for generating codes, which are the smallest units of analysis capturing interesting features, and themes, which are larger patterns of meaning underpinned by a central organising concept. The aim is to identify and interpret key features of the data, guided by the research question. This approach is flexible and can be used to identify patterns within and across data about participants' views and perspectives. Given the explanatory nature of understanding the factors mentioned by participants and AI, an inductive (data-driven) approach was adopted.

4.2.1 Perceptions of police in the neighbourhood.

The first open question elicited responses reflecting participants' immediate thoughts on the safety and crime situation in their local areas and the perceived role of the police. The key themes identified included a description of varying levels of crime; many of the respondents feel there is minimal crime. One participant notes, "In my neighbourhood, there is almost no crime. If something does happen, the police are involved". Another participant states, "Low. I live in a quiet neighbourhood, where generally not much happens". Others highlight the presence of specific crimes, including "burglaries", "drugs", "youth causing nuisance. The perceived crime levels directly influence the participants' feelings of safety in their neighbourhood, as one participant notes that in their neighbourhood, it is quiet, and they feel safe there. Conversely, the presence of crime contributes to feelings of unsafety, one respondent states, "An unsafe feeling". Some express a "double feeling where safety and not feeling safe come together", explaining, "I feel safe because the police are present, but this comes because I feel unsafe by for instance the burglaries or nuisance in the neighbourhood". Participants connect their observations of the police presence and capability to the crime situation to their feelings of safety. Some participants, however, note a lack of police presence and associate this with low crime. One participant states, "police are fortunately not seen much in our neighbourhood. Others report that seeing the police often makes them feel unsafe, suggesting that police presence is a response to existing problems. Concerns are frequently raised about the police's capacity and timeliness, with phrases such as "I don't think the police will come in time" or "the police are not up for the task".

Regarding the police and crime in the neighbourhood, with the AI-generated text, a prominent theme is the complex awareness of police presence and crime. The AI-responses convey a balance or contradiction, noting that patrols offer reassurance while also highlighting ongoing issues like theft, vandalism, or noise complaints, which police presence sometimes seems merely to underscore rather than prevent. There is a recurring suggestion that focusing solely on enforcement or reaction after crime may not be the most effective approach, with ideas emerging around the need for greater community engagement, prevention programs, or community-oriented policing to address underlying issues. The AI response also reflects an awareness of crime experienced or heard about informally (word of mouth) rather than just through official channels or news, adding a layer of lived uncertainty alongside visible policing efforts.

4.2.2 Perceived police communication with civilians

The second open question covers perceptions of how the police interact and communicate with the public. A significant theme in the answers to this question was the tone and attitude of communication. Where some of the participants perceive police officers as having an "authoritative attitude towards citizens". One respondent mentions the police displaying "haantjes gedrag" (cocky behavior). In contrast, positive communication is also notes, such as; "good and calm way of addressing the citizens by the police. The importance of communicating "direct and clear" is also mentioned. Several responses highlight the variability in communication depending on the individual officer; "It differs per officer, some can be real jerks, but there are also many who are human". Some respondents recognized the difficulties inherent to police-civilian communication, including public behavior and the police's need for transparency. One respondent notes: "people always have something to complain about, or they think they have the 'right to information' ". Participants also mention some of the various ways police communicate with the public. A contrast is drawn with past communication methods, as participants mention the loss of a neighbourhood police officer who would have visiting hours, to now communicate mostly via various social media platforms.

On the topic of police communication with civilians, the central theme according to AI models is communication as a foundation for trust and safety. The AI responses strongly emphasise the crucial impact of the quality of communication on public safety, de-escalation, and relationship-building. Effective communication is consistently described as respectful, clear, calm, empathetic, and involving active listening, which helps people feel heard and builds rapport. Conversely, communication that is perceived as distant, aggressive, dismissive, one-sided, or unclear is seen as creating distance, fear, escalating tensions, and undermining trust. The responses also acknowledge that communication experiences can vary significantly for different individuals and communities, highlighting the challenges of potential bias or pre-existing mistrust in interactions.

4.2.3 Perceptions of police trust and authority.

The third open question delved into the criteria participants use to evaluate police trustworthiness and legitimacy. Responses cover police conduct, competence, fairness, and the nature of interactions. Participants frequently cited police conduct and behaviour as crucial, including whether police are respectful, listen, show empathy, de-escalate situations, and handle interactions fairly, particularly noting concern about the treatment of minorities or protestors. Police effectiveness in creating safety, being present when needed, addressing importantneighbourhoodd issues, and solving crimes were also significant factors. Personal experiences, as well as experiences shared by family and friends, played a large role in shaping trust. The professionalism of individual officers, encompassing their appearance, language, attitude, competence, honesty, and ability to project authority, was also mentioned as influential. Concerns about internal issues within the police, such as perceived misconduct, discrimination, or corruption, were noted as impacting trust. The uniform itself was seen by some respondents as symbolic of the authority that they hoped the police would uphold.

Concerning trust in police as authority, the AI models converge on the theme of earned authority and conditional trust. They explicitly state that trust is not automatic but must be built over time through consistent actions. Key factors identified as fostering trust include fair treatment, respect for individuals' rights, transparency, accountability for actions, consistent adherence to legal procedures, and visible commitment to serving all community members without bias. Conversely, experiences of discrimination, bias, excessive force, misconduct, unequal treatment, or lack of transparency and accountability are highlighted as actions that significantly shake or erode this trust. The AI models both suggest that authority is earned through integrity and behaviour, not simply held under the uniform.

