
From lesson-learning to accountability?

*Dutch development aid evaluation in transnational
perspective, 1964-1977*

ABSTRACT

The effectiveness and efficiency of modern development aid has been criticized since its post-war inception. Commentators note the lack of feedback and accountability; the sector does not learn from its past mistakes, and no one takes responsibility in case of clear failures. Evaluation, then, is an instrument that enhances both feedback and accountability. Aid evaluation in one form or another takes place on a massive scale nowadays. How and why did it come into being? This thesis traces the roots of aid evaluation in the Netherlands in the 1960s and 1970s. It shows that, in the context of growing academic knowledge and international consultation, discussions on the purpose and the locus of evaluation led to the establishment of a dual system of aid evaluation at the Dutch Ministry of Foreign Affairs. Civil servants shaped these debates without significant and direct influence of parliament.

Paul Schilder (3978478)

Supervisor: dr. L. van de Grift

Second reader: prof. dr. B.A. de Graaf

Research MA Modern History (1500-2000)

Utrecht University

30 835 words

15 August 2019

ACKNOWLEDGMENTS

It would have taken even more time and effort to finish this thesis without the help and presence of a number of people.

One of the professors convincing me to dig deeper and think more critically was Oscar Gelderblom. My debt of gratitude to him is incredibly big, not just for inspiring me intellectually but also for keeping faith in my abilities, for hiring me as his research assistant and for not kicking me off the research master's programme.

I would also like to thank, of course, my thesis supervisor Liesbeth van de Grift with whom I had the pleasure of working for the past year. She warned me it would be hard to write a thesis next to doing an internship. She was absolutely right, and without her critical guidance along the road, I would not have finished my thesis before September 2019.

Many others deserve a word here. I had a wonderful time with my friends at Odysseus and UUMUN during the master's programme. They reminded me that there is much more out there than studying.

My library buddies Jasper, Hans, Guus, Duco, and Alessandra set a great example by finishing their own theses, and offered great relief when I needed a break in the library.

I want to thank my colleagues at the Ministry of Foreign Affairs who introduced me to the world of diplomacy during my internship.

And lastly, of course, I thank my family and (former) roommates who accepted that I annoyingly asked for attention during many stages of the process.

Paul Schilder

TABLE OF CONTENT

INTRODUCTION.....	3
I. COMING OF AGE: DEVELOPMENT AID BETWEEN 1964 AND 1977.....	11
II. AID EVALUATION ACROSS THE BORDER, 1966-1975.....	39
III. FROM IDEA TO PRACTICE: INSTITUTIONALIZING DUTCH AID EVALUATION, 1964-1977.....	59
CONCLUSION.....	85
LIST OF ABBREVIATIONS.....	90
BIBLIOGRAPHY.....	91

INTRODUCTION

'Do good and do not look back.'¹ With this characterization historians Peter Malcontent and Jan Nekkers introduced their study of fifty-year history of Dutch development cooperation in 1999. The study indicated that aid evaluation at the Dutch Ministry of Foreign Affairs (MFA) was scarce and played only a limited role in policy changes. Eight years later, former Minister for Development Cooperation Jan Pronk (1973-1977, 1989-1998) replied in a speech for the Evert Vermeer Stichting.² Aid evaluation existed, he argued, at least since the 1960s. Moreover, Pronk emphasized that he had been an important figure in the establishment of aid evaluation in the Netherlands. He was a member of the Werkgroep Janssen, a group of university researchers commissioned by the MFA in 1965 to conduct the first overall evaluation of Dutch aid policy. Pronk allegedly followed the recommendations of the Werkgroep when he became Minister in 1973.³ Despite his serious counterarguments, Pronk's speech was not picked up by historians, nor has there been a study on these early days of Dutch aid evaluation.⁴

As a matter of fact, more critics have pointed out a lack of evaluation within the aid sector. In 2006 former World Bank economist William Easterly published the bestselling *The White Man's Burden*. Easterly denounces development aid for its lack of feedback from below and accountability from above. He explains his argument by a simple dichotomy: foreign aid employs planners where it requires searchers. Planners, a stereotype for

visionaries, celebrities, presidents, chancellors of the exchequer, bureaucracies, (...) announce good intentions but don't motivate anyone to carry them out; Searchers find things that work and get some reward. Planners raise expectations but take no responsibility for meeting them; Searchers accept responsibility for their actions. Planners determine what to supply; Searchers adapt to local conditions.⁵

¹ 'Doe wel en zie niet om.' In: P.A.M. Malcontent, J.A. Nekkers, 'Inleiding. "Doe wel en zie niet om"' in: J.A. Nekkers, P.A.M. Malcontent (eds.), *De geschiedenis van vijftig jaar Nederlandse ontwikkelingssamenwerking 1949-1999* (The Hague 1999) 11.

² Former Dutch ngo, allied with the Labour Party and currently incorporated into the Foundation Max van der Stoep. The Evert Vermeer Stichting was a think tank and lobby organization for social democratic aid policy.

³ J.P. Pronk, 'Doe wel en zie niet om' (1 February 2007) <http://www.janpronk.nl/speeches/dutch/doe-wel-en-zie-niet-om.html> (consulted 12 March 2019).

⁴ Interestingly, the efficiency and effectiveness of development aid frequently resurfaces in Dutch media, e.g.: J. Born, "Effect van miljoenen aan ontwikkelingshulp moet beter worden onderzocht" (9 July 2019) <https://eenvandaag.avrotros.nl/item/effect-van-miljoenen-aan-ontwikkelingshulp-moet-beter-worden-onderzocht/> (consulted 12 August 2019); Dr. Kelder en Co, 'Het "rendement van ontwikkelingssamenwerking', (1 December 2018) <https://www.nporadio1.nl/dr-kelder-en-co/onderwerpen/481922-het-rendement-van-hulp> (consulted 12 August 2019).

⁵ W. Easterly, *The White Man's Burden. Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good* (New York 2006) 5-6.

Among the planners is, for example, Jeffrey Sachs, Columbia University professor and special advisor to the United Nations (UN) Secretary-General. Easterly takes Sachs' plan for a Big Push to end poverty as an example of the simplicity of thought present in the development community. In Easterly's view such a plan is doomed for failure and, moreover, failure would not result in resignations and dismissals. After all, evaluations are carried out not by independent researchers but by development agencies themselves, leading them to overestimate the successes of aid.⁶ Development scholar Arturo Escobar added that when evaluation occurs, it takes 'place as part of a process that is largely self-referential, to the extent that these documents are written not to illuminate a given problem but to ensure their insertion into the ongoing flow of organizational texts.'⁷ There is, Escobar argued, no honest and thorough communication between evaluators, policymakers, and local communities.

More recently, scholars have focused on the way these actors have engaged with each other and cooperated in policymaking and policy implementation. Scholars have also put more effort in understanding the actual impact of policies.⁸ This renewed focus shows that development stands for various ideas and practices that have always been highly contested both inside policymaking circles, among officials implementing policies on the ground, and within local communities. Historian Marc Dierikx follows this new wave of scholarship in studying the actual implementation of Dutch aid on the ground.⁹ He concludes that policymakers in The Hague were often unaware of the local context, execution and results of individual projects. Officials executing the projects then discovered that putting ideas into practice proved to be much harder than expected.¹⁰ In order to learn from such experiences and to hold relevant Ministers accountable, commentators promoted the institutionalization of evaluation. At the Dutch MFA, evaluation

⁶ Easterly, *The White Man's Burden*, 3-4.

⁷ A. Escobar, *Encountering Development: the making and unmaking of the Third World* (Princeton 1995) 112.

⁸ E.g. C. Unger, *Entwicklungspfade in Indien. Eine internationale Geschichte, 1947-1980* (Göttingen 2015).

⁹ Dierikx also edited a publication of sources on Dutch development aid: *Nederlandse ontwikkelingssamenwerking: bronnenuitgave*, red. M.L.J. Dierikx (Den Haag 2002-2009). Other scholars writing on the history of Dutch development aid are, e.g., P. Hoebink, *Geven is nemen. De Nederlandse ontwikkelingshulp aan Tanzania en Sri Lanka* (Nijmegen 1988); E.H. Arens, 'Multilateral Institution-Building and National Interest: Dutch development policy in the 1960s' *Contemporary European History* 4 (2003) 457-472, 458; J.J.C. Voorhoeve, *Peace, Profits and Principles. A Study of Dutch Foreign Policy* (Leiden 1985); M. Kuitenbrouwer, *De ontdekking van de Derde Wereld: beeldvorming en beleid in Nederland, 1950-1990* (Den Haag 1994); Nekkers, Malcontent (eds.), *De geschiedenis van vijftig jaar Nederlandse ontwikkelingssamenwerking*.

¹⁰ M.L.J. Dierikx, 'Like a fish out of water: experimenting fishery in Lake Victoria, 1964-1974' *Special Working Paper Series on 'Unintended Effects of International Cooperation'* 1 (2017) 6. Dierikx further suggests that the corporate sector obtained project contracts quite easily, since ministries involved in development policy had no interest in assessing results. He notes that historians should look into this matter to see whether this was actually the case.

finally emerged in the form of the *Inspectie Ontwikkelingssamenwerking te Velde* (IOV), established by Minister Pronk in 1977.

Nonetheless, most historical studies have focused on their own evaluations of Dutch foreign aid. One exception is the history of the IOV, written by investigative journalist Jos van Beurden and historian Jan-Bart Gewald in 2003.¹¹ Van Beurden and Gewald focused on the proceedings of the IOV since its inception, but they did not explain why the IOV was established in the first place, nor what ideas and practices concerning aid evaluation circulated in the Netherlands and the international development community before 1977. The existing literature still lacks an assessment of the academic and governmental debates leading to the institutionalization of aid evaluation. Its findings on aid evaluation in the 1960s and 1970s lack historical evidence and in-depth analysis of the international and national discussions on aid evaluation. We do not know how and why aid evaluation actually came into being. Therefore, this thesis aims to understand why aid evaluation at the Dutch MFA developed the way it did. In order to answer this question, it is important to focus on three debates on aid evaluation in the national and international development community between 1964 and 1977. These debates, occurring between the establishment of bilateral aid in the Netherlands (1964) and the creation of the IOV (1977), most significantly shaped the institutionalization of aid evaluation at the Dutch MFA.

Firstly, actors discussed the purposes of aid evaluation. They distinguished between two purposes of aid evaluation: lesson-learning and accountability.¹² Lesson-learning is the expansion of knowledge and the future use of such knowledge at all levels within the MFA. Through lesson-learning, evaluation is an instrument to improve the efficiency and effectiveness of aid. Thus, evaluation for the purpose of lesson-learning is for internal use. If evaluation serves accountability, on the other hand, it is by definition external. In such a case, parliament has the power to hold the Minister accountable for the results of his or her aid policies. It is also possible that not parliament but the government pushes for evaluation. The government could decide to publish information on aid results either to influence public opinion deliberately, or as a legal or moral duty to provide honest and objective information that also includes shortcomings of aid.

Secondly, actors debated the question whether evaluation should be internal or external to the MFA and other aid agencies. Many commentators in the 1960s and 1970s emphasized the need for impartiality among evaluators. Evaluation conducted by staff of aid agencies¹³

¹¹ After renaming the IOV became the *Directie Internationaal Onderzoek en Beleidsevaluatie* (IOB). J. van Beurden, J.B. Gewald, *From Output to Outcome? 25 Years of IOB Evaluations* (Amsterdam 2004).

¹² E.g. B.E. Cracknell, *Evaluating Development Aid: Issues, Problems and Solutions* (London 2000) 54-66; T. Brodin, 'Conceptualizations of "results" in Swedish policy for development cooperation from the 1960s to the 2000s' *Bulletin of Geography. Socio-economic Series* 33 (2016) 7-22.

