**A state-of-the-art review on police accountability: What do we know from empirical studies?**

1. **Introduction**

Police officers are granted influential powers and possess a certain amount of discretion when making decisions during their daily activities. Since these actions can have far-reaching consequences, police accountability is essential (Kelling, Wasserman, & Williams, 1988; Lamboo, 2010; Millen & Stephens, 2011; Ransley, Anderson, & Prenzler, 2007).

Police accountability is a reflexive, complex and multifaceted phenomenon that is defined in multiple ways (Cheung, 2005; Manning, 2012; Walker & Archbold, 2014). A large amount of literature has been written on the subject, using different definitions. These definitions range from control over police and regulating policing activities to explaining and justifying certain behavior, as well as evaluations of those decisions (Busuioc, Curtin, & Groenleer, 2011; Chan, 1999; Eijkman, 2006; Findlay, 1994; Ransley et al., 2007). For the purposes of this review, we follow Chan’s (1999) definition of accountability: “*being answerable to audiences for performing up to certain prescribed standards, thereby fulfilling obligations, duties, expectations, and other charges… When people are accountable, they can be made to explain and justify their conduct, and their behaviour can be scrutinized, judged, and sanctioned by audiences*” (p. 253). Hence, accountability refers to being accountable for police conduct, to the oversight of policing activities and to the evaluation of those activities. Police officers need to account for certain (mis)behavior, either in the course of a misconduct investigation or in the course of their daily activities (not necessarily for misconduct). As such, complaints and investigative procedures, early warning systems, civilian oversight, independent commissions and many more topics were all (possibly) relevant for our review. Since authors relate the topic of police accountability to these subjects (De Angelis, 2009; De Angelis & Kupchik, 2007; McGregor, 2015; Schaible, De Angelis, Wolf, & Rosenthal, 2012; Walker & Archbold, 2014), the inclusion of these themes reduces the possibility of missing out on relevant publications.

Ponsaers (2010) remarks that there are many different research themes in regard to police accountability, such as police culture or evaluations of police. Muller (2002), Myhill (2007) and Walker (2007) argue that police accountability has regained more scholarly and policy attention. However, Mazeika et al. (2010) state that, despite a fairly constant attention for police accountability as a research subject, there has been a small decrease in attention regarding the topic in 2006-2007. Recently, the debate concerning police accountability mainly focusses on external oversight (Harris, 2013).

Despite the complexity of the subject, police accountability has some general characteristics. Cheung (2005) distinguishes four dimensions, of which the first two refer to the questions *who* is accountable and about *what*. Individual police officers are accountable for their own behavior, whereas the police organization is accountable for the performance of the department and the quality of the offered services (Ransley et al., 2007; Walker, 2007). Cheung’s third dimension refers *to whom* one must account. Multiple bodies may be able to impose accountability by monitoring and assessing police officers’ behavior (Busuioc et al., 2011). In this regard, a distinction is often made between internal and external accountability. Internal accountability is associated with organizational responsibility, supervision, performance evaluations, codes of conduct, the disciplinary system, loyalty and control, whereas external accountability refers to, among others, public hearings, annual reports, conferences, civilian oversight agencies and criminal and/or civil litigation (den Boer, 2002; Ransley et al., 2007; Walker, 2007). Fourth, Cheung asks *how* police accountability can be achieved. This accountability can be achieved in different ways, such as the use of body-worn cameras (Coudert, Butin, & Le Métayer, 2015; Harris, 2010) or early intervention systems (Harris, 2009). For this study, we chose to include all four dimensions.

Reviews on topics concerning the police have already been conducted, for example regarding plural policing and community-oriented policing (Boels & Verhage, 2016; Gill, Weisburd, Telep, Vitter, & Bennett, 2014). Although a *systematic* overview of empirical research on police accountability can complement these previous reviews, such a review is, to our knowledge, missing. Given the importance of this subject, we aim to provide such an overview in this article. This review was conducted from December 2015 till September 2016. First, the method will be discussed, after which the general characteristics of the literature as well as the methods used to study police accountability will be highlighted. Subsequently, the results of the thematic synthesis are briefly presented. The article ends with some concluding remarks and recommendations for future research.

1. **Method**

As we aimed to shed a light on the empirical research concerning police accountability, a scoping review, based on the principles of a systematic literature review, was conducted. A systematic literature review entails different steps (Verhage & Boels, 2015), which are presented in Figure 1. However, this review also entails characteristics of a narrative meta-review (Greenhalgh, Robert, Macfarlane, Bate, Kyriakidou, & Peacock, 2005; Wong, Greenhalgh, Westhorp, Buckingham, & Pawson, 2013). Initially, we aimed to answer multiple questions related to the topic of police accountability, but due to the relatively limited information in the publications, we were obliged to rephrase the research questions. Eventually, we decided to act on two general, broad questions, of which the first one is “What is police accountability?”. This question aims to answer what accountability is and which topics are (empirically) studied in this regard. The second research question was formulated as follows: “Which perceptions exist on police accountability?” (e.g. what do citizens, police officers and other actors regard as ‘accountability’?). The remainder of this article entirely focuses on the first research question.

Figure 1: Steps in a systematic literature review (Verhage & Boels, 2015)

In order to be selected for further analysis in the review, publications had to meet the following inclusion criteria: empirical research, published from 1990 onwards1, published in English or Dutch and have an association with one of the following topics: (a) internal or external accountability, (b) transparency and oversight or (c) integrity and police misconduct. These topics were chosen because of their link with the theme of accountability, as explained in the introduction of this article. Only English and Dutch studies were included because these languages are well-understood by the authors. To disclose research in the native language of the researchers to a non-Dutch-speaking network, publications in Dutch were consciously included in the analysis.