In contrast to AI, human responses to Q7 displayed greater variability and often incorporated personal perspectives and emotional language. While human participants are also citing factors related to police conduct, their accounts were frequently grounded in direct experience and personal interaction. Many responses referenced "my personal experience", including both positive and negative encounters, and the experiences of family, friends, and colleagues. The specific way an officer approaches an individual, including communication style and listening skills, was frequently highlighted as influential. Participants also focused on the police's observable behaviour and conduct, mentioning aspects like communication quality ("taalgebruik", "Communicatie"), the ability to listen and respond appropriately ("luisteren en empathisch vermogen", "kunnen luisteren"), effective de-escalation, and overall professionalism and competence. Beyond individual interactions, human responses also revealed an awareness of systemic issues and accountability within the police force, touching upon concerns like racism, discrimination, misconduct ("wangedrag") and perceived lack of transparency in handling of reports. Respondents expressed an expectation that these broader issues should be addressed. A key distinction in human data was the explicit declaration of personal trust or respect for the police. Unlike the AI's descriptive criteria, human responses included direct statements saying they have trust in the police; some participants also articulated respect for the police uniform itself, contingent upon the officer's behaviour.

4.2.4 Perceptions of media portrayals of the police.

The final open-ended question measured how the media portrays the police resonates with the respondent's perspective. There was a common pattern that media portrayal is often perceived as negative, focusing on police errors, use of force, lack of control, or portraying them as victims. While some agreed with this negative portrayal, feeling it aligned with their view of police inadequacies or power abuse, many others disagreed, feeling the police are portrayed too negatively or inaccurately. They felt the media did not fully capture the challenges the police face, such as aggression against officers, high workloads, or staff shortages. One participant noted that the image by the police is affected/damaged by reality TV series like *Bureau Maastricht* or *Bureau Rotterdam*, according to the respondent they show less authority and are everyone's friend: "De politie wordt afgespiegeld als een autoriteit en dat moet zeker zo blijven. Zij moeten ook vertrouwen blijven uitstralen. Dat is goed. Zij gebruiken ook de media om belangrijke berichten aan de mensen te geven. Deze communicatie is essentieel in sommige gevallen. Alleen is mijn beeld van de politie door tv series zoals bureau Maastricht of Rotterdam aangetast omdat zij daar minder gezag tonen. Politie de allemansvriend. Dat hoeft voor mij niet" ["The police are portrayed as an authority, and that should certainly remain so. They should also continue to build trust. That is good.

They also use the media to give important messages to people. This communication is essential in some cases. However, my image of the police has been damaged by TV series like Bureau Maastricht or Bureau Rotterdam because they show less authority there. Police, everyone's friend. That is not necessary for me. Some respondents acknowledged that the portrayal can be mixed or dependent on context. A positive portrayal depicting police as heroes or helpful also exists in the media, in which some participants resonate with their view of police as hard-working individuals undertaking a difficult but noble profession.

When discussing the media portrayal of the police, a significant theme is the polarised media reflection of complex realities. The AI models observe that the media often presents a dual or extreme narrative, swinging between heroism and public safety efforts on one hand, and exposing instances of misconduct, systemic issues, or controversy on the other. This polarised portrayal is seen as resonating with and amplifying the diverse and sometimes conflicting real-life experiences and existing societal attitudes towards the police. The AI models suggest that media frames policing directly influences public perception, trust levels, and broader discussions about justice, accountability, and the legitimacy of law enforcement, depending on which aspects are emphasised and how they align with personal or community narratives.

Responses from GPT and Gemini provided structures, lists or paragraphs detailing factors influencing trust. The themes identified consistently focused on the attributes and actions of police as the foundation of trust and authority. A prominent theme centred on the ethical conduct and fairness of the police, emphasising the importance of equitable treatment, unbiased law enforcement, and respectful interactions regardless of an individual's background or status. Closely related was the theme of accountability and transparency, highlighting the necessity for police to be held responsible for their actions and mistakes, and for their operations and complaint handling processes to be open to scrutiny. The absence of transparency was explicitly noted as detrimental to trust. Furthermore, the AI models frequently mentioned the significance of community interaction and engagement, suggesting that police should listen to residents, understand the communities they serve, and actively work to build connections. Underlying these specific factors was the overarching idea that trust is earned, not inherent, built gradually through consistent ethical behaviour and potentially eroded by actions such as unequal treatment or excessive force. The AI-responses tended to articulate these points by detailing the criteria or conditions upon which trust and authority are based, presenting a set of objective standards for police conduct. This descriptive, criteria-focused nature aligns with the finding that all AI responses for Q7 were categorised as neutral in sentiment analysis.

5. Conclusion

This research set out to explore and compare the perceptions of police officers as expressed by human participants and advanced LLMs, specifically GPT-4 and Gemini. Utilising a mixed-methods approach, the study combined quantitative sentiment analysis with qualitative thematic analysis to provide a nuanced understanding of how these different sources frame their understanding of police and related concepts like trust and authority.