¹³ Including multilateral organizations, NGOs and national aid agencies. Note that in some countries, unlike in the Netherlands, the MFA is not responsible for the planning and execution of aid policy. Countries have either a separate aid ministry or semi-autonomous aid agencies. This applies for example

themselves would lack the impartial judgment required for both lesson-learning and accountability. Internal evaluation, also called built-in evaluation, is conducted by staff of operational departments as part of the regular policy cycle. External evaluation can be conducted by independent evaluators within the MFA staff or by consultants or academics outside the MFA hired for a particular evaluation.

Thirdly, actors discussed options for international cooperation in matters of aid evaluation. Two options were deemed most feasible. Joint evaluation is bilateral evaluation, and thus requires the participation of various stakeholders in evaluation of a particular project or programme. In addition to representatives of the donor government, stakeholders are often the residents of the country where the project is implemented. Such residents can be representatives of all levels of the recipient government, the private sector, the civil society or simply individual citizens somehow involved in the project. Furthermore, actors discussed possibilities of organizing aid evaluation multilaterally. Organizations such as the UN, World Bank or the Organisation for Economic Cooperation and Development (OECD) would be in a better position to evaluate aid results than national governments. Moreover, multilateral evaluation would also resolve concerns of impartiality.

Methodology and sources

The 1960s and 1970s were characterized by the growing power of technocratic governance. Political scientist Peter Haas mentions a global growth from 850 ministries in 1950 to 2500 in 1975. The number of ministries grew parallel to the establishment of planning and policy research institutes. Bureaucracies expanded, leading to the emergence of ‘the policy role of the knowledge elite’, as sociologist Dorothy Nelkin notes.¹⁴ Civil servants regarded themselves not as representatives of the people, but increasingly as experts in a specific field of policy.¹⁵ They came to be known the fourth power in Montesquieu’s *trias politica*.¹⁶ In addition, to civil servants at ministries, this fourth power also consists of international civil servants shaping national policy agendas and policymaking. This applies to an even larger extent to policy fields that are international by definition, such as development aid. Within the development, civil servants of organizations such as the UN, the World Bank and the OECD have enormous financial or

to the US, UK, Sweden, Norway, Denmark, Canada and Germany. See, e.g., O. Stokke (ed.), *Western Middle Powers and Global Poverty. The Determinants of the Aid Policies of Canada, Denmark, the Netherlands, Norway and Sweden* (Uppsala 1989).

¹⁴ D. Nelkin, ‘Scientific Knowledge, Public Policy, and Democracy’ *Knowledge, Creation, Diffusion, Utilization* 1 (1979) 106-122, 107; P.M. Haas, ‘Introduction: Epistemic Communities and International Policy Coordination’ *International Organization* 1 (1992) 1-35, 8-10.

¹⁵ J.D. Aberbach, R.D. Putnam, B.A. Rockman, *Bureaucrats and Politicians in Western Democracies* (Cambridge, MA 1981).

¹⁶ M. Bovens, *De Vierde Macht Revisited. Over ambtelijke macht en publieke verantwoording* (Utrecht 2000).

normative power. The importance of expert communities arises from the fact that they possess policy-relevant knowledge, though more importantly, their knowledge is recognized as authoritative in the field. Since the expert communities in the field of development aid often work transnationally, they influence public policy in multiple countries.¹⁷

When aid evaluation emerged during the 1960s and 1970s, the general idea of policy evaluation was relatively new in public administration and academia. Evaluation had just become part of the new policy planning cycle of a number of western countries. The lack of subject knowledge automatically led to experimentation with aid evaluation and intensive communication between government circles and expert communities. Evaluation produced knowledge with the ability to directly influence the approach of development actors in government, politics or academia, either through formal procedures or personal experiences.¹⁸ Due to the absence of subject knowledge, it is important to realize that civil servants need time to transform political ideas into feasible practices before policies are finally institutionalized. In such a period, civil servants test hypotheses and conduct experiments. Piet Hein Donner typified this as policy intimacy (*beleidsintimiteit*) when he was the Dutch Minister of the Interior in 2011. Policy intimacy entails the possibility to experiment with new ideas in public administration with confidentiality, without interference of the general public through parliament and media. It is characterized, therefore, as a period of very limited accountability.¹⁹ However, since development aid in general was a relatively young government practice in the 1960s and 1970s, civil servants might have had serious considerations against transparency at all.

This thesis typifies the 1960s and the 1970s as the period of policy intimacy concerning aid evaluation and aims to uncover the tension between policy intimacy and public accountability. It does so by analysing three types of source material. The first type of source material are memos from the Dutch MFA. They demonstrate the lively debates concerning aid evaluation among the Ministers for Development Cooperation and their staff members. The memos range from 1964 to 1978 and mainly include staff members in The Hague, since they were most involved in the institutionalization of aid evaluation. Therefore, this thesis did not analyse sources from Dutch missions.²⁰ The only exceptions are reports and letters from Dutch missions that were included in the archival inventories on aid evaluation. Moreover, this thesis does not assess the actual evaluations conducted during the given period of time, since the analysis only focuses on the debates leading up to the institutionalization of evaluation. Lastly, aid evaluation

¹⁷ P.M. Haas, 'Introduction: Epistemic Communities and International Policy Coordination' *International Organization* 1 (1992) 10-11.

¹⁸ C. Unger, *International Development. A Postwar History* (London 2018) 3-4, 10.

¹⁹ S. Keulen, *Monumenten van Beleid. De wisselwerking tussen Nederlands rijksoverheidsbeleid, sociale wetenschappen en politieke cultuur, 1945-2002* (Hilversum 2014) 9-11; further definition in P. Frissen, *Het geheim van de laatste staat. Kritiek van de transparantie* (Amsterdam 2016) 128-130.

²⁰ Embassies, consulate generals and permanent representations.

also occurred in the governance of Dutch NGOs.²¹ This thesis only focuses on the evaluation procedures inside the MFA.

Secondly, the thesis focuses on published sources, such as parliamentary debates and newspaper articles, to assess the public debate on aid evaluation. These include the minutes of both the Tweede and the Eerste Kamer. All newspapers analysed here had national coverage and are accessible on Delpher, such as De Telegraaf, De Volkskrant, Het Parool and Trouw. Moreover, the *Nationale Adviesraad inzake de Ontwikkelingshulp* (NAR), a civil society advisory group for the MFA, published a number of recommendations on aid evaluation that reveal the broader public opinion on the topic. The analysis excludes party manifestos and other party documents, assuming that parliamentary debates are more useful to assess the impact of political parties on policymaking.

The first two types are Dutch sources. The case of the Netherlands is highly relevant, despite the modest size of its aid budget in absolute numbers. For several years, the Netherlands actually topped the list of biggest contributor of net Official Development Assistance (ODA) relative to its gross national income.²² The Netherlands also had a strong presence in multilateral institutions engaged in development. It was an early member state of the OECD, which paid great attention to aid evaluation. Furthermore, Dutch research institutes have played an important role in both policy making and policy analysis compared to foreign counterparts.²³ In aid policy too, numerous experts in economics and agricultural science obtained key positions in advisory councils on Dutch aid policy, such as Jan Tinbergen and Egbert de Vries.

Thirdly, to demonstrate the international context of Dutch experiments in aid evaluation, this thesis analyses reports of international conferences on aid evaluation and evaluations conducted by consultants and academics. These include the reports of four conferences on aid evaluation in Berlin (1966), Paris (1968), Wassenaar (1970) and Amsterdam (1973). In addition, the source material consists of a selection of studies on the institutionalization of aid evaluation within national and international governance. These include studies of the OECD, the UN and independent consultants. The justification of this selection is given in Chapter II. The analysis excludes archival material of multilateral organizations and national aid agencies other than the Dutch MFA. Moreover, the analyse does not cover the large body of scientific literature on the

²¹ E.W. Hommes, 'Aid Evaluation in the Netherlands' in: Stokke (ed.), *Evaluating development assistance*, 154-155.

²² M. Kuitenbrouwer, 'Nederland gidsland? De ontwikkelingssamenwerking van Nederland en gelijkgezinde landen, 1973-1985' in: Nekkers, Malcontent (eds.), *De geschiedenis van vijftig jaar Nederlandse ontwikkelingssamenwerking*, 183-200.

²³ F. van Nispen, P. Scholten, 'Introduction' in: F. van Nispen, P. Scholten, *Policy analysis in the Netherlands* (Bristol 2014) 1-4; S.S. Blume, R.P. Hagendijk, A.A.M. Prins, 'Political culture and the policy orientation in Dutch social science' in: P. Wagner (ed.), *Social sciences and modern states: national experiences and theoretical crossroads* (Cambridge 1991) 168-190; E. Hueting, *De permanente herstructurering in het welzijnswerk* (Zutphen 1989) 337-351.

actual conduct of aid evaluation. Questions of how to measure social development, how to trace the multiplier effect of financial investment and many other key elements of development aid are irrelevant to answer the main research question.

Definitions

Firstly, it is important to note that the definition of evaluation was itself part of the discussion in the 1960s and 1970s. Political scientist Evert Vedung coined a minimal definition that is, anachronistically, the common denominator of the different opinions in the 1960s and 1970s. According to Vedung,

[e]valuation is a nebulous concept (...) defined as careful retrospective assessment of public-sector interventions, their organization, content, implementation and outputs or outcomes, which is intended to play a role in future practical situations.²⁴

By 'retrospective assessment', Vedung defines evaluation as an *ex post* activity, thus taking place after termination of a project or programme. This thesis uses the definition since it excludes *ex ante* evaluation, commonly called 'appraisal' in the literature and the source material. *Ex ante* evaluation or appraisal often had the purpose of studying the feasibility or profitability of a project or programme.²⁵ Expert communities in the 1960s and 1970s furthermore included progress reports – nowadays called monitoring – in their definition. Progress reports are written during the implementation phase of a project or programme. As Chapters II and III will point out, especially built-in evaluation contained such progress reports. These documents often had the same purposes as *ex post* evaluation, i.e. lesson-learning and/or accountability. Lesson-learning and accountability are already defined above.

It should be emphasized that evaluation is different from audit. While audit serves for financial and management control, evaluation assesses outputs, outcomes and impacts. Evaluation studies results and analyses what works, why and how.²⁶ In other words, 'auditing begins from a normative position of asking whether or not particular rules, procedures, or mandates have been followed. Evaluation begins from a different set of assumptions – testing

²⁴ E. Vedung, 'Four Waves of Evaluation Diffusion' *Evaluation* 3 (2010) 263-277, 264.

²⁵ Some appraisals were conducted in order to create benchmarks for progress reports or *ex post* evaluations, which is why can be argued that these appraisals served accountability and lesson-learning too. However, these appraisals will not be considered in this thesis because such appraisals did not study implementation and results. This makes appraisal fundamentally different than progress reports and *ex post* evaluation.

²⁶ C. Wisler (ed.), 'Evaluation and Auditing: Prospects for Convergence' *New directions for evaluation* 3 (1996) 1-67.

hypotheses, ruling out competing explanations, and attempting to answer the causal question'.²⁷ In contrast to auditing, evaluation focuses on effectiveness and efficiency. Effectiveness is the extent to which a project or programme has reached its *a priori* stated objectives. Efficiency is the extent to which a project or programme has reached appropriate results in relation to the amount of used resources.

Lastly, the term 'aid' is preferred over 'cooperation' due to two reasons. Nowadays, 'aid' is used as the encapsulating term for humanitarian activities – food aid, medical aid et cetera, mostly in the context of a natural disaster or conflict-ridden area. Cooperation, on the contrary, is not used for direct relief but rather for long-term social and economic development. Until the 1970s, however, the common term for both activities was aid – development aid, foreign aid, international aid or other denominations. In addition, aid evaluation is much easier to read than cooperation evaluation.