English studies were searched in 17 databases2, using 7 keywords3. Dutch studies were searched in 14 databases4, using 9 keywords5. We included two English terms (‘policing’ and ‘accountability’) in the search for Dutch literature because this terminology is also regularly used in Dutch publications6. This explains why we have two additional keywords in this language; not including these terms would have limited the number of (relevant) Dutch publications found. First, titles of publications were analyzed on their relevance for the research questions. This resulted in an English longlist of 862 publications and a Dutch longlist of 86 publications7. Subsequently, the authors determined whether or not the abstracts and/or full texts met the aforementioned inclusion criteria. This process resulted in an English shortlist of 77 publications. The exclusion of so many studies was mainly due to the lack of empirical research. A few publications were written in another language (e.g. Spanish) and some studies turned out to be irrelevant for the research question (e.g. explanations of police misconduct). Of these 77 publications, 41 were relevant for the first research question (accountability); 36 were relevant for the second research question (perceptions on accountability). Regarding the Dutch publications, this process resulted in a shortlist of 3 publications, which were all relevant for the first research question. In order to determine to which research question a particular publication would be appointed, the authors took the research questions, aim and/or description of the study into account. As such, only the publications that specifically announced they would study perceptions on accountability were appointed to the second research question. Nevertheless, when analyzing the studies included in this review, the authors noticed that in two of these publications, both accountability and perceptions on accountability were studied, which is why we decided to include them for both research questions8.

Subsequently, a quality or critical appraisal was executed, based on reporting criteria (Thomas & Harden, 2008; Verhage & Boels, 2016). We used five reporting criteria found in the literature, but in order to enhance the consistency of the assessment, we added sub-criteria. As such, we assessed which of the publications on our shortlist reported on at least three of the following reporting criteria:

1. Are the *research questions/objectives* clearly mentioned?
2. Is the *selection of participants* described explicitly? For this criterion, two sub-criteria were added: (a) Why were certain people (not) selected? and (b) Which sort of sample was used (e.g. purposeful, convenience etc.)?
3. Is a full description of the characteristics of the *sample* given? These characteristics were assessed in terms of (a) The number of participants and (b) An adequate description of the sample (e.g. a distinction between gender, age, function, ethnicity etc.). Both sub-criteria needed to be reported for this criterion to be qualified as a ‘yes’.
4. Are details of the *data collection* method given? This criterion was given a ‘yes’ when sufficient information was reported concerning the questions that were asked, information request procedures, search strategies etc.
5. Are details of *data analysis* presented? In this regard, we assessed if (a) Information was reported on the way in which the data were saved (e.g. field notes, tape recordings, transcriptions etc.), (b) Statistical methods were explained (quantitative) and (c) The authors clearly mentioned how the themes were created and how certain information was appointed to a particular code (qualitative).

If three or more of these reporting criteria were reported in the publication, the publication was included in the thematic synthesis; if not, it was excluded. As such, we only included studies that adequately reported their methodology in our further analysis. The main reason for this decision was our aim to review and disclose high quality empirical research. In order to determine whether or not the inclusion of the other studies would have made any difference with regard to the topics found, we conducted a sensitivity analysis on the basis of reporting criteria. The purpose of this analysis was to assess if the exclusion of inadequately reported studies limits the content of the thematic synthesis (Carroll, Booth, & Lloyd-Jones, 2012; Franzel, Schwiegershausen, Heusser, & Berger, 2013; Noyes & Popay, 2007; Thomas & Harden, 2008; Verhage & Boels, 2016). We found that excluding inadequately reported studies does make a difference regarding the number of topics found and the variety of information within these topics (Feys, Boels, & Verhage, 2017). Therefore, we would discourage future scoping reviews to exclude inadequately reported studies and instead encourage the inclusion of all relevant publications. However, as we conducted a quality appraisal, we did exclude these studies from the thematic synthesis reported in this article9. As such, only 10 publications in English and 1 Dutch publication regarding accountability remained. In sum, only 11 publications are included in the thematic synthesis. This is a striking finding: although one would expect that in peer reviewed publications this basic methodological information is provided, this clearly is not the case. Based on this finding, we expressly encourage researchers to report their methodology more adequately and recommend journals and reviewers to be critical on this point. A summary of the selection process throughout this review is presented in Figure 2.

Figure 2: Summary literature selection (based on research results, 2016)

Finally, the findings from the selected studies were analyzed and summarized by means of a thematic synthesis (Thomas & Harden, 2008). During this thematic synthesis, we limited ourselves to the two first phases (i.e. line-by-line coding and developing descriptive themes), resulting in the development of descriptive themes.

This research was initially conducted by two independent researchers, under close supervision of two other researchers. However, when the English shortlist was finished, one of the main researchers left the project, from which point on the first author conducted the subsequent phases on her own. Despite the dropout of one of the main researchers, the authors tried to compensate by holding regular meetings and seeking independent feedback whenever possible. The two senior researchers gave independent advice. When needed, the researchers conducted the phases independently from each other and compared their assessments.

1. **Findings**
   1. **General characteristics of the literature**

The studies that remained after the critical appraisal of the publications were published between 2004 and 2015. Generally, the timeframe of the fieldwork contained one to four years, with a time gap between (the beginning of) the fieldwork and actual publication between several months and 10 years, but not all authors provide sufficient details to determine this information. Two studies do not mention when the fieldwork was conducted, whereas seven others only mention when the research started or when a specific part of the research took place (e.g. the authors only mention when the observations took place, but do not give any information concerning the interviews or other methods used). Thus, only two studies specifically report the timeframe of the empirical research.

The majority of the studies in this review were published as journal articles (9). The journals in which these studies were published, are the British Journal of Criminology (2), the Australian and New Zealand Journal of Criminology (2), Police Practice and Research (2), Asian Criminology (1), European Journal of Social Psychology (1) and Cahiers Politiestudies (1). Aside from these articles, one book and one doctoral dissertation were included.

Most of the studies focused on one specific country, such as the United States (2), the Netherlands (1), Canada (1), China (1), the United Kingdom (1), Australia (1) and the Philippines (1), but some authors focused on a combination of countries, such as Australia and New Zealand (1) or China and the United States (1). One study did not specifically mention which country was studied. Nevertheless, the results section focused on the United States, the United Kingdom, Canada and Australia. This variety in study sites makes it difficult to compare the results of the studies given the differences in organizational structure of the police, legal frameworks and accountability systems.

* 1. **Methods used to study police accountability**

The majority of the studies were qualitative in nature (8), but a small minority also used quantitative methods (2). In one study, a mixed methods approach was used (quantitative and qualitative). More specifically, the following methods were used: interviews (7), document search and analysis (7), observations (4), questionnaires (2), informal conversations (2), participant observation (1) and focus groups (1). Figure 3 summarizes the methods used to study police accountability.