The core research question sought to determine if significant differences in sentiment and thematic content concerning police officers between human respondents and AI-generated responses. While the quantitative analysis revealed that the overall sentiment distribution across all prompts (Q5-Q8), when aggregated, did not show a statistically significant difference between human participants, GPT-4, and Gemini, a critical distinction emerged when examining specific dimensions of police perceptions. Notably, for question 7, which explicitly explored perceived trust and authority, a statistically significant difference in sentiment was observed. When prompted to trust in the police, GPT-4 generated content with a distinctly more negative sentiment compared to the human respondents in the study's sample. In contrast, Gemini produced neutral answers for this question, and human respondents, while offering a broader range of sentiments, were predominantly neutral. The sentiment analysis involved manually categorising responses into Negative, Neutral, and Positive valence categories. However, the study identified significant correlations between the data of human participants. A positive correlation was found between average sentiment and overall media consumption, as well as time spent watching online news websites.

The qualitative thematic analysis complemented these quantitative findings by elucidating the 'why' behind these differences. For instance, the AI-generated responses for Q7 offered structured, criteria-focused descriptions of factors influencing trust, such as ethical conduct, fairness, accountability, transparency, and community engagement, consistently presenting trust as something "earned, not inherent". Conversely, human responses to Q7 displayed greater variability, frequently incorporating personal perspectives, emotional language, and direct experiences with police, including positive and negative encounters. They also highlighted specific officer behaviours like communication style and listening skills, and demonstrated an awareness of systemic issues like racism and misconduct. This finding, where AI responses lean towards lower (more negative or neutral) average valence scores, is particularly significant. If AI-generated content, with its capacity for “mass producing goods and cheap semantic artefacts”, with its universal use, this embedded negative sentiment could contribute to the cultivation of a more critical or distrustful public perception of police. This suggests that AI, trained on vast datasets that reflect existing media and societal content, may inadvertently cultivate specific representations of police officers, potentially reinforcing or altering existing attitudes towards police officers.

The findings of this research carry significant theoretical implications, particularly considering cultivation theory and social location theory. Cultivation theory posits that prolonged exposure to media content shapes an individual's perception of reality. In the context of this research, LLMs can be considered powerful new “media” forms, absorbing and representing vast datasets. The observed generated negative sentiment from ChatGPT regarding police trust for Q7 suggests that these models may not merely reflect a neutral aggregation of data but could be reproducing and potentially intensifying existing societal narratives or biases prevalent in their immense training datasets. This outcome aligns with the understanding that LLMs, by design, encapsulate patterns from the data they are trained on, which inherently include various forms of bias. The marked difference from human responses in this study’s sample points to a potential “digital cultivation” effect, where AI-generated content, if widely consumed, might contribute to or even reinforce public perceptions, even if those perceptions are from a specific human demographic’s general sentiment. The study by Mustafaj and Van den Bulck (2021) further supports this by showing how broader societal context can influence broader societal contexts can influence cultivation processes, underscoring the importance of considering the ‘training context’ of AI models as a digital proxy for societal influence (p.731).

The divergence between human and AI sentiment highlights the relevance of social location theory. The human participants in this study, primarily female respondents, primarily white with more than 40% reporting to have an academic background, represent a distinct social location, whose backgrounds and experiences would shape their responses. In contrast, LLMs lack a personal ‘social location’; instead, they draw from an immense, heterogeneous dataset representing a vast range of perspectives and biases present across the internet and other sources. The varying sentiment suggest that the AI models might be reflecting a more aggregated, or perhaps, more critical, perspective on police trust than the specific human sample examined, or even a particular 'tone and perspective' with 'underlying biases and prejudices' embedded in the language they use, as identified in cognitive bias analysis for text (Wen & Younes, 2024, p.5-7). The thematic analysis provides a rich context for understanding why the AI models' sentiment regarding police trust diverged from human responses, reflecting a potentially more aggregated or critical perspective derived from their training data. The tendency of AI models, like GPT-4 and Gemini, to sometimes "overestimate" bias in other contexts, interpreting content as biased where humans see it as neutral, due to exposure to bias during training, supports the idea that their generated sentiment can reflect these embedded patterns. In conclusion, this research explores how LLM-generated texts, particularly generated by GPT-4, can exhibit a negative sentiment regarding police perception, contrasting with human responses in this sample. This outcome reaffirms the necessity for rigorous, ongoing evaluation of AI language models to mitigate the reproduction and amplification of biases present in their training data, especially as these models are becoming increasingly integrated into public information dissemination.

However, the findings also highlight the complexities and conditions under which cultivation effects manifest, aligning with insights from reception research. While overall social media consumption was found to be significantly related to the fear of crime in one study (Intrivia et al, 2017, p.166), the lack of uniform statistical significance across all media types in prior research, particularly for specific social media news consumption and traditional formats like TV crime dramas, point to the mediating role of audience characteristics and content specifics (Glascock, 2023, p580). Social location theory further illuminates these conditional effects, proposing that individuals construct meaning from media based on their social position, demographics, and personal experiences.

Despite the valuable insights, this study has several limitations that warrant consideration for future research. A primary limitation stems from the sampling method employed to gather human respondents. This method involves sharing the questionnaire link through personal networks and is generally discouraged in academic research due to its high likelihood of biased results and non-representativeness. This small and potentially biased human sample size also meant that for several questions, the differences in sentiment distributions between humans and AI models, though present, were not statistically significant. The analysis of the AI study found that the impact of media consumption on attitudes towards police is nuanced, with the effect potentially being conditional on the viewer's gender and personal experiences, and a lack of significant findings for other media types.