Outline

The research will be presented as followed. Chapter I focuses on general trends in the history of development between 1964 and 1977, based on a study of the existing literature. Chapter I analyses who the important actors in the development community were, as well as the major policy changes. These questions will be answered in the international context first, after which the thesis focuses on these questions within the Dutch context. Chapter II aims to answer the question how evaluation ideas were discussed in international conferences and academic publications. Chapter III specifically focuses on the discussion on aid evaluation in the Netherlands, at the MFA, in parliament and the written media. The chapter is chronologically split up into three parts: the preparation and report of the Werkgroep Janssen, its reception at the MFA and in the public sphere, and the initiatives taken during Pronk's tenure at the MFA. Lastly, the conclusion sums up the answers to the question central to this thesis, and provide suggestions for further research.

²⁷ C. Rist (ed.), *Program Evaluation and the Management of Government. Patters & Prospects across Eight Nations* (New Brunswick 1999) 8.

I. COMING OF AGE: DEVELOPMENT AID BETWEEN 1964 AND 1977

Er is geen reden tot pessimisme; men kan slechts tot de slotsom komen dat er nog veel werk te doen valt.²⁸

- Ferdinand van Dam, 1964.

As mentioned in the introduction to this thesis, the past decade saw a new wave of historical research on development. Roughly twenty years earlier, postdevelopment literature emerged out of a fundamental critique on development among academics, politicians and the general public. Development was criticized because, allegedly, it did not deliver. Many recipient countries had, for instance, immensely indebted themselves in the late 1970s and 1980s due to the global economic recession. Development became the scapegoat, denounced as a holistic concept that implied the top-down restructuring of society.

Most postdevelopment scholars were mainly interested in the intellectual history of development. Nowadays, however, research on other subtopics of development history is increasingly recognized essential for understanding development's complex and practical nature.²⁹ Current scholarship goes 'deeper' and 'wider', historian Joseph Hodge notes. Through 'deeper' research in local archives of recipient countries scholars aim to discover how aid policy was implemented on the ground. Research is 'wider', as scholars tend to focus on more than just American and Western European foreign aid. Scholars now study development policies of the Soviet Union and Eastern European states, as well as initiatives from recipient countries themselves.³⁰ Historian Monica van Beusekom, for example, demonstrates in her history of the Office du Niger that French development officials were unable to implement the policies they and their colleagues in Paris had devised because of opposition by the local population. The actual development practice in the region was an outcome of negotiations between local farmers and French experts. In conclusion, Van Beusekom comments that scholars cannot understand how development ideas win or lose support without taking local experiences into consideration. Feedback from below – from local communities and development practitioners - is incredibly valuable for historical research.³¹ Therefore, the new wave of scholarship shows that the history

²⁸ F. van Dam, *Onbehegen rond de ontwikkelingshulp* (Groningen 1964) 20.

²⁹ E.g. C.A. Bayly, V. Rao, S. Szreter, M. Woolcock (eds.), *History, Historians and Development Policy* (Manchester 2011).

³⁰ J.M. Hodge, 'Writing the History of Development (Part 2: Longer, Deeper, Wider) *Humanity 2* (2016) 125-174.

³¹ M. van Beusekom, *Negotiating Development: African Farmers and Colonial Experts at the Office du Niger, 1920-1960* (Oxford 2002).

of development is characterized by contest, rather than homogeneity, between ideas, practices and actors. The new literature shows that development is not a holistic, inherently bad idea created in Europe and implemented in the developing world. New scholarship rather makes use of a non-normative perspective to analyse development. Such studies are brought together in a number of general histories on development that were recently published.³²

These works, combined with the existing Dutch literature, provide the basis for this chapter. The chapter illustrates the context of international and Dutch development aid in which the experiments of aid evaluation were embedded. Section I examines the changes in the international context of development between 1964 and 1977, looking at the main actors engaged in international development and the major policy changes during that period. Section II outlines the history of Dutch development and focuses on the same questions as the first section. Both sections include an analysis of the existing knowledge on the history of public policy evaluation. Despite the richness of the literature written on development history, the thesis' focus on Dutch development aid limits the depth of this chapter's analysis of the international context. Nonetheless, by the end of this chapter the reader should have a general overview of trends in purposes, actors, policies and evaluation in the early development community.

1.1 INTERNATIONAL DEVELOPMENT FROM POINT FOUR TO MATURITY

For a long time scholars designated 1949 as the start of international development as we know it today. In that year, US President Harry Truman announced the Point Four Program. Point Four embodied the provision of technical assistance to countries in Africa, Asia and Latin America, encapsulated in a 'bold new program for making the benefits of our scientific advances and industrial progress available for the improvement and growth of underdeveloped areas.'³³ Through Point Four, the US set the tone for a new era of development aid. The UN embarked on a similar campaign later that year, when they announced their Expanded Programme for Technical Assistance (EPTA).³⁴ In the decades that followed, development aid transformed into an industry handling substantial sums of money. If we want to understand why aid evaluation emerged within such an environment, we must highlight the influential persons who shaped the international

³² Unger, *International Development*; Macekura, Manela (eds.), *The Development Century*.

³³ Harry Truman, 'Inaugural Address' (edition 20 January 1949)

<http://www.let.rug.nl/usa/presidents/harry-s-truman/inaugural-address-1949.php> (15 August 2019).

³⁴ EPTA was the first, centrally coordinated, programmatic expression of the UN Charter's paragraphs 55 and 56, which state that member states 'shall promote higher standards of living, full employment, and conditions of economic and social progress and development'. Before EPTA, many of the UN's development agencies had already come into existence. Examples include the United Nations Children's Fund (UNICEF), the United Nations Relief and Works Agency for Palestine Refugees (UNRWA) and the Food and Agricultural Organization (FAO). Organizations such as the International Labor Organization (ILO), the International Monetary Fund (IMF) and the World Bank were also founded earlier, but were later incorporated under the UN umbrella. One of EPTA's objectives was to bring central coordination to the full scope of UN development activities.

development community, in both thought and practice, and what the various policies of the 1960s and 1970s generally entailed.

1.1.1. The actors in international development

Turning to the actors in the international development community, we must first distinguish between the channels of aid. Aid is either bilateral, directly from government to project or indirectly through implementing partners³⁵, or multilateral, by contributions to international organizations such as the UN. If we exclude colonial development policies, most foreign aid during the 1950s was given through multilateral channels. Bilateral aid often replaced colonial development policies after the wave of decolonization around 1960. At the same time, many donor governments established or expanded their aid programmes by founding separate ministries or aid agencies – the US (Agency for International Development, USAID), Sweden (Swedish International Development Cooperation Agency, SIDA), the UK (Department for International Development, DFID) and Germany (Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, BMZ).³⁶ Most donor governments gave the majority of their bilateral aid to their former colonies. Countries with few colonial territories, such as the Nordics, spent a relatively large percentage of their aid through multilateral channels. The US, being the largest donor during the 1950s, was the exception to this and mostly preferred bilateral aid.

A number of multilateral institutions are important to mention here, since they will be discussed extensively in Chapters II and III. The introduction already referred to EPTA, which incorporated into the United Nations Development Programme (UNDP) late 1965. The establishment of UNDP was an impetus for the reorganization of multilateral aid. In financial terms, the World Bank was the biggest multilateral agency, especially when taking account its subsidiary, the International Development Association (IDA). The latter provided soft-loans particularly intended for development projects in the Global South. The World Bank was not just a key player in financial terms: the organization also invested in the development of aid evaluation. This led the Bank to adopt the Standards and Procedures for Operations Evaluation in 1976.³⁷

Two other institutions were important, though not necessarily in financial terms. The UN Committee on Development Planning (CDP³⁸) and the OECD Development Assistance Committee (DAC) set the tone as fora where donor governments discussed development ideas and policies.

³⁵ Most such aid is nowadays given through NGOs.

³⁶ D.C. Engerman, 'The Romance of Economic Development and New Histories of the Cold War' *Diplomatic History* 1 (2004) 23-54; M.L.J. Dierikx, 'Policy versus Practice. Behind the Scenes in Dutch Development Aid, 1949–1989' *The International History Review* 4 (2017) 638-653; 641.

³⁷ T. Chianca, 'The OECD/DAC Criteria for International Development Evaluations: An Assessment and Ideas for Improvement' *Journal of Multidisciplinary Evaluation* 9 (2008) 41-51, 41.

³⁸ Later renamed Committee on Development Policy.

Despite the CDP's continuing activity inside the UN system, little research is available on its history. The CDP was created in 1966 as a body of independent academics that reviewed and steered worldwide development efforts. Jan Tinbergen was named its first chair and retained that position until 1972. In 1970, the CDP published its review of progress during the 1960s – the First Development Decade – and its recommendations for development policy in the 1970s – the Second Development Decade. This report, colloquially named the Tinbergen Report, became an internationally renowned document that formed the basis for negotiations on the Second Development Decade strategy within the UN. It will be discussed in the analysis of Chapter II.

The DAC was established in 1961³⁹ to enhance the efficiency and effectiveness of aid through joint donor cooperation. Among western countries it was and still is known 'as an "international think tank", serving as a hub of knowledge, and a "policy path finder" for its members and the international community at large.'⁴⁰ The DAC had a significant impact on development thought and practice until at least the mid-1970s, comparable to the impact of the World Bank and UNDP.⁴¹ Historians sometimes neglect the organization because of its lack of hard power. However, the DAC influenced aid policies of its member states, particularly through the introduction of standards, data generation and donor peer reviews. The DAC created tools to better measure the quantity of aid flowing from North to South, for instance by developing the concept of Official Development Assistance (ODA).⁴² Prior to these initiatives, donor governments were easily able to overestimate the size of their aid budget and use that overestimation as leverage during negotiations with other donors, recipient countries, multilateral organizations and even their own parliaments. Eventually, in the late 1970s and early 1980s, the DAC became a catalyst for the professionalization of aid evaluation. This led to the establishment of the DAC's Criteria for International Development Evaluations in 1991.

In addition, non-governmental organizations (NGOs) and philanthropic organizations participated in development aid since pre-war colonial times. The American government even turned to private institutions when it expanded its aid budget after the Second World War. After

³⁹ The DAC was the successor of the Development Assistance Group, founded in 1960.

⁴⁰ C. Ydsen, J. Verschaeve, 'The OECD Development Assistance Committee and Peace: Instituting Peace by Economic Means' in: A. Kulnazarova, V. Popovski (eds.), *The Palgrave Handbook of Global Approaches to Peace* (Basingstoke 2019) 479.

⁴¹ M. Schmelzer, 'A Club of the Rich to Help the Poor? The OECD, "Development", and the Hegemony of Donor Countries' in: M. Frey, S. Kunkel, C.R. Unger (eds.), *International Organizations and Development 1945-1990* (Basingstoke 2014) 171-187.