Figure 3: Methods used to study police accountability (based on research results, 2016)

* 1. **Thematic synthesis**

Only few authors specifically studied the subject of this review (e.g. van Duijneveldt, 2015). Most of them studied police accountability in relation to another theme, such as intelligence-led policing (Sanders, Weston, & Schott, 2015) or police misconduct (Kelly, 2012). The research subjects studied in the empirical literature concerning police accountability are quite diverse. In general, three broad themes can be distinguished. The first main theme refers to historical facts about accountability or integrity-related changes, as will be explained below. A second theme that was identified was police accountability itself, which more generally refers to being accountable for what one does or how things are done. More specifically, authors studied what police accountability is, to whom police officers must account and how police accountability can be achieved. The third theme is integrity, which was a regularly studied topic. In the included studies, integrity often referred to the absence of police misconduct. Subcategories of this theme are the meaning of integrity, which factors influence whether or not the police (department) is integer, types of misconduct, how this (alleged) misbehavior can be investigated and which outcome and consequences it can generate and the satisfaction concerning these investigative procedures and consequences. Figure 4 presents an overview of the themes found in the included publications. In appendix 2, the main findings of the included studies are briefly summarized.

Figure 4: Themes found in the respective publications (based on research results, 2016)

*History*

Several authors described certain historical facts in order to have a better understanding about the current situation of accountability or, more regularly, integrity (Klockars et al., 2006; Prenzler, 2004). Examples are the externalization of police misconduct investigations in New South Wales (Prenzler, 2004) and the introduction of independent oversight bodies in Australia in the 90s (den Heyer & Beckley, 2013).

*Accountability*

*What is police accountability?* Van Duijneveldt (2015) was the only author who specifically studied the meaning of accountability. He interviewed 23 district agents and 10 chiefs in the Netherlands about the meanings they assign to the term and to whom they need to account. He concluded that police officers associate accountability with, among others, the judgement of (wrongful) conduct of police officers, explaining certain situations, the sharing of information, transparency and integrity, receiving feedback and reflecting on the job. The sharing of information and being accountable for the execution of the job are the most common meanings attached to the term. In other studies, both the researchers and respondents referred to financial or economic accountability (Cronin & Reicher, 2009; Sanders et al., 2015).

Accountability is important because it can influence the decision making process of police officers (Cronin & Reicher, 2009). In this regard, van Duijneveldt (2015) highlights that accountability has different functions, such as learning, reflecting, developing a common vision in the team and professional development. He concludes that some police officers see accountability as an inherent part of their job and hence prefer the term ‘professional dialogue’, which refers to a conversation in which police officers and chiefs reflect on their work.

*To whom do police officers account?* Police officers account to multiple audiences. On the one hand, there is *internal accountability*, which refers to the direct chief, colleagues and executives (Cronin & Reicher, 2009; van Duijneveldt, 2015). On the other hand, *external accountability* can be identified. This form of accountability is associated with the local community, politics, civilian review agencies, police auditors, the media and the public (Cronin & Reicher, 2009; den Heyer & Beckley, 2013; Kelly, 2012; Klockars et al., 2006; Sanders et al., 2015; van Duijneveldt, 2015).

In addition, van Duijneveldt’s (2015) study showed that some police officers believe they need to be accountable to themselves. They should be able to explain to themselves why certain decisions were made, especially considering the autonomy and freedom they are granted when performing their tasks.

In the United Kingdom, Cronin & Reicher (2009) found, on the basis of interviews with commanders and observations of police officers, that the different actors involved in accountability issues sometimes have different priorities, which can lead to accountability dilemmas. Commanders wiLL, for example, be more responsive to audiences that can sanction them or their organization. In this regard, several respondents in van Duijneveldt’s (2015) study reported that they experienced difficulties in deciding to which audience they are primarily accountable: the chief or the district.

*How can police accountability be achieved?* Police accountability can be achieved in multiple ways, such as the use of intelligence-led policing (ILP) and crime analytics to justify decisions (Sanders et al., 2015); video recordings to simplify investigations of complaints and improve accountability in interactions between the police and civilians, e.g. by making sure that police misconduct is sanctioned (Kelly, 2012); by using independent oversight bodies (den Heyer & Beckley, 2013) or making the police more publicly accountable by making every record of police misconduct and its investigation public (Klockars et al., 2006); or peer debriefings, which enable police officers to reflect on their behavior and can possibly result in the learning of (new) techniques in dealing with certain situations (Klockars et al., 2006).

*Integrity*

Despite many authors studying the topic of integrity, only one of the studies included in this review specified the meaning of integrity. Klockars, Ivkovic, & Haberfeld (2006) used the following definition: “*the normative inclination among police to resist temptations to abuse the rights and privileges of their occupation*” (p. 251). According to these authors, integrity is – among others – linked with (the absence of) corruption and other misconduct, the creation and communication of organizational rules and investigation of violations of these rules.

*Affecting factors.* Different factors can impact upon a police officer’s or police department’s integrity. Examples are the organizational culture and code of silence (den Heyer & Beckley, 2013; Jiao, 2010; Klockars et al., 2006; Prenzler, 2004), colleagues (both direct colleagues and supervisors or commanders) (Klockars et al., 2006), receptivity to citizen complaints (Klockars et al., 2006; Ransley et al., 2007) and the triangle of integrity (Klockars et al., 2006). This triangle refers to the recruitment, selection and training phase of police officers.

*Police misconduct.* Multiple forms of police misconduct could be identified throughout the publications. Examples are abuse of power, police corruption, excessive or unnecessary use of force, discourtesies, false arrests (arrests without probable cause), failure to identify as a police officer, sexual assault and verbal and physical abuse, racism or racial profiling (e.g. in traffic stops) and discrimination within the police (Jiao, 2010; Kelly, 2012; Klockars et al., 2006; Prenzler, 2004; Ransley et al., 2007; Xu, 2013). There is, however, no full consensus on appropriate behavior. For example, sleeping on duty can be tolerated in certain situations (e.g. in case of a lack of sleep due to the death of a relative) (Klockars et al., 2006).