Future studies should aim for larger and more diverse human samples, utilising more robust sampling methods to ensure greater representativeness and reduce bias, potentially incorporating social location theory more directly into the recruitment and analysis.

Discussion

This increased engagement with online news websites and more positive sentiment towards police officers presents an intriguing contrast to much of the existing research in this field. This outcome stands in contrast to a significant body of research, predominantly conducted within American contexts, which often suggests a negative relationship between online news consumption and perceptions of police legitimacy. For instance, Intravia et al. (2017) found that individuals who spent more time reading news online had significantly fewer positive attitudes about police legitimacy (p.163). Similarly, Glascock (2023) noted that exposure to liberal media outlets, including online news sources like *CNN*.com and *NYtimes*.com, was negatively associated with attitudes towards the police (p 79-581). Previous research also indicated that news media, particularly network news, tend to focus on negative news like police misconduct, potentially cultivating frequent misconduct among heavy viewers. The dataset that was used to conduct this research consists of European respondents. Herman and Eisand (2020) conducted a study with a German student sample, which found different cultivation effects for Facebook use compared to typical US findings (p.1128-1134). It is plausible that the nature, content, and framing of online news websites in Europe may differ from their American counterparts. European online news, particularly, may tend to focus more on factual reporting of emergencies of everyday police work, which could foster more positive perceptions among viewers.

This research embarked on an exploratory study, aiming to answer the central research question: how do large language models (LLMs) and human respondents differ in their articulation of perceptions regarding police officers? Drawing on cultivation theory and social location theory, this research provides critical insights into the nuanced ways AI models like GPT-4 and Gemini generate text about police officers compared to human perspectives, highlighting potential areas where AI may reflect or even amplify societal biases. Through a mixed-methods approach, combining quantitative sentiment analysis and qualitative thematic analysis, this research offers a nuanced understanding of these dynamics. This descriptive, criteria-focused nature, while articulating objective standards, co-exists with quantitatively observed significance negatively from ChatGPT for this specific question, suggesting that while the AI models might articulate what builds or erodes trust, their overall sentiment on the topic can be negative.

References

Anthias, F. (2012). Hierarchies of social location, class and intersectionality: Towards a translocational frame. *International Sociology*, *28*(1), 121–138. https://doi.org/10.1177/0268580912463155

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional accuracy Disparities in commercial gender classification. *Proceedings of Machine Learning Research*, *81*, 1–15. https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

Braun, V., & Clarke, V. (2016). Thematic analysis. *The Journal of Positive Psychology*, *12*(3), 297–298. https://doi.org/10.1080/17439760.2016.1262613

Dowler, K., & Zawilski, V. (2007). Public perceptions of police misconduct and discrimination: Examining the impact of media consumption. *Journal of Criminal Justice*, *35*(2), 193–203. https://doi.org/10.1016/j.jcrimjus.2007.01.006

Doyle, A. (2018). "COPS": Television Policing as policing Reality. In *Routledge eBooks* (pp. 95–116). https://doi.org/10.4324/9781351328203-6

Fiske, J. (1986). Television: Polysemy and popularity. *Critical Studies in Mass Communication*, *3*(4), 391–408. https://doi.org/10.1080/15295038609366672

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its Nature, scope, Limits, and Consequences. *Minds and Machines*, *30*(4), 681–694. https://doi.org/10.1007/s11023-020-09548-1

García-Castro, J. D., & Pérez-Sánchez, R. (2018). Fear of crime and cultivation effects. social and psychological predictors. *Universitas Psychologica*, *17*(3), 7-20.

Gerbner, G. (1998). Cultivation Analysis: An Overview. *Mass Communication & Society*, *1*(3–4), 175–194. https://doi.org/10.1080/15205436.1998.9677855

Glascock, J. (2023). Relationship of media usage to attitudes toward police. *Communication Quarterly*, *71*(5), 568–587. https://doi.org/10.1080/01463373.2023.2239314

Hermann, E., Eisend, M., & Bayón, T. (2020). Facebook and the cultivation of ethnic diversity perceptions and attitudes. *Internet Research*, *30*(4), 1123–1141. https://doi.org/10.1108/intr-10-2019-0423

Intravia, J., Wolff, K. T., & Piquero, A. R. (2017). Investigating the effects of media consumption on attitudes toward police legitimacy. *Deviant Behavior*, *39*(8), 963–980. https://doi.org/10.1080/01639625.2017.1343038

Intravia, J., Thompson, A. J., & Pickett, J. T. (2020). Net legitimacy: internet and social media exposure and attitudes toward the police. *Sociological Spectrum*, *40*(1), 58–80. https://doi.org/10.1080/02732173.2020.1720554

Katz, Y., & Shifman, L. (2017). Making sense? The structure and meanings of digital memetic nonsense. *Information Communication & Society*, *20*(6), 825–842. https://doi.org/10.1080/1369118x.2017.1291702

Keil, S. I., & Gabriel, O. W. (2012). *Society and democracy in Europe*. https://doi.org/10.4324/9780203079812

Kirk, H. R., Jun, Y., Volpin, F., Iqbal, H., Benussi, E., Dreyer, F., Shtedritski, A., & Asano, Y. (2021, December 6). *Bias Out-of-the-Box: An empirical analysis of intersectional occupational biases in popular generative language models*.