⁴² 'ODA consists of flows to developing countries and multilateral institutions provided by official agencies, including state and local governments, or by their executive agencies, each transaction of which meets the following test: a) it is administered with the promotion of economic development and welfare of developing countries as its main objective, and b) it is concessional in character and contains a grant element of at least 25 per cent (calculated at a rate of discount of 10 per cent).' In: OECD, *The Story of Official Development Assistance. A History of the Development Assistance Committee and the Development Co-operation Directorate in Dates, Names and Figures* (Paris 1996) 24.

all, private institutions had been working in recipient countries for decades and possessed a level of knowledge and experience unmatched in government circles.⁴³ Historian Corinna Unger shows that in the 1960s, when agricultural development programmes gained prominence, ‘most of the research on which the Green Revolution technology rested was funded by private companies and by non-governmental organizations, particularly by the Rockefeller Foundation.’⁴⁴ From the 1960s onwards, NGOs took on a much larger role than they had previously, and gave development a new image.⁴⁵ British NGOs remained active and jumped in to fill the void after the independence of the UK’s former colonies.⁴⁶ A decade later, USAID chose to invest in partnerships with both the corporate sector and NGOs. Other donor governments copied this policy shift. In total, between 1973 and 1980 OECD member states increased their funding for NGOs by roughly 90 percent.⁴⁷

Other powerful actors in the development community were academics. The 1950s and 1960s were the heyday of the belief that science and planning could overcome social and economic policy problems. The narrative was popular across the globe. Decolonized countries aimed to become equals to those in the Global North through centrally planned rapid industrialization based on rational and scientifically proven methods.⁴⁸ The concept of gross national product (GNP) became an important policy tool for multilateral agencies and governments all over the globe to measure a country’s economic progress. GNP’s popularity was partly fostered by the UN’s initiative to standardize national accounting in 1953. GNP enabled measuring economic development and became the most important indicator of policy success.⁴⁹

Academics also became involved in policymaking. For example, scholars from Harvard and MIT largely contributed to American foreign aid policy.⁵⁰ Walt Whitman Rostow, the historian and economist from MIT who wrote about ‘the five stages of economic growth’, served as an important foreign policy advisor to Eisenhower, Kennedy and Johnson. His role has been analysed by scholars time and again. ‘Scientific optimism’, which emerged in the 1950s and 1960s, ensured that scholars such as Rostow could take centre stage as ‘public intellectuals’.⁵¹ ‘Liberal anti-communism’ served as the basis of thinking for development economists such as Max F. Millikan,

⁴³ D. Ekbladh, *The Great American Mission: Modernization and the Construction of an American World Order* (Princeton 2009) 21.

⁴⁴ Unger, *International Development*, 113.

⁴⁵ J. Adelman, ‘Epilogue: Development Dreams’ in: Macekura, Manela (eds.), *The Development Century*, 337.

⁴⁶ M. Hilton, ‘Charity and the End of Empire: British Non-Governmental Organizations, Africa and International Development in the 1960s’ *The American Historical Review* 2 (2018) 493-517.

⁴⁷ Macekura, ‘Point Four and the Crisis of U.S. Foreign Aid Policy in the 1970s’, 86-89.

⁴⁸ Unger, *International Development*, 8-9.

⁴⁹ S.J. Macekura, ‘Development and economic growth: an intellectual history’ in: I. Borowy, M. Schmelzer (eds.), *History of the Future of Economic Growth: Historical Roots of Current Debates on Sustainable Degrowth* (Abingdon 2017) 114-115.

⁵⁰ N. Gilman, *Mandarins of the Future: Modernization Theory in Cold War America* (Baltimore 2003).

⁵¹ Unger, *International Development*, 7.

Lucian Pye and Rostow, who 'felt (...) that their knowledge could help to improve and strengthen the position of the United States in the Cold War.'⁵²

In Europe, academic centres focused on development were founded in the 1950s and 1960s as well. The Dutch government established the International Institute of Social Studies in 1952. Similarly, the University of Sussex founded the Institute of Development Studies in 1966. The Soviet Union also focused on development research, for example at the Institute of Oriental Studies at the Soviet Academy of Sciences. Outside Europe, India set up the Centre for the Study of Developing Societies in 1963, and Tanzania had the Institute of Development Studies at the University of Dar es Salaam a decade later. In addition to the new centres, academics also founded a large number of journals dedicated to development studies.⁵³ To conclude, it can be argued that the coming-together of academia and development was truly a global movement:

[I]t is crucial to remember that development as a field was distinctly international and interdisciplinary in character. Individual experts, many of them academics – engineers, economists, sociologists, geographers, anthropologists, demographers, historians, urban planners, agronomists – produced the knowledge and wrote the textbooks that became the theoretical basis of development practice.⁵⁴

The development community is and already was too extensive in the 1960s and 1970s to be discussed entirely in this thesis. To narrow the scope, I left out development policies of the EEC and the many regional development banks. They unquestionably left their mark on the shape of development. The same applies for UN Conference on Trade and Development (UNCTAD), *the* forum for discussions on international trade and development in the 1960s and 1970s. However, as Chapter II and III will demonstrate, discussions on aid evaluation often took place within the institutional frameworks of other organizations, such as the UNDP, World Bank, CDP and DAC.

1.1.2. Major changes in international aid policy

Throughout the 1960s and 1970s, new insights and experiences shaped the aid policies of the actors named above. Scholars generally agree that the aid community came under more public scrutiny by the end of the 1960s. Commentators tried to understand why the results of the First Development Decade did not meet expectations and why modernization theory, a grand plan to holistically transform society, failed to deliver the 'sweeping changes' it had promised. Some

⁵² Unger, *International Development*, 7.

⁵³ Ibidem, 74-77; A.S. Park, 'Does Development Learn from History?' *World Development* (2017) 52-64.

⁵⁴ Unger, *International Development*, 7.

historians even speak of the first crisis of development.⁵⁵ But what did aid policies in the 1960s look like in the first place? And what implications did the crisis have for aid policies in the 1970s?

To answer these questions, an outline of the phenomenon of modernization theory is necessary. Modernization theory is often attributed to early American aid policies, where it served the American strategic interests during the Cold War. In the minds of US policymakers, modernization theory gave liberalism the objectives and tools to counter similar, grotesque aims set by fascism and communism.⁵⁶ But modernization theory was much more than a (geo)political instrument. It was a set of beliefs, founded on scientific research, that centrally planned interventions would catalyse rapid catch-up growth for the developing world. Modernization theory in the aid sector was fundamentally based on the idea that interventions should mostly include technical and financial assistance for large scale projects in infrastructure and manufacturing. Social interventions, such as programs for direct poverty alleviation, were not included in the thoughts of modernization theorists:

Social scientists of the Cold War period emphasized that a concentrated “big push” of directed investment and foreign aid would allow postcolonial societies to reach the crucial “take-off ” point, after which they would enter the period of “self-sustaining growth,” ready and able to advance without recourse to external help.⁵⁷

It is important, however, to look beyond American aid policies to identify the global impact of the narrative. DAC member states also drew on modernization theory and on Rostow in particular. Development aid was allegedly meant to take recipient countries to the stage of ‘self-sustained growth’ through rapid economic growth. This type of thought had no eye for poverty reduction.⁵⁸ The same applied for many recipient countries. Their national development plans generally focused on capital accumulation destined for industrialization.⁵⁹

⁵⁵ E.g. S. Macekura, ‘Point Four and the Crisis of U.S. Foreign Aid Policy in the 1970s’ in: R.H. Geselbracht (ed.), *Foreign aid and the legacy of Harry S. Truman* (Kirksville 2014) 73-100; J.J.P. de Jong, ‘In het kielzog van Multatuli. Van koloniaal welvaartsproject naar ontwikkelingssamenwerking’ in: B. de Graaff, D. Hellema, B. van der Zwan (eds.), *De Nederlandse buitenlandse politiek in de twintigste eeuw* (Amsterdam 2003) 55; M.E. Latham, *Modernization as Ideology: American Social Science and “Nation Building” in the Kennedy Era* (Chapel Hill 2000); Ekbladh, *The Great American Mission*.

⁵⁶ Brilliant analyses of modernization theory can be found in Ekbladh, *The Great American Mission*; Gilman, *Modernization Theory in Cold War America*; Latham, *The Right Kind of Revolution*.

⁵⁷ Latham, *The Right Kind of Revolution*, 7.

⁵⁸ Schmelzer, ‘A Club of the Rich to Help the Poor?’ in: Frey, Kunkel, Unger (eds.), *International Organizations and Development*, 182-183.

⁵⁹ M. Finnemore, ‘Redefining Development at the World Bank’ in: R. Packard, F. Cooper (eds.), *International Development and the Social Sciences: Essays on the History and Politics of Knowledge* (Berkeley 1997) 205-209.

But somewhere in the late 1960s and early 1970s, criticism gained ground. American scholars emphasized the destructive effects of the Vietnam War on public confidence in American foreign interference, including development aid.⁶⁰ More generally, Unger distinguishes between three key criticisms that put modernization theory and its policies in contempt: the lack of success of development policies in previous decades; criticism on the role of the state as a catalyst; and the growing confidence of decolonized countries and, subsequently, their political and socio-economic struggles.⁶¹

Firstly, public awareness on the lack of success grew due to two reasons at the end of the 1960s. Multilateral institutions and academic circles had gathered more data, which displayed the enormous extent of poverty in recipient countries.⁶² Additionally, a number of publications influenced public opinion. Behind the scenes, studies of Lester Pearson, Robert Jackson and the CDP reset the stage at the World Bank and the UN. Books of Gunnar Myrdal and Albert Hirschman also had a profound impact in the wider international community.⁶³ Although many recipient countries reached staggering rates of GNP growth during the 1960s, the data and the commentators illustrated that the economic boost did not trickle down to the poorer parts of developing societies.

Secondly, criticism on modernization theory also originated in a declining faith in centralized planning. Commentators viewed state-driven development as 'paternalistic and overly bureaucratic'.⁶⁴ This criticism was a precursor to the more general popularity of anti-statist thought during the neoliberal turn in the late 1970s and 1980s. NGOs and the private sector gained support because they allegedly delivered projects with a more tailor-made approach. Grand modernizing schemes were detested and replaced with a preference for stereotypes of more 'natural', 'organic' stimulation of development. This criticism developed further into the late 1970s, to focus on 'bad governance' by governments in the Global South, which included 'the lack of efficiency, responsibility, and reliability.'⁶⁵

On the other hand, critics from the Global South still favoured big plans in infrastructure and manufacturing but argued that modernization theory failed because of structural power imbalances in the international economic order. Members of the G77 voiced this opinion during UNCTAD conferences based on the theories of Argentinian economist Raul Prébisch and German-

⁶⁰ Ekbladh, *The Great American Mission*, 23.

⁶¹ Unger, *International Development*.

⁶² Finnemore, 'Redefining Development at the World Bank', 209.

⁶³ E.g. A.G. Frank, *Capitalism and Underdevelopment in Latin America. Historical Studies of Chile and Brazil* (New York 1967); G. Myrdal, *Asian Drama. An Inquiry into the Poverty of Nations* (New York 1968); A.O. Hirschman, *Development Projects Observed* (Washington 1967). For a reception of Myrdal's work, see T. Lankester, "'Asian Drama': the pursuit of modernization in India and Indonesia' *Asian Affairs* 3 (2004) 291-304.

⁶⁴ Unger, *International Development*, 133-137, 146.

⁶⁵ *Ibidem*, 144.

British economist Hans Singer.⁶⁶ Despite their voting majority in the UNCTAD and the UN General Assembly, the G77 was not able to find a breakthrough to substantially alter the global trading system. Nonetheless, its criticism on the existing system of trade and development undoubtedly influenced a shift in development policies of OECD member states.⁶⁷

Not specifically mentioned by Unger was the environmentalist criticism on modernization theory. The popularity of introducing modern agricultural technology waned in the 1970s when several reports about the negative external effects came to light. Such concerns were raised many times, most famously in the report *Limits to Growth* by the Club of Rome in 1972. Less known is the collection of ecological disasters described by a group of researchers in 1969, titled *The Careless Technology: Ecology and International Development*.⁶⁸ Through the Conference on the Human Environment in Stockholm in 1972, the UN and several member states aimed to find common ground to increase the sustainability of international and national development programmes.