*Investigation*. Similar to internal and external accountability, the investigation of police misconduct can either be internal or external (Jiao, 2010). Such an investigation can take on many forms, e.g. an internal affairs investigation, a civil litigation or a criminal trial/lawsuit, or an investigation by a civilian review agency, an independent police auditor or city bookkeepers (e.g. discrepancies in payroll records) (Kelly, 2012; Klockars et al., 2006; Prenzler, 2004; Ransley et al., 2007). Some complaints are not investigated (Kelly, 2012; Ransley et al., 2007). However, in almost half of the cases in Prenzler’s study in which the police decided not to investigate the complaint, an investigation would have been appropriate.

Prenzler (2004) analyzed the results of different studies concerning the investigation of police misconduct. In many of these studies, an independent body to perform audits or execute investigations is recommended. One of the main purposes of police independent oversight bodies is to improve public trust in the police (den Heyer & Beckley, 2013). However, some police officers believe that outside investigators do not contain the necessary background or skills to conduct a fair investigation (Prenzler, 2004).

One can also distinguish between more and less severe types of misconduct. Serious complaints could be investigated by an independent body or internal affairs, less serious complaints could be investigated by police themselves (e.g. supervisors) (den Heyer & Beckley, 2013; Klockars et al., 2006). The question then remains which types of behavior are more/less severe. Nonetheless, some people point to the (personal) importance of less serious complaints to the complainants, in which case an independent investigation might seem more appropriate as well (den Heyer & Beckley, 2013; Prenzler, 2004). Another possibility is a mixed investigation team, dominated by civilians (Prenzler, 2004). A mixed composition of investigators could give the impression that the investigation is more thorough and fair (de Guzman, 2007).

*Outcome and consequences*. Although the study of the **outcome** was rather limited, it can be noted that three solutions were discovered: a complaint can be sustained, it can be regarded as unfounded or there may be insufficient evidence available to come to a conclusion (Klockars et al., 2006). A police officer can hence be disciplined or even convicted, cleared or acquitted or the charges can be dropped (Kelly, 2012).

There is a wide variety of possible **consequences**, which are situated on the level of the individual police officer, on the level of the police (department) or an even broader level. However, not every (alleged) case of police misconduct is disciplined, or, exceptionally, the consequences can be reversed (e.g. a police officer is reinstated) (Kelly, 2012; Klockars et al., 2006). Prenzler (2004) found that 22% of the cases in which no discipline was given, discipline would have been appropriate.

An individual police officer can, among others, be fired, suspended with(out) pay, resign, receive a written or oral reprimand, be demoted and lose his/her reputation. Besides that, some police officers commit suicide after an (alleged) misbehavior (Jiao, 2010; Kelly, 2012; Klockars et al., 2006; Prenzler, 2004). Klockars, Ivkovic, & Haberfeld (2006) found that a written reprimand and suspension are more common, whereas dismissal is quite rare.

The police department or entire organization can experience some consequences as well, such as a loss of public respect or changes in policy (Jiao, 2010; Kelly, 2012; Klockars et al., 2006; Prenzler, 2004). More in general, public protests can be provoked, press conferences can be held and indemnifications are possible (Kelly, 2012).

*Satisfaction*. The last theme that was identified in this review was the satisfaction in regard to (the outcome or consequences of) investigations of police misconduct, which was only scarcely studied. Prenzler (2004) studied sources of opinion, institutional websites and surveys from complainants, the police and the public opinion. He noticed that serious misconduct is sometimes punished by light sentences, which can lead to public disquiet. Complainants sometimes believe that investigation processes favor the police. Many complainants are thus advocates for an independent body. Moreover, many of these complainants seemed to be dissatisfied with the outcome of the investigation (Prenzler, 2004).

1. **Discussion and conclusions**

This review has shed a light on both the process of a scoping review, based on the principles of a systematic literature review, and the content of accountability-related, empirical studies. Mainly, police accountability is studied in regard to police misconduct or different aspects that are somehow linked with the theme, but go beyond solely accountability-focused research.

The results of this scoping review brought up some points of discussion. The most striking finding throughout this entire study was the limited number of empirical studies included in the thematic synthesis. For this review, only 11 publications remained after the quality or critical appraisal. These were the only (empirical) studies that reported sufficiently on their methodology and still contained valuable information with regard to the research question. The latter was formulated rather broadly, which is why the authors predicted that more publications would have been included. The longlist contained 948 potentially relevant publications, but only 8% (N=80) of these publications remained after analysis of the inclusion criteria. One of these criteria was the need for empirical research. From the 80 remaining publications on the shortlist, only 14% (N=11) remained after the quality appraisal. As such, only 1% of the initial longlist remained for the thematic synthesis in regard to the first research question (‘What is accountability?’). The two main reasons identified for this finding, were a lack of empirical research on the topic and insufficient reporting on methodology. More empirical research into police accountability and, moreover, more attention for methodological considerations when reporting research results, are recommended.

The limited number of empirical studies included in the thematic synthesis, could partly be the result of a methodological artefact. First, the keywords and databases served as a first selection. Approximately 30 databases were searched based on rather generally formulated keywords (e.g. ‘accountability’, ‘oversight’), resulting in a range of different themes related to the topic. However, future researchers might benefit from an even wider search than the scope of this current study, considering that important themes/subjects in regard to police accountability were not included in the thematic synthesis of this review (e.g. body-worn cameras). We did retrieve some of these themes on our longlist, but they were mostly excluded based on the inclusion criteria (second selection) or critical appraisal (third selection). As such, there might be more empirical research on police accountability than this review suggests. It could be beneficial to revise the keywords used in the course of this review and to think of some additional, perhaps more specific terms, such as ‘professional dialogue’.

We excluded 33 publications based on the quality assessment because they contained insufficient methodological information. Our sensitivity analysis showed that, had we included these 33 studies in our analyses, the results would have been different. More themes would have been found, meanwhile enriching the information within the themes. Assuming that there is a relatively limited number of empirical studies on police accountability as main study subject, we can wonder why this is the case. As in many countries that adhere the principle of community-oriented policing, accountability is central to the policing philosophy (Kelling, Wasserman, & Williams, 1988; Skolnick & Bayley, 1988a; Skolnick & Bayley, 1988b), so one could expect that this topic provides enough body for research. Perhaps ‘police accountability’ is rather difficult to study due to the sensitivity of the topic? There is such a variety of accountability mechanisms that it seems quite hard to get a full image of what accountability is or how it is achieved in different countries. This might also explain why many researchers study the topic in a certain country (or limited number of countries), always taking into account the specific context. Or is accountability mostly an issue relevant to practice and less to the academic world, which is often more theory oriented?