https://proceedings.neurips.cc/paper/2021/hash/1531beb762df4029513ebf9295e0d34f-Abstract.html

Kurita, K., Vyas, N., Pareek, A., Black, A. W., & Tsvetkov, Y. (2019, June 18). *Measuring bias in contextualized word representations*. arXiv.org. https://arxiv.org/abs/1906.07337

Mustafaj, M., & Van Den Bulck, J. (2021, January 14). *Revisiting Cultivation as a gravitational Process: A Cross-National comparison of the cultivation of fear and Mistrust*. Mustafaj | International Journal of Communication. https://ijoc.org/index.php/ijoc/article/view/15230

Nadeem, M., Bethke, A., & Reddy, S. (2020, April 20). *StereoSet: Measuring stereotypical bias in pretrained language models*. arXiv.org. https://arxiv.org/abs/2004.09456

Navigli, R., Conia, S., & Ross, B. (2023). Biases in large language models: origins, inventory, and discussion. *Journal of Data and Information Quality*, *15*(2), 1–21. https://doi.org/10.1145/3597307

Sheng, E., Chang, K., Natarajan, P., & Peng, N. (2019, September 3). *The woman worked as a babysitter: on biases in Language Generation*. arXiv.org. https://arxiv.org/abs/1909.01326

Succar, R., Ramallo, S., Das, R., Ventura, R. B., & Porfiri, M. (2024). Understanding the role of media in the formation of public sentiment towards the police. *Communications Psychology*, *2*(1). https://doi.org/10.1038/s44271-024-00059-8

Soulliere, D. M. (2004). Policing on prime-time: A comparison of television and real-world policing. *American Journal of Criminal Justice*, *28*(2), 215–233. https://doi.org/10.1007/bf02885873

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is All you Need*. https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

Wen, Z., & Younes, R. (2023). ChatGPT v.s. media bias: A comparative study of GPT-3.5 and fine-tuned language models. *Applied and Computational Engineering*, *21*(1), 249–257. https://doi.org/10.54254/2755-2721/21/20231153

APPENDIX A: Survey – Questionnaire

Start of Block: Introduction & Informed Consent

Q1 Dear participant,

Q2 Thank you for taking the time to participate in this survey! To make sure you are most comfortable you have the ability to either take the survey Dutch/English, you can change the language in the right top corner. My name is Mieke Dohmen and I would like to invite you to participate in the research process of my MA Thesis. I am a student at Erasmus University of Rotterdam currently completing my Master's degree in Media, Digitalisation & Social Impact, this research is aimed at studying and comparing language generated by humans and AI generated text regarding police officers. This survey consists of a total of 4 open-ended scenarios which need to be completed by the participant. You are requested to complete these prompts by adding sentences/words. This means that the added text has to be a consistent continuation of the original scenario. There are no right or wrong answers and you are free to add whatever continuation you want; the responses are meant to be subjective and therefore instinctiveness is appreciated. It is requested that you complete the statements as truthfully as possible in relation to your personal values and ideologies. Respondent's anonymity is guaranteed, and the use of the collected data is strictly limited to this research's purposes and therefore will be securely stored. This entails that this data will not be shared with any other third parties. The survey takes approximately 10 minutes to complete. If at any point you feel like you want to stop participating, you are of course free to do so. The following question will ask you whether you have understood that your statements will be anonymous and treated confidentially. In addition, it is also meant to inform whether you consent to the collection and processing of the data for the purpose of this research. Thank you again for your participation! :)

Page Break

Q3 By clicking "I agree" below, you indicate that you have read and understood the information provided, and you agree to participate in this study. By signing this form, I: 1. Consent to participate in this research; 2. Consent to the use of my personal data; 3. Confirm that I am at least 18 years old; 4. Confirm that I understand that participating in this research is completely voluntary and that I can stop at any time; and 5. Confirm that I understand that my data will be anonymised for publication, educational purposes and further research.

- o I agree (1)
- o I do not agree (2)

Skip To: End of Survey If Q3 = I do not agree

End of Block: Introduction & Informed Consent

Start of Block: Intrasentence CAT

Q4 This section consist of 4 scenarios which I request you to complete. There are no right or wrong anwers and you are free to add whatever continuation you feel is most appropriate. You can complete the scenarios as truthfully as possible in relation to your personal values and ideologies.

Page Break

Q5 Thinking about police and crime in my neighbourhood, what often comes to mind is....
Please complete the scenario below.

____

____

____

____

____

Page Break

Q6 When I hear people talk about police's communication with civilians. I think of... *Please complete the scenario below.*

____

____

____

____

____

Page Break

Q7 Whether I trust the police and see them as authority, comes down to things like... *Please complete the scenario below.*

____

____

____

____

____

Page Break

Q8 The portrayal of police officers in media which is.... , resonates with... *Please complete the scenario below.*

____

____

____

End of Block: Intrasentence CAT

Start of Block: Media Consumption & Crime related

Q9 Most individuals have limited personal experience with crime or the criminal justice system, popular media often serves as a primary source of information about these topics. To help investigate this relationship, in the following 3 questions you will be asked about your media consumption.