In addition, the international community of the 1970s paid more attention to direct poverty alleviation and basic needs.⁶⁹ International Relations scholar Martha Finnemore offers a concise but clear description of the shift in focus brought about by the first crisis in development:

When development first emerged as a transnational mission in the 1950s, the goal of development was understood to be raising GNP. Now that goal has been expanded to include concerns about income distribution and poverty, environmental degradation and sustainability, the status of women, maintaining cultural integrity of indigenous populations, and human rights.⁷⁰

In Finnemore's opinion, it was the World Bank who drove the agenda of poverty alleviation. More specifically, World Bank President Robert McNamara made poverty alleviation a top priority of the Bank's lending policy. This included focusing on individuals instead of countries, and on education, small-scale farming and (mal)nutrition. Then again, scholars disagree as to what exactly characterized the dominant development thought in the decades after the demise of modernization theory.⁷¹ An important take-away from this thesis, nonetheless, is that confidence

⁶⁶ Often dubbed as the dependency theory, centre-periphery theory or the Prébisch-Singer thesis.

⁶⁷ Macekura, 'Development and economic growth' in: Borowy, Schmelzer (eds.), *History of the Future of Economic Growth*, 121-122.

⁶⁸ M. Taghi Farvar, J.P. Milton (eds.), *The Careless Technology: Ecology and International Development* (New York 1969).

⁶⁹ Finnemore, 'Redefining Development at the World Bank' in: Packard, Cooper (eds.), *International Development and the Social Sciences*, 210-217.

⁷⁰ Ibidem, 203.

⁷¹ Ekbladh, *The Great American Mission*, 230-231.

in existing development thought had decreased by the end of the First Development Decade. This may have been the ground for the emergence of evaluation as an instrument for policy analysis.

1.1.3. The history of public policy evaluation

Starting in the late 1970s, many OECD countries initiated widespread bureaucratic reforms captured by the overarching term New Public Management (NPM). Key to this was a new style of public administration entailing both 'explicit standards and measures of performance' and 'greater emphasis on output controls'.⁷² Effects of public policy should be measured, and continuation of policy depended on its results. NPM focused on output and outcome, rather than the process of policy preparation and execution. Supporters of NPM also assumed that private-sector management techniques in the public sector would enhance the latter's efficiency.⁷³ This was deemed necessary because of large budget deficits, a greater desire for accountability and public disaffection of government policies.⁷⁴ In theory, evaluation fits the framework of the NPM, being an instrument to measure the effectiveness and efficiency of public policy.⁷⁵ It is not a coincidence, then, that members of the OECD's Development Assistance Committee (DAC) revitalized their joint efforts in aid evaluation in 1979.⁷⁶ But where can we find the roots of public policy evaluation? This paragraph first focuses on public administration in general, before analysing the literature on aid evaluation in particular.

Most scholars distinguish between several waves of policy evaluation. The US, Sweden, Germany and the UK initiated the first wave of public policy evaluation in the context of expanding welfare state activity in the 1960s. Such programmes sparked the establishment of a new policy cycle, including centralized planning and evaluation.⁷⁷ The first wave, the 'science-driven wave', was related to the narrative

that the world can be made more humane if capitalism and the market economy can be reined in by appropriate doses of central policy planning and public intervention (...). In

⁷² C. Hood, 'A public management for all seasons?' *Public Administration* 1 (1991) 3-19.

⁷³ S.P. Osborne, 'The New Public Governance?' *Public Management Review* 3 (2006) 377-387, 378-379.

⁷⁴ M.J. Hatton, K. Schroeder, 'Results-based management: friend or foe?' *Development in Practice* 3 (2007) 426-432, 427; L. Berlage, O. Stokke, 'Evaluating Development Assistance: State of the Art and Main Challenges Ahead' in: L. Berlage, O. Stokke (eds.) *Evaluating Development Assistance. Approaches and Methods* (London 1992) 1.

⁷⁵ P. Dahler-Larsen, *The Evaluation Society* (Stanford 2012) 3; J.E. Lane, *The Public Sector* (London 1993) 132.

⁷⁶ OECD, *A History of the DAC Expert Group on Aid Evaluation* (Paris 1993) 7.

⁷⁷ H. Wollmann, 'Policy Evaluation and Evaluation Research' in: F. Fischer, G.J. Miller, M.S. Sidney (eds.), *Handbook of Public Policy Analysis: Theory, Politics, and Methods* (Boca Raton 2007).

public-sector thinking, this was hailed as a victory of a kind of rationality. Public policy should be made more scientific and sensible.⁷⁸

Evaluation could make government policy more rational and scientific, and thereby more effective and efficient. Evaluation dealt with the question 'what works'.⁷⁹ In the US, evaluation originated in the establishment of the Program, Planning and Budgeting System (PPBS) in the Department of Defense in 1961. Robert McNamara, the later World Bank President, personally encouraged the use of PPBS data when he assumed the position of Secretary of Defense. McNamara found in PPBS a tool through which he received valuable information on 'problem identification, development of options or alternatives, delineation of objectives and criteria, evaluation of impacts of these options, estimate of future effects, and – of course, recommendations for action.'⁸⁰ The staff conducting PPBS were in touch with McNamara directly, and many of them eventually became known as his 'Whiz Kids'. Some of them spread out to other federal departments after President Johnson instructed his cabinet members by executive order to use PPBS in their departments in 1965.⁸¹ In Johnson's opinion, PPBS could become a 'very revolutionary system (...) so that through the tools of modern management the full promise of a finer life can be brought to each American at the lowest possible cost.'⁸²

Purposes and practices of evaluation differed per country. Scholars discussed to which extent American experiments took root in other countries. Canada, Sweden and the UK all had favourable domestic environments for the institutionalization of evaluation: their governments were progressive; they all national social reform programmes; they had many cabinet members with a background in social science; and they had increasing government budgets.⁸³ In the UK, several departments experimented with PPBS. They were driven by the Fulton Committee research on the state of the British civil service, requested by the House of Commons in 1965. The research revealed 'the amateurishness of senior civil servants and recommended, inter alia, that departments should set up policy planning units.'⁸⁴ This is similar to the US, where evaluation in programs of social intervention was promoted to eliminate 'the chaos and inefficiency in existing

⁷⁸ Vedung, 'Four Waves of Evaluation Diffusion', 265.

⁷⁹ Also acknowledged by V. Pattyn, S. van Voorst, E. Mastenbroek, C.A. Dunlop, 'Policy Evaluation in Europe' in: E. Ongaro, S. van Thiel (eds.), *The Palgrave Handbook of Public Administration and Management in Europe* (London 2018) 578-580.

⁸⁰ B.A. Radin, *Beyond Machiavelli. Policy Analysis Comes of Age* (Washington D.C. 2000) 9-30.

⁸¹ P.H. Rossi, H.E. Freeman, M.W. Lipsey, *Evaluation. A systematic approach* (Thousand Oaks 1999) 17; R.H. Haveman, 'Policy Analysis and Evaluation Research After Twenty Years' *Policy Studies Journal* 2 (1987) 191-218.

⁸² Radin, *Beyond Machiavelli*, 17.

⁸³ C. Laporte (Ph.D. thesis), *L'évaluation, un objet politique: le cas d'étude de l'aide au développement*, Institut d'études politiques de Paris, March 2015, 41

⁸⁴ C. Pollitt, 'Occasional excursions: A brief history of policy evaluation in the UK' *Parliamentary Affairs* 3 (1993) 353-363, 361.

welfare services'.⁸⁵ In the FRG, it was parliamentary influence that played at least some role in the institutionalization of evaluation. The Bundestag 'started to require the federal government to report on the implementation and impact of various socioeconomic and tax programs.'⁸⁶

The second wave of policy evaluation erupted shortly after the first oil price rise in 1973. The cause here was not expanding government activity but rather its contraction: 'the mandate of policy evaluation got accordingly redefined with the aim to reducing the costs of policies and programs'.⁸⁷ This expansion of evaluation for cost-efficiency of public policy happened in numerous countries, for example in the Netherlands and the US:

Öffentlichkeit und Parlament versprachen sich eine verbesserte Kontrolle staatlicher Maßnahmen, die Berücksichtigung von Neben- und Folgewirkungen; der Verwaltung eröffnete Evaluierung Eingriffschancen, um verbesserte Zielgenauigkeit der Programme, verbesserte Wirtschaftlichkeit der Maßnahmen und Senkung der Kosten zu erreichen.⁸⁸

Evaluation expanded almost across every western country. In the UK, Conservative Prime Minister Edward Heath instituted a Central Policy Review Staff in 1970, responsible for the evaluation of large, mostly interdepartmental government programs. Such policy reviews then disappeared when budget cuts took away the capacity for evaluation in the second half of the 1970s.⁸⁹ Either way, the practice of evaluation served as a response to external influences on government departments. External interference increased with the emergence of NPM from the late 1970s onwards. The transfer of implementing responsibilities from government to semi-public or private agents partly transformed government departments from programme operators into mere supervisors. Evaluation was a tool to perform this new role and to hold implementing partners accountable. Similarly, within the NPM thought, government activity was regarded as a service to customers. Evaluation, then, was a test to find out whether customers are served well:

⁸⁵ A. Oakley, 'Public Policy Experimentation: Lessons from America' *Policy Studies* 2 (1998) 93-114, 93.

⁸⁶ 'Reports included those on the Labor Market and Employment Act (1969), the General Educational Grants Law (1969), the joint federal-state program to improve the regional economic structure (1970), the hospital investment program (1971), and various reports on subsidies and taxes.' In: L.G. Morra Imas, R. Rist, *The road to results: Designing and conducting effective development evaluations* (Washington 2009) 21; H.U. Derlien, 'Program Evaluation in the Federal Republic of Germany' in: R.C. Rist (ed.), *Program Evaluation and the Management of Government: Patterns and Prospects across Eight Nations* (New Brunswick 1999) 37-52.

⁸⁷ Wollmann, 'Policy Evaluation and Evaluation Research' in: Fischer, Miller, Sidney (eds.), *Handbook of Public Policy Analysis*, 396-397.

⁸⁸ Stockmann, 'Kleine Entwicklungsgeschichte der Evaluationsforschung Nachholende Entwicklung in Deutschland' in: S. Brüne (ed.), *Erfolgskontrolle in der entwicklungspolitischen Zusammenarbeit* (Hamburg 1998) 31.

⁸⁹ Pollitt, 'Occasional excursions', 361-362.

In the neo-liberal wave, it is regarded as imperative that the fundamental principal in a representative democracy, the demos, has a right to know how her agents spend her money. (...) Evaluation has taken on new expressions in the form of accountability assessments, performance measurement and consumer satisfaction appraisal. Quality assurance and benchmarking are also recommended.⁹⁰

Many scholars cited above studied the onset of evaluation within public administration in general. In contrast, only little is known about the history of aid evaluation. The most recent and extensive study was conducted by Camille Laporte, who received her Ph.D. for a study on the history, state of art and effects of aid evaluation. Focusing mainly on France and the UK she concludes that in the 1990s, after decades of little activity, a 'culture of evaluation' came into being as a consequence of domestic political pressures and a normalization of evaluation in multilateral development programmes. Unfortunately, Laporte is very concise about the 1960s and 1970s, a period that she describes as a time in which the effectiveness of aid was barely doubted. Aid evaluation attracted only little interest from political decision-makers in a time when aid budgets seemed unlimited. Aid agencies simply felt no need for transparency towards the general public.⁹¹

Laporte also argues that when evaluation took place, it was enshrined in an 'evolutionist ideology' proscribing the application of western economic models in the Global South.⁹² Similar to other scholars, Laporte further argues that the US played a catalysing role in experimenting with public policy evaluation. She refers to the large influence of Edward Schuman's *Evaluative research*, published in 1967, and the establishment American Evaluation Association in 1976. Laporte states more explicitly than other scholars that the primary function of public policy evaluation was to find proof of a programme's success, in order to legitimize political choices and to show transparency. Evaluation had a clear parliamentary aspect: it offered science-based guidance to ministers and politicians. Evaluation gave politicians '*savoir-pouvoir*':

[L]es rares évaluations qui étaient mises en oeuvre avaient pour objectif de confirmer la bonne réalisation des activités d'aide par rapport aux objectifs fixés en amont des projets et surtout aux montants engagés. Très peu de considération était donnée à l'évaluation des processus de mise en oeuvre des projets d'aide et de leurs résultats.⁹³

⁹⁰ Vedung, 'Four Waves of Evaluation Diffusion', 270-273.