Another point of discussion relates to the finding that it was quite hard to get a clear understanding of the meaning of police accountability based only on the publications included in the thematic synthesis. No single, unanimously agreed upon definition could be found. Such a definition would be beneficial for the comparability of different studies and the possibilities to measure the concept. Then again, an unanimously accepted definition is probably wishful thinking due to the differences in policing structures and different sorts of accountability mechanisms. After all, multiple ways of achieving accountability can be identified. This review has discussed some of these methods, but a lot of options exceed this study. It would be useful to map all these accountability mechanisms, their effectiveness and efficiency and whether or not stakeholders believe it is a good way to pursue accountability. We lack this kind of overview. We need to be willing to question the existing accountability mechanisms and try to propose other, perhaps even better accountability structures. Discussions about this theme need to be enhanced, and accountability mechanisms need to be evaluated. Harris (2013) states, for example, that external police oversight is spreading throughout the world, but we do not know whether or not such a system is efficient.

In order to be able to propose advanced accountability structures, we need to start by studying police accountability more frequently, particularly in an empirical manner and preferably as main study subject. This kind of research could be extremely beneficial to society. Policing, and especially the use of their powers (most importantly the use of force), is a heavily debated topic. As such, more academic (empirical) research on the topic is crucial to get a clear view on the ways in which these police powers are applied and which factors impact on this. This can contribute to the legitimacy of and trust in the police (cfr. procedural justice). Satisfaction with police accountability is, according to De Angelis & Wolf (2016), a strong and consistent predictor of satisfaction with the police. As a consequence, when police are regarded as a legitimate institution, citizens might be more likely to cooperate with police officers. Hence, this kind of research could potentially have a positive influence on police practices. Achieving accountability is challenging, but studying this topic can enhance police accountability and, hence, impact upon policing in general (Walker & Archbold, 2014).

In our opinion, more empirical research in regard to police accountability is desirable. An example of a striking lacuna is the existence of and solution to accountability dilemmas. Such research seems rather useful and could be studied through different methodologies. One the one hand, one could quantitatively measure which solution is preferred. Does the majority believe that a police officer should still report to all relevant stakeholders? On the other hand, qualitative research would allow researchers to broaden our understanding as to why one makes a certain decision and how. Perhaps a mixed methods design would be most appropriate, which would allow researchers to do both. Other examples of study subjects are preferences in regard to accountability mechanisms and the effectiveness of these mechanisms. Systematically conducted literature reviews could prove beneficial in this regard. Since the literature on police-related topics is quite unstructured, more such reviews would enable us to get an overview as to which topics have already been studied (and how) and which gaps are yet to be filled. After all, police research is often fragmented due to varying policy issues and urgencies, which is why it is crucial to first map existing research and identify gaps and lacunas, before adding to the scatterplot of police research. Only then can we add to the existing body of police research and build on firm fundaments.

**Funding**

This study was funded by FWO Vlaanderen.

**Notes**

1. The purpose of this project was to map the research from the last 25 years (i.e. 1990-2015, when the study started). This boundary was defined at the outset of the study, partly because of feasibility reasons (limited amount of time and limited resources).

2. Web of Science, Sociological Abstracts, Campbell Library, International Bibliography of the Social Sciences (IBSS), Google Scholar, Policing: A Journal of Policy and Practice, Policing: An International Journal of Police Strategies and Management, Journal of Police and Criminal Psychology, The Police Journal: Theory, Practice and Principles, Police Practice and Research, Policing and Society, European Journal of Policing Studies (EJPS), Cahiers Politiestudies (CPS), American Society of Criminology (ASC), European Society of Criminology (ESC), Criminology Library Grey Literature (CrimDoc) and Social Science Research Network (SSRN) e-library.

3. Polic\* AND accountability, polic\* AND misconduct, polic\* AND integrity, polic\* AND transparency, polic\* AND oversight, polic\* AND control.

4. Vlaams Academisch Bibliografisch Bestand (VABB), Antilope, Belgische Bibliografie, Tijdschrift voor Criminologie, Tijdschrift voor Veiligheid, Justitiële Verkenningen, Panopticon, Politie en Wetenschap, website Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC), European Journal of Policing Studies (EJPS), Cahiers Politiestudies (CPS), LibHub, Google Scholar and Handboek Politiediensten.

5. Politie\* EN verantwoording, politie\* EN integriteit, politie\* EN optreden, politie\* EN controle, politie\* EN toezicht, politie\* EN transparantie, policing, accountability.

6. The authors of this article experimented with keywords before the actual start of the project as to identify the most appropriate keywords. We questioned the option to include specific terms (such as body-worn cameras), considering such keywords would steer the study in one (or some) specific direction(s). Missing out other relevant, specific keywords, would be truly negative in this case. As such, we chose to include more general keywords, to keep our options open and to make sure we did not specifically include certain subtopics of accountability while unconsciously excluding other subtopics. If one aims to conduct an exhaustive review on this topic, it could be recommended to include more specific keywords and different terminology for each country (e.g. ‘internal affairs’ in the USA, ‘interne toezicht’ for Belgium). However, this is only feasible when the review is limited to one or few countries. As the review on which this article is based was not limited to a certain group of countries, it would not have been feasible (given the amount of time and resources available) to do so. This could be an interesting study in the future though.

7. Every contribution was counted separately. When a book chapter or an edited book was found for the longlist, the entire edited book was taken into account and the book chapters were counted separately. However, not all publications were available to the authors. If an edited book was not fully available to the authors, only the book chapters that were initially found and/or available were counted.

8. As such, one English study (Klockars, Ivkovic, & Haberfeld, 2006) and one Dutch study (van Duijneveldt, 2015) will also be included for the second research question, although being initially appointed to the first research question.

9. The sensitivity analysis was conducted as to find out whether or not authors should include studies of lesser quality in their scoping review. Based on the sensitivity analysis, we would recommend to include all studies, despite their quality. However, the results presented in this article are based on the thematic synthesis of our review. In this synthesis, we analyzed each text and identified theme meticulously, in depth. This was not the case in the sensitivity analysis; such an analysis aims to find out if new themes would be found if studies of lesser quality would be included and if so, if any information would be lost. In other words: the sensitivity analysis analyzed the identification of (new) themes (despite the specific content within these themes), whereas the thematic synthesis (reported in this article) analyzed these themes more in depth.