Page Break

Q10 In a typical week, how much time do you spend watching:

	Never (1)	Less than once a week (2)	About once a week (3)	Several times a week (4)	Every day or almost every day (5)
Local news (1)	o	o	o	o	o
National news (2)	o	o	o	o	o
Social media (3)	o	o	o	o	o
Online news websites (4)	o	o	o	o	o
Newspaper (print/online) (5)	o	o	o	o	o
Entertainment (such as Netflix, HBO, Amazon prime, Disney+) (6)	o	o	o	o	o

Page Break

Q11 How often does the media you watch cover crime, policing or the criminal justice system?

	Never (1)	Less than once a week (2)	About once a week (3)	Several times a week (4)	Every day or almost every day (5)
Local news (1)	o	o	o	o	o
National news (2)	o	o	o	o	o
Social media (3)	o	o	o	o	o
Online news websites (4)	o	o	o	o	o
Newspaper (print/online) (5)	o	o	o	o	o
Entertainment (such as Netflix, HBO, Amazon prime, Disney+) (6)	o	o	o	o	o

Page Break

Q12 When consuming media, how often would you say you see content about the following topics:

	Never (1)	Sometimes (2)	Often (3)
Police misconduct (such as excessive	o	o	o

force, corruption etc) (1)			
Police discrimination (such as treating certain groups different based on race/wealth) (2)	o	o	o
Police effectively preventing or solving crimes (3)	o	o	o
Police treating citizens fairly and with respect (4)	o	o	o

End of Block: Media Consumption & Crime related

Start of Block: Direct Contact with police officer(s)

Q13 Have you been in direct contact with the police in the past 5 years?

o No (1)

o Yes (2)

Page Break

Display this question:

If Q13 = Yes

Q14 Please indicate your satisfaction of that interaction with the officer(s)

	Dissatisfied	Neutral	Satisfied
()			

End of Block: Direct Contact with police officer(s)

Start of Block: Demographics

Q15 The following questions will ask you about your age, gender, education level, nationality, religion and employment status. This information helps to understand the context of your media consumption patterns in relation to broader societal factors and contributes to the analysis of the survey results.

Q16 Please indicate your age

____

Page Break

Q17 Please indicate your gender

- o Male (1)
- o Female (2)
- o Non-binary / third gender (3)
- o Prefer not to say (4)

Page Break

Q18 Please indicate your education level

- o Primary (1)
- o High School (2)
- o Secondary (3)
- o Bachelor (4)
- o Master (5)
- o Doctoral (6)

Page Break

Q19 Please indicate your nationality

▼ Afghanistan (1) ... Zimbabwe (193)

Page Break

Q20 Please indicate your religion

- o Christianity (1)
- o Islam (2)
- o Hinduism (3)
- o Buddhism (4)
- o Judaism (5)
- o Sikhism (6)
- o Traditional / Indigenous beliefs (7)
- o Atheist / No religion (8)
- o Other : (9) ____
- o Prefer not to say (10)

Page Break

Q21 What is your current employment status?

- o Employed full-time (1)
- o Employed part-time (2)
- o Self-employed (3)
- o Unexployed (4)
- o Retired (5)
- o Student (6)
- o Homemaker (7)
- o Unable to work (8)

End of Block: Demographics

Start of Block: Block 9

APPENDIX A1: SPSS – Output

Frequencies – Demographic data

Statistics

		Please indicate your gender	Please indicate your education level	List of Countries	Please indicate your religion - Selected Choice	What is your current employment status?
N	Valid	45	45	45	45	45
	Missing	32	32	32	32	32
Mean		1,80	3,60	116,38	5,67	2,51
Median		2,00	4,00	121,00	8,00	2,00
Std. Deviation		,548	1,009	21,675	3,674	2,018
Variance		,300	1,018	469,786	13,500	4,074
Minimum		1	2	17	1	1
Maximum		4	5	121	10	8
Percentiles	25	1,50	3,00	121,00	1,00	1,00
	50	2,00	4,00	121,00	8,00	2,00
	75	2,00	4,00	121,00	8,00	3,00

Please indicate your gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	11	14,3	24,4	24,4
	Female	33	42,9	73,3	97,8
	Prefer not to say	1	1,3	2,2	100,0
	Total	45	58,4	100,0	
Missing	System	32	41,6		
Total		77	100,0		

Please indicate your education level

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	High School	8	10,4	17,8	17,8
	Secondary	11	14,3	24,4	42,2
	Bachelor	17	22,1	37,8	80,0
	Master	9	11,7	20,0	100,0
	Total	45	58,4	100,0	
Missing	System	32	41,6		
Total		77	100,0		

List of Countries

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Belgium	2	2,6	4,4	4,4
	Netherlands	43	55,8	95,6	100,0
	Total	45	58,4	100,0	
Missing	System	32	41,6		
Total		77	100,0		

Please indicate your religion - Selected Choice

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Christianity	16	20,8	35,6	35,6
	Islam	1	1,3	2,2	37,8
	Atheist / No religion	18	23,4	40,0	77,8
	Other :	7	9,1	15,6	93,3
	Prefer not to say	3	3,9	6,7	100,0
	Total	45	58,4	100,0	
Missing	System	32	41,6		
Total		77	100,0		