⁹¹ Laporte, *L'évaluation, un objet politique*, 39.

⁹² *Ibidem*, 41-42.

⁹³ *Ibidem*, 58.

Thus, evaluation served mainly as a tool to show that policies reached their goals. Evaluation was not used to analyse the implementation and management of policies, nor was there an objective eye for results. Unfortunately, Laporte does not solidify her argument with much historical evidence.

Other scholars mentioned evaluation practices in individual countries. In the case of the US, development scholar Basil Cracknell argued that ‘very little of this “domestic” experience [in public policy evaluation] seems to have spilled over into the evaluation of aid programmes or projects.’⁹⁴ European diplomat and development scholar Gerard van Bilzen, however, shows that two US development agencies had already created evaluation units in 1952 and 1954.⁹⁵ The activities of these units were not systematically institutionalized by USAID until the late 1960s, and produced unsatisfactory results. Therefore, USAID conducted a study on why evaluation had not led to clear policy advice. The study found out that a major issue lied in the low degree of clarity in the statement of policy goals. Evaluations could, in sum, not assess the success of policies as policymakers had not set standards or benchmarks.⁹⁶

But according to Annette Binnendijk, evaluator at USAID, her agency’s evaluations in the 1960s were primarily conducted to measure the economic rates of return. Since GNP growth was aid’s most prominent objective, only the effects of aid on a region’s growth rate had to be assessed. When focus shifted towards basic human needs and poverty alleviation, USAID devised a new evaluation methodology called the Logical Framework Approach (LFA).⁹⁷ She further argues that, although a number of agencies took over the LFA, ‘little formal coordination existed in the 1970s among the various donors in terms of sharing evaluation findings or experiences with M&E methodologies and procedures.’⁹⁸ Moreover, host governments showed little interest in and understanding of evaluation. Nor did donor governments invest in evaluation capacities of host governments: ‘[o]ther than routine internal monitoring and evaluation, most evaluation efforts were by external teams; collaborative evaluations were far too rare.’⁹⁹

These claims resemble the findings of Laporte. A former official of the UK’s aid department Overseas Development Administration (ODA), now Department for International Development (DFID), told Laporte that he used students to conduct evaluation during the 1970s.

⁹⁴ Cracknell, *Evaluating Development Aid*, 41.

⁹⁵ G. van Bilzen, *The Development of Aid* (Cambridge 2015) 148. The two agencies were the Foreign Operations Administration and the Indian Planning Commission.

⁹⁶ M. Hageboeck, ‘United States Agency for International Development (USAID)’ in: B.E. Cracknell (ed.), *The Evaluation of Aid Projects and Programmes. Proceedings of the Conference organised by the Overseas Development Administration in the Institute of Development Studies, at the University of Sussex 7-8 April 1983* (London 1984) 102.

⁹⁷ A.L. Binnendijk, ‘Donor agency experience with the monitoring and evaluation of development projects’ *Evaluation Review* 3 (1989) 206-222.

⁹⁸ *Ibidem*, 213.

⁹⁹ *Ibidem*, 213.

There was simply no money available to hire consultants, nor was there qualified personnel at the ODA to conduct evaluations. The UK and other donor countries allegedly showed little interest in evaluating aid, especially compared to their interest in evaluating other parts of public policy.¹⁰⁰

In the case of Sweden, however, '[t]he need for follow-up, evaluation and assessment of international development cooperation to increase its efficiency was recognized already in Government Bill 1962:100'.¹⁰¹ Ten years later, SIDA founded its evaluation unit. In the 1960s and early 1970s, Swedish evaluation focused primarily on gaining experience with aid and learning lessons to improve future policy making. This changed in the second half of the 1970s, when 'there was a shift in the approach to partner countries, with an increased focus on Swedish development objectives; more emphasis is given to accountability and the importance of informing taxpayers in Sweden about the results of Swedish development cooperation'.¹⁰²

In the FRG, aid evaluation was taken up seriously in 1970. Similar to other sectors of German public policy, evaluation of its foreign aid originated in parliamentary requests. The BMZ established its *Inspektionsreferat* in 1971. Alexandra Caspari and Gudrun Lachenmann argue that the main goal of evaluation was policy improvement. West-German aid developed a significant measure of integration with American aid programmes. This also led the BMZ to take over American experiments of aid evaluation. Moreover, there was obvious criticism from academics and taxpayers towards the effectiveness of aid policies. From early on, then, aid evaluation at the BMZ was catalysed by external factors.¹⁰³

Remarkably, not all Western countries joined the movement towards aid evaluation. The French government assumed that its civil servants would not learn anything new from reviewing aid: '[I]n no case monitoring and evaluation [is] carried out in a systematic way (...). The French themselves tend to take a sceptical view of the need for extensive monitoring and evaluation, arguing that it will reveal little that the experienced staff does not already know.'¹⁰⁴

At the UN, evaluation gained ground already in the 1950s.¹⁰⁵ UNESCO commissioned a research on aid project evaluation, which resulted in the publication of a report in 1959.¹⁰⁶ Stokke mentions that the UN Economic and Social Council (ECOSOC) requested member states and

¹⁰⁰ Laporte, 'L'évaluation, un objet politique', 53.

¹⁰¹ Brolin, 'Conceptualizations of "results" in Swedish policy for development cooperation from the 1960s to the 2000s', 15.

¹⁰² Ibidem, 16.

¹⁰³ A. Caspari, *Evaluation der Nachhaltigkeit von Entwicklungszusammenarbeit. Zur Notwendigkeit angemessener Konzepte und Methoden* (Wiesbaden 2004) 11-43; G. Lachenmann, 'Evaluierungsforschung: Historische Hintergründe, sozialpolitische Zusammenhänge und wissenschaftliche Einordnung' in: D. Kantowsky (eds.), *Evaluierungsforschung und -praxis in der Entwicklungshilfe* (Zürich 1977) 68.

¹⁰⁴ S.H. Arnold, *Implementing Development Assistance. European Approaches to Basic Needs* (Boulder 1982) 28; Van Bilzen, *The Development of Aid*, 365.

¹⁰⁵ Van Bilzen, *The Development of Aid*, 156.

¹⁰⁶ S.P. Hayes, Jr., *Evaluating development projects* (Louvain 1966 [1959]).

international organizations to increase cooperation throughout the 1960s. The strategy for the Second Development Decade followed ECOSOC's request and promoted the institutionalization or strengthening of national evaluation systems in every developing country. Member states agreed that a review and appraisal should take place every two years. Moreover, UNDP became responsible for regular review of development progress at the level of its headquarters and regional bureaux:

Evaluation of performance ranked high on the UNDP agenda during the first years of its existence. The focus was not to be restricted to projects but should include the overall impact and effectiveness of development assistance. This was an integral part of the administrative reform proposed by the capacity study and the 1970 consensus.¹⁰⁷

Stokke further refrains from any in-depth historical analysis. He does not explain what such developments actually meant for the ideas and practices of aid evaluation. For example, he leaves out whether member states or UN officials pushed for the evaluation initiatives of EPTA and UNDP. This would reveal crucial information about the objectives behind the institutionalization of evaluation. Moreover, Stokke argues that UNDP aimed to implement a tripartite model – international organizations, donor, and beneficiary governments – but he does not delve into the specifics of the institutionalization.¹⁰⁸

Another multilateral organization committed to policy evaluation was the World Bank. This might not come as a surprise, since it was McNamara who introduced a management reform when he became World Bank President in 1968. He brought along his experience with introducing PPBS at the American Department of Defense. The 'first crisis of development' in the late 1960s prompted McNamara to reform the institution 'along the lines of a multinational company.'¹⁰⁹ His creation of the Operations Evaluation Department (OED) was part of this. Christopher Willoughby, former director of the OED, argues that '[f]rom the earliest stages in developing a system for more strategic management of the organization, it was recognized that once the foundations had been laid, feedback loops would need to be introduced, so that plans could be revised on the basis of actual outcomes.'¹¹⁰ Another reason why McNamara institutionalized evaluation lay in McNamara's desire to keep the World Bank as independent as

¹⁰⁷ O. Stokke, *The UN and Development: From Aid to Cooperation* (Bloomington 2009) 218.

¹⁰⁸ *Ibidem*, 69-70, 169, 191, 207.

¹⁰⁹ Unger, *International Development*, 137.

¹¹⁰ C. Willoughby, 'First Experiments in Operations Evaluation: Roots, Hopes, and Gaps' in: P. Grasso, S.S. Wasty, R. Weaving (eds.), *World Bank Operations Evaluation Department: The First 30 Years* (Washington 2003) 3-4.

possible from national governments. After all, he allegedly started prioritizing evaluation when the US announced that they would initiate their own evaluations of World Bank operations.¹¹¹

A third platform for aid evaluation was the DAC. The DAC put great effort in advancing evaluation techniques. In the first year of its existence, member states agreed to the Resolution of the Common Aid Effort. With that document, DAC member states set a goal 'to secure an expansion of the aggregate volume of resources made available to the less-developed countries and to improve their effectiveness.'¹¹² DAC member states often discussed the topic of evaluation at their joint meetings, and eventually convened several times solely for the purpose of discussing evaluation.¹¹³ This led to the establishment of the DAC Expert Group on Aid Evaluation in 1982 and the DAC Criteria for International Development Evaluations in 1991. Chapter II will demonstrate that these initiatives have a history of intensive DAC engagement, primarily through its nature as forum for donor coordination and sharing of best practices.

All in all, the existing literature shows us that some countries at least experimented with aid evaluation during the 1960s and 1970s. These experiments originated due to various motives and because of internal and external factors. The existing literature also illustrates that, to a certain extent, there was a caesura in development thought and practice in the late 1960s. The first crisis in development altered policies and caused a redirection of focus away from modernization theory-oriented projects.

1.2 DUTCH DEVELOPMENT AID

Zooming in on the Netherlands, historian Duco Hellema described Dutch development aid as the only part of foreign policy that had not yet matured by the end of the 1950s.¹¹⁴ That is not a surprise when keeping in mind that the Netherlands participated in EPTA from its inception, but remained relatively inactive in bilateral aid until 1963. From that year on, aid evaluation was continuously on the agenda of one State Secretary and four Ministers until 1977.¹¹⁵ To have a better understanding of the context in which Dutch aid evaluation started, we need to answer a number of questions about the Dutch aid apparatus in general. Who were the main actors and

¹¹¹ P.A. Sharma, *Robert McNamara's Other War: The World Bank and International Development* (Philadelphia 2017) 47-48; Grasso, Wasty, Weaving (eds.), *World Bank Operations Evaluation Department*, 166.

¹¹² OECD, *DAC in Dates. The History of the OECD's Development Assistance Committee* (Paris 2006) 10; OECD, *A History of the DAC Expert Group on Aid Evaluation*, 7.

¹¹³ T. Chianca, 'International aid evaluation: An analysis and policy proposals' (edition April 2008) <https://scholarworks.wmich.edu/dissertations/758/> (31 May 2018).