**References**

Boels, D., & Verhage, A. (2016). Plural policing: A state-of-the-art review. *Policing: An International Journal of Police Strategies & Management, 39*(1), 2-18.

Busuioc, M., Curtin, D., & Groenleer, M. (2011). Agency growth between autonomy and accountability: The European Police Office as a 'living institution'. *Journal of European Public Policy, 18*(6), 848-867.

Carroll, C., Booth, A., & Lloyd-Jones, M. (2012). Should we exclude inadequately reported studies from qualitative systematic reviews? An evaluation of sensitivity analyses in two case study reviews. *Qualitative Health Research, 22*(10), 1425-1434.

Chan, J. B. L. (1999). Governing police practice: Limits of the new accountability. *British Journal of Sociology, 50*(2), 251-270.

Cheung, J. (2005). Police accountability. *The Police Journal, 78*(1), 3-36.

Coudert, F., Butin, D., & Le Métayer, D. (2015). Body-worn cameras for police accountability: Opportunities and risks. *Computer Law & Security Review, 31*(6), 749-762.

Cronin, P., & Reicher, S. (2009). Accountability processes and group dynamics: A SIDE perspective on the policing of an anti-capitalist riot. *European Journal of Social Psychology, 39*(2), 237-254.

De Angelis, J. (2009). Assessing the impact of oversight and procedural justice on the attitudes of individuals who file police complaints. *Police Quarterly*, *12*(2), 214-236.

De Angelis, J., & Kupchik, A. (2007). Citizen oversight, procedural justice, and officer perceptions of the complaint investigation process. *Policing: An International Journal of Police Strategies & Management*, *30*(4), 651-671.

De Angelis, J., & Wolf, B. (2016). Perceived accountability and public attitudes toward local police. *A Critical Journal of Crime, Law and Society, 29*(3), 232-252.

de Guzman, M. C. (2007). Integrity of civilian review: A contemporary analysis of complainants' and police officers' views in the Philippines. *Police Practice and Research, 8*(1), 31-45.

den Boer, M. (2002). Towards an accountability regime for an emerging European policing governance. *Policing and Society, 12*(4), 275-289.

den Heyer, G., & Beckley, A. (2013). Police independent oversight in Australia and New Zealand. *Police Practice and Research, 14*(2), 130-143.

Eijkman, Q. (2006). To be held accountable: Police accountability in Costa Rica. *Police Practice and Research, 7*(5), 411-430.

Feys, Y., Boels, D., & Verhage, A. (2017). Sensitivity analysis in a scoping review on police accountability: Assessing the feasibility of reporting criteria in mixed studies reviews. In K. Hannes, A. Heylighen, B. Dierckx de Casterlé, F. Truyen, & A. Declercq (Eds.), *ECQI 2017 Proceedings: Quality and reflexivity in qualitative inquiry* (pp. 114-126). Leuven: KU Leuven NQRL.

Findlay, M. (1994). The ambiguity of accountability: Deaths in custody, and the regulation of police power. *Current Issues in Criminal Justice, 6*(2), 234-251.

Franzel, B., Schwiegershausen, M., Heusser, P., & Berger, B. (2013). How to locate and appraise qualitative research in complementary and alternative medicine. *BMC Complementary and Altnernative Medicine, 13*(125), 1-10.

Gill, C., Weisburd, D., Telep, C. W., Vitter, Z., & Bennett, T. (2014). Community-oriented policing to reduce crime, disorder and fear and increase satisfaction and legitimacy among citizens: A systematic review. *Journal of Experimental Criminology, 10*(4), 399-428.

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., Kyriakidou, O., & Peacock, R. (2005). Storylines of research in diffusion of innovation: A meta-narrative approach to systematic review. *Social Science & Medicine, 61*(2), 417-430.

Harris, D. A. (2009). How accountability-based policing can reinforce - or replace - the fourth amendment exclusionary rule. *Ohio State Journal of Criminal Law*, *7*, 149-213.

Harris, D. A. (2010). Picture this: Body-worn video devices (head cams) as tools for ensuring fourth amendment compliance by police. *Texas Law Review*, *43*, 357-371.

Harris, F. C. (2013). *Holding police accountability theory to account* [Doctoral dissertation]. Portsmouth: University of Portsmouth.

Jiao, A. Y. (2010). Controlling corruption and misconduct: A comparative examination of police practices in Hong Kong and New York. *Asian Journal of Criminology, 5*(1), 27-44.

Kelling, G. L., Wasserman, R., & Williams, H. (1988). *Police accountability and community policing*. Washington: National Institute of Justice.

Kelly, D. A. (2012). *Cops, cameras and accountability: User-generated online video and public space police-civilian interactions* [Doctoral dissertation]. Maine: The University of Maine.

Klockars, C. B., Ivkovic, S. K., & Haberfeld, M. R. (2006). *Enhancing police integrity*. Dordrecht: Springer.

Lamboo, T. (2010). Police misconduct: Accountability of internal investigations. *International Journal of Public Sector Management, 23*(7), 613-631.

Manning, P. K. (2012). Trust and accountability in Ireland: The case of *An Garda Síochána*. *Policing and Society, 22*(3), 346-361.

Mazeika, D., Bartholomew, B., Distler, M., Thomas, K., Greenman, S., & Pratt, S. (2010). Trends in police research: A cross-sectional analysis of the 2000-2007 literature. *Police Practice and Research, 11*(6), 520-547.

McGregor, A. J. (2015). Politics, police accountability, and public health: Civilian review in Newark, New Jersey. *Journal of Urban Health, 93*(1), 141-153.

Millen, F., & Stephens, M. (2011). Policing and accountability: The working of police authorities. *Policing and Society, 21*(3), 265-283.

Muller, E. R. (2002). Policing and accountability in the Netherlands: A happy marriage or a stressful relationship? *Policing and Society, 12*(4), 249-258.

Myhill, A. (2007). Police authorities' public accountability role: Learning from three community engagement demonstration projects. *Policing: A Journal of Policy and Practice, 1*(2), 173-183.

Noyes, J., & Popay, J. (2007). Directly observed therapy and tuberculosis: How can a systematic review of qualitative research contribute to improving services? A qualitative meta-synthesis. *Journal of Advanced Nursing, 57*(3), 227-243.