What is your current employment status?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Employed full-time	17	22,1	37,8	37,8
	Employed part-time	16	20,8	35,6	73,3
	Self-employed	3	3,9	6,7	80,0
	Unexployed	1	1,3	2,2	82,2
	Retired	3	3,9	6,7	88,9
	Student	2	2,6	4,4	93,3
	Unable to work	3	3,9	6,7	100,0
	Total	45	58,4	100,0	
Missing	System	32	41,6		
Total		77	100,0		

Valence Aggregate – Crosstabulation (separate dataset to maintain categorical valence variable)

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Valence_aggregrate * source	212	100,0%	0	0,0%	212	100,0%

Valence_aggregrate * source Crosstabulation

			source			
			Human	ChatGPT	Gemini	Total
Valence_aggregrate	Negative	Count	47	8	6	61
		% within source	26,1%	50,0%	37,5%	28,8%
	Neutral	Count	89	7	9	105
		% within source	49,4%	43,8%	56,3%	49,5%
	Positive	Count	44	1	1	46
		% within source	24,4%	6,3%	6,3%	21,7%
Total		Count	180	16	16	212
		% within source	100,0%	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	7,772[a]	4	,100
Likelihood Ratio	8,654	4	,070
Linear-by-Linear Association	5,332	1	,021
N of Valid Cases	212		

a. 4 cells (44,4%) have expected count less than 5. The minimum expected count is 3,47.

Q5 – Valence, Crosstabulation

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Valence_Q5 * source	53	68,8%	24	31,2%	77	100,0%

Valence_Q5 * source Crosstabulation

			source			
			Human	ChatGPT	Gemini	Total
Valence_Q5	Negative	Count	18	2	2	22
		% within source	40,0%	50,0%	50,0%	41,5%
	Neutral	Count	23	2	2	27
		% within source	51,1%	50,0%	50,0%	50,9%
	Positive	Count	4	0	0	4
		% within source	8,9%	0,0%	0,0%	7,5%
Total		Count	45	4	4	53
		% within source	100,0%	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	,876[a]	4	,928
Likelihood Ratio	1,466	4	,833
Linear-by-Linear Association	,560	1	,454
N of Valid Cases	53		

a. 7 cells (77,8%) have expected count less than 5. The minimum expected count is ,30.

Q6 – Valence, Crosstabulation

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Valence_Q6 * source	53	68,8%	24	31,2%	77	100,0%

Valence_Q6 * source Crosstabulation

			source			
			Human	ChatGPT	Gemini	Total
Valence_Q6	Negative	Count	11	2	3	16
		% within source	24,4%	50,0%	75,0%	30,2%
	Neutral	Count	16	1	1	18
		% within source	35,6%	25,0%	25,0%	34,0%
	Positive	Count	18	1	0	19
		% within source	40,0%	25,0%	0,0%	35,8%
Total		Count	45	4	4	53
		% within source	100,0%	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	5,677[a]	4	,225
Likelihood Ratio	6,300	4	,178
Linear-by-Linear Association	5,094	1	,024
N of Valid Cases	53		

a. 6 cells (66,7%) have expected count less than 5. The minimum expected count is 1,21.

Q7 – Valence, Crosstabulation

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Valence_Q7 * source	53	68,8%	24	31,2%	77	100,0%

Valence_Q7 * source Crosstabulation

			source			
			Human	ChatGPT	Gemini	Total
Valence_Q7	Negative	Count	6	4	1	11
		% within source	13,3%	100,0%	25,0%	20,8%
	Neutral	Count	28	0	3	31
		% within source	62,2%	0,0%	75,0%	58,5%
	Positive	Count	11	0	0	11
		% within source	24,4%	0,0%	0,0%	20,8%
Total		Count	45	4	4	53
		% within source	100,0%	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	17,921[a]	4	,001
Likelihood Ratio	16,196	4	,003
Linear-by-Linear Association	4,923	1	,026
N of Valid Cases	53		

a. 6 cells (66,7%) have expected count less than 5. The minimum expected count is ,83.

Q8 – Valence, Crosstabulation

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Valence_Q8 * source	53	68,8%	24	31,2%	77	100,0%

Valence_Q8 * source Crosstabulation

			source			
			Human	ChatGPT	Gemini	Total
Valence_Q8	Negative	Count	12	0	0	12
		% within source	26,7%	0,0%	0,0%	22,6%
	Neutral	Count	22	4	3	29
		% within source	48,9%	100,0%	75,0%	54,7%
	Positive	Count	11	0	1	12
		% within source	24,4%	0,0%	25,0%	22,6%
Total		Count	45	4	4	53
		% within source	100,0%	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	5,193[a]	4	,268
Likelihood Ratio	7,571	4	,109
Linear-by-Linear Association	,501	1	,479
N of Valid Cases	53		

a. 6 cells (66,7%) have expected count less than 5. The minimum expected count is ,91.

ANOVA – Valence, Age

ANOVA

Age Group

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	822,739	8	102,842	,417	,903
Within Groups	8869,261	36	246,368		
Total	9692,000	44			

ANOVA Effect Sizes[a,b]

		Point Estimate	95% Confidence Interval	
			Lower	Upper
Age Group	Eta-squared	,085	,000	,084
	Epsilon-squared	-,118	-,222	-,120
	Omega-squared Fixed-effect	-,116	-,216	-,117
	Omega-squared Random-effect	-,013	-,023	-,013

a. Eta-squared and Epsilon-squared are estimated based on the fixed-effect model.

b. Negative but less biased estimates are retained, not rounded to zero.