¹¹⁴ D. Hellema, *Neutraliteit & Vrijhandel. De geschiedenis van de Nederlandse buitenlandse betrekkingen* (Amsterdam 2002) 243.

¹¹⁵ State Secretary Isaac Diepenhorst (1963-1965) and Ministers Theo Bot (1965-1967), Berend-Jan (Bé) Udink (1967-1971), Kees Boertien (1971-1973) and Jan Pronk (1973-1977).

what were the major policy changes? This section also illuminates what is missing in the existing literature on the history of Dutch public policy evaluation.

1.2.1. The organization and actors behind Dutch development aid

Pressure from parliament and employers' organizations VNO and NCW caused an expansion of the aid budget and the introduction of bilateral aid in the early 1960s. The expansion required intense reorganization. Isaac Diepenhorst was named State Secretary of Development Cooperation in 1963 under supervision of the Minister of Foreign Affairs, Joseph Luns. Diepenhorst established and then coordinated the Directorate General for International Cooperation (*Directoraat-Generaal voor Internationale Samenwerking*, DGIS). DGIS itself consisted of a number of departments (*directies*) responsible for policy planning and execution: *Directie Internationale Technische Hulp* (DTH), *Directie Internationale Organisaties* (DIO) and *Directie Financieel-Economische Hulp* (DFO). In addition, the MFA added a *Bureau Beleidsvoorbereiding* (BE) to DGIS which would function as a centre of study and research. Furthermore, the *Interdepartementale Coördinatie Commissie inzake Hulpverlening aan Minder Ontwikkelde Landen* would harmonise policies of the various ministries involved in aid, primarily the Ministry of Economic Affairs and the Ministry of Agriculture and Fisheries.¹¹⁶

Despite the organizational reforms, Luns still criticised the alleged unwillingness of other ministries to follow the course of the MFA. Most Dutch aid in the 1960s was handled by the Ministry of Economic Affairs, since aid policy was intensely tied to the interests of the private sector.¹¹⁷ Development aid was 'booming business' and it seemed to Luns as if every section of the government pushed for influence.¹¹⁸ In order to eliminate the chaos, the State Secretary was replaced by a Minister for Development Cooperation¹¹⁹ in 1965. Another, more important reason to replace the State Secretary, was a political one: the extra ministerial seat gave the Catholic Party a majority in the Council of Ministers.¹²⁰

In addition, the MFA established a *Nationale Adviesraad inzake de Ontwikkelingshulp* (NAR). Tinbergen chaired the NAR, which consisted of around 60 individuals from the private sector and academia. BE operated as the NAR's secretariat. The NAR published a large number of recommendations to the Minister for Development Cooperation, two of which will be discussed in Chapter III. Although there is some discussion among historians about the impact of the NAR on aid policy, the recommendations reveal significant insights into the discussion about aid

¹¹⁶ Hellema, "'Beschamende situaties". De eerste staatssecretaris en de oprichting van DGIS' in: Nekkers, Malcontent (eds.), *De geschiedenis van vijftig jaar Nederlandse ontwikkelingssamenwerking*, 127-144.

¹¹⁷ P. Hoebink, 'The Humanitarianisation of the Foreign Aid Programme in the Netherlands' *The European Journal of Development Research* 1 (1999) 176-202.

¹¹⁸ Hellema, "'Beschamende situaties"', 140.

¹¹⁹ The position is a Minister without portfolio and still resides under the Minister of Foreign Affairs.

¹²⁰ Voorhoeve, *Peace, Profits and Principles*, 255.

evaluation. In addition to the NAR, most individuals with an expertise or interest in development aid could participate in other public or semi-public institutions in the field, such as the *Nationale Commissie Ontwikkelingsstrategie '70-'80* (NCO) or the *Financierings-Maatschappij voor Ontwikkelingslanden* (FMO). The NCO, founded by Minister Udink in the context of the Second Development Decade, was headed by Prince Claus. Its primary work focused on promoting the positive image of development aid among the general public.¹²¹ The FMO was a public-private partnership serving as a development bank. Lastly, the scope of Dutch NGOs active in the field grew substantially following the establishment of NOVIB in 1956. NGOs received public funding through the *Medefinancieringsprogramma* (MFP). This programme started in 1965 with 5 million guilders as budget for subsidies, but grew substantially to 70 million in 1975 and 240 million in 1981.

Behind such a large organization of public and semi-public institutions were various influential individuals. If we exclude the ministers and the state secretary, scholars agree that Jan Meijer and Ferdinand van Dam were most influential in shaping Dutch development aid in the 1960s and 1970s. Jan Meijer, 'the godfather of modern Dutch development aid', became DGIS and Van Dam headed BE.¹²² Prominent social democrat Meijer had been a major figure in the Dutch resistance during World War II, when he cofounded newspaper *Het Parool*. He bore the nickname Napoleon, deriving from his commanding behaviour within the Directorate General.¹²³ Van Dam became professor of international institutions and economics of developing countries at the University of Groningen in 1964. The two coordinated the first experiments to assess the effectiveness of Dutch aid.¹²⁴ They were eventually replaced by other officials but none of these public officials would leave such a mark on Dutch aid as Meijer and Van Dam.

Other academics performed a key role in the Dutch aid sector too. Tinbergen and De Vries are both mentioned before previously. De Vries presided over NUFFIC, the organization for international cooperation among universities, and served as the first rector of the International Institute for Social Studies in The Hague, the institute established in 1952 partly to educate students from recipient countries.¹²⁵ One of the most influential critics of Dutch development aid was also an academic – professor of non-western sociology Willem Frederik Wertheim, who taught at the University of Amsterdam.¹²⁶ While economist Tinbergen taught Minister Pronk,

¹²¹ Hellema, *Neutraliteit en vrijhandel*, 248.

¹²² DGIS is not only the abbreviation for the Directorate General, but also for its leading civil servant. P.A.M. Malcontent, "Het blijft voortmodderen. Maar ook dat kun je goed en slecht doen" Ferdinand van Dam (1931)' in: B. van der Zwan, B. de Graaff, D. Hellema (eds.), *In dienst van Buitenlandse Zaken. Achttien portretten van ambtenaren en diplomaten in de twintigste eeuw* (Amsterdam 2008) 190.

¹²³ M. Brandsma, P. Klein, *Jan Pronk. Rebel met een missie* (Utrecht 1996) 48.

¹²⁴ M. Dierikx, "Mister Ontwikkelingshulp" Jan Meijer (1914-1997)' in: Van der Zwan, De Graaff, Hellema (eds.), *In dienst van Buitenlandse Zaken*, 154.

¹²⁵ Kuitenbrouwer, *De ontdekking van de Derde Wereld*, 159-177.

¹²⁶ Arens, 'Multilateral Institution-Building and National Interest', 469-470.

sociologist Wertheim had taught Jan Breman, who also turned out to be a vocal critic of Tinbergen's ideas and Udink's policies. Both Pronk and Breman became important figures in the New Left movement and were particularly interested in turning Dutch aid policy upside down.¹²⁷

From the late 1960s onwards, criticism expanded to more popular media. Hollands Maandblad published an article by Arabist and former diplomat Jan Brugman titled 'De illusie van ontwikkelingshulp'.¹²⁸ De Telegraaf published two articles on a crisis in the Stichting Nederlandse Vrijwilligers (SNV), the equivalent of the US Peace Corps, in May 1974.¹²⁹ Despite the criticism, the policy direction and financial expansion of aid received broad support in parliament during the 1970s.¹³⁰ Parties mainly debated the involvement of private sector in decision-making and policy execution, for instance through the option of tying aid, and the selection of recipient countries. This enabled the Ministers for Development Cooperation to operate relatively independent from parliament. Most relevant here, is their suggestion that the lack of parliamentary attention had a negative impact on the incentive to assess the achievements of aid. There was little parliamentary stimulus to evaluate aid.¹³¹ Chapter III will check the validity of this claim on the basis of parliamentary proceedings and the source material of the MFA.

1.2.2. Major changes in Dutch aid policy

Several scholars discussed the motives behind Dutch aid. The discussion fits into a much larger academic debate about two questions on Dutch foreign policy in general. Firstly, has Dutch foreign policy predominantly been determined by international or domestic influences? And secondly, in case of the latter, has the merchant (*koopman*) or the minister (*dominee*) been more influential? Since a thorough discussion on the motives behind Dutch aid could encapsulate the entire length of the thesis, this paragraph remains on a more abstract level and distils only the general trends in the 1960s and 1970s from the literature.

The first relevant policy note (*Kamerbrief*) to parliament, published in 1962 by Minister Luns, can be seen as a radical break from the Dutch policy of the 1950s. By announcing the establishment of Dutch bilateral aid, the note exemplifies the economic or commercial motive behind Dutch aid. Luns pledged to initiate bilateral aid after pressure from employers' federations VNO and NCW.¹³² The new markets of the developing world opened up opportunities for export

¹²⁷ Brandsma, Klein, *Jan Pronk*, 33-37.

¹²⁸ J. Brugman, 'De illusie van ontwikkelingshulp', *Hollands Maandblad*, mei 1968.
https://www.dbnl.org/tekst/_hol006196901_01/_hol006196901_01_0059.php.

¹²⁹ H. de Mari, 'Hulpstichting smijt al jaren met miljoenen' *De Telegraaf*, 3 May 1974; H. de Mari, 'Vrijwilligers misbruikt voor eigen glorie', *De Telegraaf*, 4 May 1974.

¹³⁰ Brandsma, Klein, *Jan Pronk*, 51.

¹³¹ C. Cooper, J. Verloren van Themaat, 'Dutch Aid Determinants, 1973-1985: Continuity and Change' in: Stokke (ed.), *Western Middle Powers and Global Poverty*, 117-158.

¹³² Unofficially, the Netherlands was already involved in bilateral aid. Both the MFA and the Ministry of Economic Affairs offered subsidized loans to development projects initiated by the Dutch private sector.

of Dutch knowledge and products. VNO and NCW feared that the Dutch corporate sector would fall behind European counterparts, whose governments had initiated bilateral aid earlier in the decade. Sections within the Ministry of Economic Affairs also favoured a more positive attitude towards bilateral aid, and had pointed out that the stubborn monopoly of multilateralism harmed Dutch exports. Because bilateral aid was mainly commercially driven, the largest part of it was given in the form of tied aid: it obliged the recipient country to spend (part of) the aid on Dutch products.¹³³ Not every Dutch cabinet held business in high regard, however. Pronk had very limited economic motives, and he decreased levels of tied aid from 56% of ODA in 1973 to 30% in 1978.¹³⁴

Political motives were also prominent. This category should be split up in international and domestic motives. Regarding the former, the Netherlands had no alternatives if it wanted to: '[M]ee doen betekende mee tellen.'¹³⁵ Participating in EPTA and later initiatives provided the Netherlands with an instrument to polish their international reputation after the Indonesian War of Independence.¹³⁶ However, participating was not just about being at the table. Development aid offered the Dutch some measure of foreign policy independence. After all, more traditional sectors of foreign policy such as international security 'were pre-empted by the larger Powers.'¹³⁷ Development aid sometimes actually became intertwined with foreign policy, especially during Pronk's tenure. He used foreign aid to Egypt as a tool to smoothen relations with that country after the oil embargo of 1973.¹³⁸ Moreover, he regarded aid as a tool to improve human rights conditions in recipient countries. This stance irritated fellow Labour Party Minister of Foreign Affairs Max van der Stoep, who challenged Pronk's continuous involvement in his foreign policy.