Ponsaers, P. (2010). De verhouding politie en wetenschappelijk onderzoek - Kritische analyse van 25 jaar politieonderzoek in Europa en de Angelsaksische wereld. In W. Bruggeman, E. Devroe, & M. Easton (Eds.), *Evaluatie van 10 jaar politiehervorming: Terugkijken in het verleden en vooruitkijken in de toekomst* (pp. 131-150). Antwerpen-Apeldoorn: Maklu.

Prenzler, T. (2004). Stakeholder perspectives on police complaints and discipline: Towards a civilian control model. *The Australian and New Zealand Journal of Criminology, 37*(1), 85-113.

Ransley, J., Anderson, J., & Prenzler, T. (2007). Civil litigation against police in Australia: Exploring its extent, nature and implications for accountability. *The Australian and New Zealand Journal of Criminology, 40*(2), 143-160.

Sanders, C. B., Weston, C., & Schott, N. (2015). Police innovations, 'secret squirrels' and accountability: Empirically studying intelligence-led policing in Canada. *British Journal of Criminology, 55*(4), 711-729.

Schaible, L. M., De Angelis, J., Wolf, B., & Rosenthal R. (2012). Denver's citizen/police complaint mediation program: Officer and complainant satisfaction. *Criminal Justice Psychology Review*, *24*(5), 626-650.

Skolnick, J. H., & Bayley, D. H. (1988a). Theme and variation in community policing. *Crime & Justice*, *10*, 1-37.

Skolnick, J. H., & Bayley, D. H. (1988b). *Community policing: Issues and practices around the world*. Washington: National Institute of Justice.

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*(45), 1-10.

van Duijneveldt, I. (2015). Wijkagenten over verantwoording. *Cahiers Politiestudies, 4*(37), 57-76.

Verhage, A., & Boels, D. (2015). Systematische reviews van kwalitatief onderzoek binnen de criminologie. *Panopticon, 36*(3), 302-311.

Verhage, A., & Boels, D. (2016). Critical appraisal of mixed methods research studies in a systematic scoping review on plural policing: Assessing the impact of excluding inadequately reported studies by means of a sensitivity analysis. *Quality & Quantity, 50*(3), 1-20.

Walker, S. (2007). Police accountability: Current issues and research needs. Washington: National Institute of Justice.

Walker, S., & Archbold, C. A. (2014). *The new world of police accountability*. Thousand Oaks: Sage.

Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J., & Pawson, R. (2013). RAMESES publication standards: Meta-narrative reviews. *BMC Medicine, 11*(20), 1-15.

Xu, J. (2013). Police accountability and the commodification of policing in China: A study of police/business posters in Guangzhou. *British Journal of Criminology, 53*(6), 1093-1117.

**Appendix**

Appendix 1: Methods used to study police accountability in the respective publications

|  |  |
| --- | --- |
| method | studies |
| interviews (7) | Cronin & Reicher (2009); den Heyer & Beckley (2013); Jiao (2013); Klockars, Ivkovic, & Haberfeld (2006); Sanders, Weston, & Schott (2015); van Duijneveldt (2015); Xu (2013) |
| document search and analysis (7) | den Heyer & Beckley (2013); Jiao (2010); Kelly (2012); Klockars, Ivkovic, & Haberfeld (2006); Prenzler (2004); Ransley, Anderson, & Prenzler (2007); Xu (2013) |
| observations (4) | Cronin & Reicher (2009); Klockars, Ivkovic, & Haberfeld (2006); Sanders, Weston, & Schott (2015); Xu (2013) |
| questionnaires (2) | de Guzman (2007); Klockars, Ivkovic, & Haberfeld (2006) |
| informal conversations (2) | Cronin & Reicher (2009); Klockars, Ivkovic, & Haberfeld (2006) |
| focus groups (1) | Klockars, Ivkovic, & Haberfeld (2006) |
| participant observation (1) | Sanders, Weston, & Schott (2015) |