Independent Samples T-Test - Valence average - Gender

Group Statistics

	Please indicate your gender	N	Mean	Std. Deviation	Std. Error Mean
valence_average	Male	11	,9091	,35834	,10804
	Female	33	1,0303	,46669	,08124

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
valence_average	Equal variances assumed	,194	,662	-,785	42	,218	,437	-,12121	,15434	-,43268	,19026
	Equal variances not assumed			-,897	22,279	,190	,379	-,12121	,13518	-,40136	,15893

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval	
				Lower	Upper
valence_average	Cohen's d	,44330	-,273	-,957	,413
	Hedges' correction	,45142	-,269	-,939	,406
	Glass's delta	,46669	-,260	-,943	,428

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control (i.e., the second) group.

ANOVA Effect Sizes[a,b]

		Point Estimate	95% Confidence Interval Lower	Upper
Please indicate your gender	Eta-squared	,513	,149	,592
	Epsilon-squared	,404	-,040	,501
	Omega-squared Fixed-effect	,399	-,039	,495
	Omega-squared Random-effect	,077	-,005	,109

a. Eta-squared and Epsilon-squared are estimated based on the fixed-effect model.

b. Negative but less biased estimates are retained, not rounded to zero.

Simple Regression analysis – Media average

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics: R Square Change	F Change	df1	df2	Sig. F Change
1	,323[a]	,104	,083	,43118	,104	5,006	1	43	,030

a. Predictors: (Constant), media_average

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,931	1	,931	5,006	,030[b]
	Residual	7,994	43	,186		
	Total	8,925	44			

a. Dependent Variable: valence_average

b. Predictors: (Constant), media_average

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	,261	,329		,793	,432
	media_average	,207	,093	,323	2,237	,030

a. Dependent Variable: valence_average

Simple Regression analysis – Q10_5" In a typical week, how much time do you spend watching: - Newspaper (print/online)"

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics R Square Change	F Change	df1	df2	Sig. F Change
1	,321[a]	,103	,082	,43150	,103	4,935	1	43	,032

a. Predictors: (Constant), In a typical week, how much time do you spend watching: - Online news websites

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,919	1	,919	4,935	,032[b]
	Residual	8,006	43	,186		
	Total	8,925	44			

a. Dependent Variable: valence_average

b. Predictors: (Constant), In a typical week, how much time do you spend watching: - Online news websites

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	,597	,186		3,216	,002
	In a typical week, how much time do you spend watching: - Online news websites	,105	,047	,321	2,221	,032

a. Dependent Variable: valence_average

Simple Regression analysis – Police related content

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics R Square Change	F Change	df1	df2	Sig. F Change
1	,022[a]	,000	-,023	,45548	,000	,020	1	43	,888

a. Predictors: (Constant), policecontent_average

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	,949	,254		3,742	<,001
	policecontent_average	,010	,074	,022	,142	,888

a. Dependent Variable: valence_average

Declaration Page: Use of Generative AI Tools in Thesis

Student Information
Name: Mieke Dohmen
Student ID: 247715
Course Name: Master Thesis CM5050
Supervisor Name:
Date: 6-6-2025

Declaration:

Acknowledgment of Generative AI Tools
I acknowledge that I am aware of the existence and functionality of generative artificial intelligence (AI) tools, which are capable of producing content such as text, images, and other creative works autonomously.

GenAI use would include, but not limited to:

- Generated content (e.g., ChatGPT, Quillbot) limited strictly to content that is not assessed (e.g., thesis title).
- ~~Writing improvements, including~~ grammar and spelling corrections (e.g., Grammarly)
- Language translation (e.g., DeepL), without generative AI alterations/improvements.
- Research task assistance (e.g., finding survey scales, qualitative coding verification, debugging code)
- Using GenAI as a search engine tool to find academic articles or books (e.g.,

☒ I declare that I have used generative AI tools, specifically GPT-4 and Gemini, in the process of creating parts or components of my thesis. The purpose of using these tools was to aid in generating content or assisting with specific aspects of thesis work.

Extent of AI Usage
☒ I confirm that while I utilized generative AI tools to aid in content creation, the

☐ I declare that I have NOT used any generative AI tools and that the assignment concerned is my original work.

Signature: [digital signature]
Date of Signature: [Date of Submission]

majority of the intellectual effort, creative input, and decision-making involved in completing the thesis were undertaken by me. I have enclosed the prompts/logging of the GenAI tool use in an appendix.

Ethical and Academic Integrity

☒ I understand the ethical implications and academic integrity concerns related to the use of AI tools in coursework. I assure that the AI-generated content was used responsibly, and any content derived from these tools has been appropriately cited and attributed according to the guidelines provided by the instructor and the course. I have taken necessary steps to distinguish between my original work and the AI-generated contributions. Any direct quotations, paraphrased content, or other forms of AI-generated material have been properly referenced in accordance with academic conventions.

By signing this declaration, I affirm that this declaration is accurate and truthful. I take full responsibility for the integrity of my assignment and am prepared to discuss and explain the role of generative AI tools in my creative process if required by the instructor or the Examination Board. I further affirm that I have used generative AI tools in accordance with ethical standards and academic integrity expectations.

Signature: Mieke Dohmen
Date of Signature: 26-06-2025