Appendix 2: Overview of the included publications

|  |  |  |  |
| --- | --- | --- | --- |
| author(s) | location | method/sample | overview of the publication and results[[1]](#footnote-1) |
| van Duijneveldt (2015) | the Netherlands | The author interviewed 23 district agents and 10 police chiefs in 2010. | The author studied the meaning police officers appoint to the term ‘accountability’ and to whom police officers must account. The most prominent results were that police officers apply multiple meanings to the term ‘accountability’ (e.g. control, assessing whether or not certain behavior was acceptable, explaining certain actions, sharing of information and getting feedback) and that they need to account to multiple audiences (e.g. the chief, colleagues, superiors and politicians). |
| Sanders, Weston, & Schott (2015) | Canada | The authors conducted in-depth interviews with 30 crime/intelligence analysts, 26 patrol officers and 30 officers/civilians working within police information technology bureaus from six different police services across Canada. These interviews were supplemented with 36 hours of observation in patrol policing and participant observation in three crime analytic workshops and two CACP workshops on police information technology and information management. | The authors mainly focus on intelligence-led policing, information regarding technology and manners in which data can be analyzed. Nevertheless, a small section concerning accountability was included. In this regard, the authors explain that police officers need to account financially, to the communities and police service boards and that intelligence-led policing can be used to achieve (external) accountability, e.g. by justifying certain decisions, for example financial decisions. |
| Xu (2013) | China | The author conducted weekly observations from 2009-2012 and in-depth semi-structured interviews with 21 police officers, 3 police scholars, 5 businessmen, 5 urban management officers, 6 ordinary citizens and 2 security guards. In addition, newspaper reports and police documents/directives about police/business posters were used. | The author explains the commodification of policing, meaning that lately the police can be purchased more often. This commodification in China lacks accountability of police power. Basically, the police cooperate with businesses in order to produce police/business posters. As such, the police can freely promote crime prevention. However, an individual police officer can also abuse these powers in order to make money. Questions regarding neutrality in law enforcement are raised. |
| Cronin & Reicher (2009) | the United Kingdom | The authors conducted observations on 18 June 1999 and (largely unstructured) interviews with the Silver and Gold commanders. They also had informal conversations with a wide range of officers at all levels. | This analysis is based on the events of J18. This was an amalgam of events with no single overall organizer, also referred to as a ‘Carnival against Capitalism’. The authors refer to the importance of police accountability (e.g. protection of police reputation) and the difference between external and internal accountability. They also refer to the possible emergence of accountability dilemmas when one needs to account to multiple audiences. |
| Ransley, Anderson, & Prenzler (2007) | Australia | The authors obtained reports by searching the Australian Legal Information Institute (AustLII) database and media reports, retrieved from Factiva. The authors used keywords which were derived from the literature and case names of known incidents of litigation against the police. | The authors suggest that civil litigation can be used as an accountability measure. The authors found that there is an upward movement in the number of claims against the police. On the one hand, citizens sue police (e.g. about excessive use of force or negligence). On the other hand, police sue police (e.g. human resource management decisions or negligence). Police officers sue the police almost as frequently as civilians do. Citizens have a high success rate when these claims involve (alleged) abuse of power. The police have a success rate of 70%. |
| Prenzler (2004) | not specified | The author systematically searched for sources of opinion and institutional websites in a policy-oriented critical review format. He used electronic search engines and a snowball method with regard to the references in these sources. The searches were conducted in 2001 and again in 2002. The author used both primary sources (such as judicial inquiry reports and annual reports) and secondary sources (including, among others, complainant surveys and public opinion surveys). | After a brief summary of the history regarding judicial inquiries, which have contributed to the rationale for external review, the author discusses (dis)advantages of external oversight. The main advantage is the public confidence that is generated by using external oversight. The main disadvantage is the lack of police experience. Complainants’ views are discussed (most of which support independent oversight), complemented by police officers’ views. Police officers’ opinions are divided. Some are supportive of independent oversight, some reject such a system. The author summarizes some studies in this regard. Besides these actors, the public opinion and civil liberties groups are also discussed, alongside government reviews and miscellaneous groups. The author suggests that independent control of complaints should be done by mixed civilian/police teams dominated by civilians. Another option is a distinction between less and more serious complaints, investigated by different investigators, respectively the police and external investigators. |
| Kelly (2012) | the United States | The author conducted a descriptive multiple case study of 14 cases regarding police-civilian interaction in American public space that have been video recorded by a civilian, on the basis of document search and analysis (e.g. video recordings, police reports, court filings etc.). All research materials were obtained from public sources. | The author studies the outcome of user-generated online video on police civilian interactions in American public space, hereby, among others, hypothesizing that such videos can possibly improve accountability in such interactions. The author describes multiple cases related to such interactions, thereby giving a lot of examples of police misconduct and possible investigation methods. In addition, the outcome and consequences of such investigations are discussed (on multiple levels: individual police officer, organization and more in genral). Indirectly, the author also refers to whom police officers must account (e.g. the press). |
| Klockars, Ivkovic, & Haberfeld (2006) | the United States | The authors started by executing a questionnaire. A convenience sample of 3,235 police officers was obtained. Subsequently, interviews, focus groups (and study groups), observations, informal conversations and a document search and analysis were conducted. Finally, a second survey was distributed, which obtained a sample of 1,544 police officers. | The authors start by conceptualizing ‘police integrity’. Subsequently, they measure police integrity, based on a questionnaire with 11 hypothetical scenarios, in 30 departments. They concluded, for example, that when police officers perceived a certain behavior to be more serious, they believed a more severe discipline should and would be received and they would be more willing to report it. On the basis of this analysis, the authors selected three agencies in which to conduct a subsequent, in-depth study. The integrity profiles of these departments are described, after which all three departments are studied more closely. The authors try to explain how the (high) levels of integrity are achieved and maintained in these agencies. In order to do so, they describe the specific characteristics of these departments and relevant historical events. As such, forms of police misconduct, investigations, outcomes and consequences etc. are all discussed. The authors also conducted a second survey, to alleviate some of the flaws in the first one. This survey enabled the authors to identify areas that could be improved (e.g. training). In the subsequent chapters, the authors describe the effect of recruitment, selection and training, processing of citizen complaints, discipline and the code of silence on a department’s integrity. Finally, they give a summary of the research findings. |
| den Heyer & Beckley (2013) | Australia and New Zealand | The authors examined official reports (such as annual reports and websites) and conducted interviews with staff from four independent oversight agencies. | The authors discuss the developments in independent oversight in New Zealand, New South Wales, Queensland and Victoria. They stress the importance of police independent oversight bodies and their role in achieving accountability and trust in the police. They also emphasize the change from police internal (pre-1980) to external oversight (1980-2000) and, eventually, to cross-public sector integrity commissions (2000 onwards). The authors conclude that independent oversight bodies are crucial for today’s society and that public confidence is essential. |
| Jiao (2010) | China and the United States | The author used inductive methods. He conducted semi-structured and non-structured interviews with 36 police officers in Hong Kong and 12 police officers from New York (during 2001-2003). In addition, he conducted a document search and analysis, to complement the personal interviews by providing historical and current information that is relevant for the research questions. | The author explains that there are different organizational approaches regarding corruption and misconduct. Corruption control can either be internal (e.g. New York) or external (e.g. Hong Kong). He states that Hong Kong’s control mechanisms indicate three important factors concerning the effective handling of police corruption: an independent agency with the ability of conducting investigations on their own, a strong police culture against corruption and a strong relationship between the police and an external investigative body. |
| de Guzman (2007) | the  Philippines | The author conducted a cross-sectional survey with 515 complainants, 206 officers who had cases in the PLEB and 514 police officers who never had a case in the PLEB. These surveys were administered through a structured interview or through self-administration. The study was done in 2001. | The author addresses the differences in opinion between complainants and police officers regarding the PLEB (People’s Law Enforcement Board). Complainants thought this board was police-dominated, whereas police officers thought it was civilian-dominated. The author contradicts previous research regarding the topic, considering he found that police without cases in the PLEB showed satisfaction with a civilian-dominated board. The author suggests a mixed model of a civilian review board, because it might give the feeling of a more fair and thorough investigation. |

1. We chose to give an overview of the included publications. As such, we describe the outset of the publications and the most prominent results. Some of the publications were thus extensive that a short summary of the results would not be meaningful (e.g. the book of Klockars, Ivkovic, & Haberfeld, 2006). We hence advise readers to consult the respective publications when interested to get a full image of the content. [↑](#footnote-ref-1)