Compared to other donor countries, the Dutch were generally not very interested in development aid for Cold War reasons. Anti-communism was much less influential in the Netherlands than in Germany 'where the general perception of a communist threat to security dominated the rhetoric of development co-operation.'¹³⁹ In contrast to the *dominee* and the

New Guinea, Surinam and the Caribbean colonies received large sums of money. Moreover, in 1955 the government decided to spend a sum of money on bilateral programmes of knowledge transfer. The size of the expenditures given to regions other than the colonies was very limited.

¹³³ Tied aid requires recipient countries to spend aid money in the donor country. Hoebink, 'The Humanitarianisation of the Foreign Aid Programme in the Netherlands', 176-202.

¹³⁴ O. Stokke, *International Development Assistance. Policy Drivers and Performance* (Basingstoke 2019) 270; F.A.J. Baneke, C.J. Jepma, 'Nederlands belang en ontwikkelingsbelang. Bedrijfsleven en ontwikkelingssamenwerking vanaf de jaren zeventig' in: Nekkers, Malcontent (eds.), *De geschiedenis van vijftig jaar Nederlandse ontwikkelingssamenwerking*, 265-266.

¹³⁵ Dierikx, 'Inleiding' in: *Nederlandse ontwikkelingssamenwerking I*, xxiii.

¹³⁶ M.L.J. Dierikx, 'Ontwikkelingssamenwerking, spagaat tussen moraliteit en handel' *Historisch Nieuwsblad* 5 (2002).

¹³⁷ Voorhoeve, *Peace, profits and principles*, 216-217.

¹³⁸ Malcontent, 'De schaduwminister van Buitenlandse Zaken', 218-223.

¹³⁹ Arens, 'Multilateral institution-building and national interest', 462-463.

koopman, the *generaal* played a minor role in Dutch development aid.¹⁴⁰ The 1966 policy note, the leading document for Dutch aid policy until 1973, even lacked any reference to the containment of communism.¹⁴¹

Finally, aid also had domestic political motives as a tool in public relations. Parliament and public opinion were generally very much in favour of development aid.¹⁴² Such motives often overlapped with humanitarian motives, especially during the Pronk years. This combination is dubbed by scholars as humane internationalism.¹⁴³ Before Pronk's tenure, we already see a clear presence of humanitarian motives in Dutch foreign aid. For instance, the participation in EPTA fit the existing policies vis-à-vis the former Dutch Indies. The humanitarian motive originated to a certain extent from feelings of guilt about the colonial past.¹⁴⁴ The Dutch held a 'paternalistic-moral notion' of duty towards the less fortunate across the border.¹⁴⁵ Surveys among the Dutch public also showed support for development aid for humanitarian reasons. Giving aid was seen a moral duty to alleviate poverty and hunger.¹⁴⁶ Human solidarity urged nations, as the 1966 policy note says, to help each other in overcoming violence and poverty. The general public 'discovered the Third World' through the television, and believed that the horrendous socioeconomic situations in recipient countries were a consequence of colonial exploitation.

The growing popularity of this humane internationalism coincided with transitions in international development thought. As noted in Section 1.1, by the late 1960s it came to light that the big push prescribed by modernization theory did not work out as expected by many leading scientists and politicians. The First Development Decade had failed. Some leading scientists and politicians had actually been aware of failures earlier than historians generally assume. Van Dam already mentioned this during his inaugural lecture at the University of Groningen in 1964: '[i]n de afgelopen maanden is in de Westelijke wereld een sfeer van twijfel en onbehagen ontstaan rond de hulpverlening aan de minderontwikkelde landen.'¹⁴⁷ Van Dam referred to an interview in the

¹⁴⁰ P. Hoebink, 'Hoe de dominee de koopman versloeg. Nederlandse ontwikkelingssamenwerking gewogen' *Internationale Spectator* 11 (2006) 578.

¹⁴¹ Because of the cabinet's fall later that year during the Nacht van Schmelzer, the Lower House never really discussed the note. Nevertheless, both Kuitenbrouwer and Hoebink argue that its ideas became the guiding principles for aid policy at least until 1973. Hoebink, *Geven is nemen*, 53; Kuitenbrouwer, *De ontdekking van de Derde Wereld*, 142.

¹⁴² Voorhoeve, *Peace, Profits and Principles*, 216-217.

¹⁴³ O. Stokke, 'The Determinants of Aid Policies: General Introduction' in: Stokke (ed.), *Western Middle Powers and Global Poverty*, 9-32; C. Cooper, J. Verloren van Themaat, 'Dutch Aid Determinants, 1973-85: Continuity and Change' in: Stokke (ed.), *Western Middle Powers and Global Poverty*, 117-158.

¹⁴⁴ M. Kuitenbrouwer, 'Een eeuwig durende ereschuld? Ontwikkelingssamenwerking en koloniaal verleden' *Internationale Spectator* 7/8 (2000) 377-382, 379-380; J.J.P. de Jong, 'In het kielzog van Multatuli. Van koloniaal welvaartsproject naar ontwikkelingssamenwerking' in: De Graaff, Hellema, Van der Zwan (eds.), *De Nederlandse buitenlandse politiek in de twintigste eeuw*, 46-47; Dierikx, 'Ontwikkelingssamenwerking, spagaat tussen moraliteit en handel'.

¹⁴⁵ Dierikx, 'Inleiding' in: *Nederlandse ontwikkelingssamenwerking* I, xxiii.

¹⁴⁶ Kuitenbrouwer, *De ontdekking van de Derde Wereld*, 129.

¹⁴⁷ Van Dam, *Onbehagen rond de ontwikkelingshulp*, 5.

Algemeen Dagblad with dr. K.P. van der Mandele in 1962, who rhetorically asked what had become of the 26 billion dollars spent by the West between 1948 and 1958 ‘*in het bodemloze vat van de ontwikkelingslanden*’.¹⁴⁸ Although Van Dam believed in the potential of aid, he criticized the expectations raised by development plans. Both policymakers and scholars did not sufficiently understand development yet.¹⁴⁹ Some years later, such doubts resurfaced in Dutch aid policy. The Den Uyl Cabinet even claimed in its 1974 budget presentation ‘*dat het ontwikkelingsvraagstuk in een impasse was geraakt*’.¹⁵⁰

Despite such doubts, Pronk aimed for a radical policy transformation when he took office in 1973. He argued that aid policy was foreign policy by definition, in the first place because development processes are embedded in international political and socio-economic relations. Secondly, Pronk regarded aid policy as a tool to influence such relations as well as a tool to influence national governance in recipient countries. Promotion of human rights was in that regard an important policy objective for Pronk.¹⁵¹ In addition to the instrumentalization of aid for foreign policy, Pronk left behind the policy objective of his predecessors focused on increasing the GNP of recipient countries. Instead, in Pronk’s opinion, aid should benefit first and foremost the least fortunate people in recipient countries. He further announced that self-reliance was a key objective.¹⁵² The Den Uyl Cabinet also agreed on a budget increase to 1.5% of net national income.¹⁵³ Although Pronk peaked with only 0.7% in 1975, the quantitative expansion of aid led to organizational problems in the 1970s. According to Hoebink, aid quadrupled during Pronk’s tenure parallel with only a minor increase in the number of personnel.¹⁵⁴

The extent to which Dutch aid policy changed during its early decades is still up for discussion.¹⁵⁵ Although one could clearly argue that Pronk set a different tone than his predecessors Boertien, Udink and Bot, historians note that many of Pronk’s ideas already existed at the MFA during the 1950s and 1960s but were never really showcased to the public. The increased presence of aid policy in the media, the polemic debating style of Pronk and growing criticism in parliament contributed to the sense that Pronk’s policy was a break from the past.

¹⁴⁸ Van Dam, *Onbehagen rond de ontwikkelingshulp*, 5-6; N.n. ‘Arme volken helpen maar geen geld in een bodemloos vat’ *Algemeen Dagblad*, 1 December 1962.

¹⁴⁹ Van Dam, *Onbehagen rond de ontwikkelingshulp*, 18-20.

¹⁵⁰ G. Ringnalda, ‘Verwachtingen en teleurstellingen. De Nederlandse rol in het VN-ontwikkelingsbeleid 1960-1980’ in: Nekkers, Malcontent (eds.), *De geschiedenis van vijftig jaar Nederlandse ontwikkelingssamenwerking*, 201-216, quote on 212.

¹⁵¹ P.A.M. Malcontent, ‘De schaduwminister van Buitenlandse Zaken. Ontwikkelingshulp als politiek instrument’ in: Nekkers, Malcontent (eds.), *De geschiedenis van vijftig jaar Nederlandse ontwikkelingssamenwerking*, 217-236.

¹⁵² Brandsma, Klein, *Jan Pronk*, 52.

¹⁵³ Dierikx, ‘Inleiding’ in: *Nederlandse ontwikkelingssamenwerking IV*, xvii.

¹⁵⁴ P. Hoebink, ‘Hoe de dominee de koopman versloeg. Nederlandse ontwikkelingssamenwerking gewogen’ *Internationale Spectator* 11 (2006) 578-584, 580.

¹⁵⁵ E.g. Malcontent, Nekkers, ‘Inleiding’, 58-60; Hoebink, ‘The Humanitarianisation of the Foreign Aid Programme in the Netherlands’, 198.

1.2.3. Dutch policy evaluation

Interestingly, former senior official of the Dutch MFA Joop de Jong mentions that Dutch development aid lacked evaluation, whereas colonial development projects were consistently evaluated. De Jong regards the absence of evaluation primarily as a political issue. In his opinion, it was impossible to criticize recipient countries for their inefficient use of foreign aid sums because of the Cold War context. Criticizing recipient countries, or demanding certain forms of governance as a condition for aid, would drive them to the communist side of the conflict.¹⁵⁶ This argument sounds valid, but is not based on historical research. It remains to be seen whether actors used such an argument against evaluation in these early years. Therefore, this paragraph focuses on the existing literature on evaluation in Dutch public policy, and aid policy in particular.

In the 1980s and 1990s, scholars touched briefly on the origins of Dutch public policy evaluation. The popularity of evaluation appears to have increased in the 1980s, partly due to the conviction that government budget cuts were necessary. In the transition from welfare state to ‘*versoberingsstaat*’, policies had to be assessed on efficiency and effectiveness. Inefficient and ineffective policies were disbanded.¹⁵⁷ The Vonhoff Committee (*commissie Hoofdstructuur Rijksdienst*) is often seen as the starting point for the expansion of evaluation activity. In the early 1980s, this committee emphasized the importance of policy evaluation.¹⁵⁸

Some scholars, however, claim that the Dutch government experimented with evaluation as early as the 1970s. The Biesheuvel Cabinets (1971-1973) established the *interdepartementale Commissie voor de ontwikkeling van beleidsanalyse* (COBA). The purpose of the COBA was to promote and expand the tools for policy analysis, based on the American practice of PPBS which, as mentioned in Section 1.1, included evaluation.¹⁵⁹ This strategy of policy analysis in the 1970s became known to contemporary commentators as the *Herwaardering*. The objective of *Herwaardering* was ‘de vergroting van het maatschappelijk rendement in ruime zin en het richtsnoer daarbij is de kosten-baten-ratio en/of het marginale nut in vergelijking met dat van andere beleidsprogramma’s.’¹⁶⁰ Nevertheless, the *Herwaardering* largely remained a set of ideas. The Den Uyl Cabinet (1973-1977) announced to continue the policy analysis initiatives of the

¹⁵⁶ J.J.P. de Jong, ‘In het kielzog van Multatuli. Van koloniaal welvaartsproject naar ontwikkelingsamenwerking’ in: De Graaff, Hellema, Van der Zwan (eds.), *De Nederlandse buitenlandse politiek in de twintigste eeuw*, 61.

¹⁵⁷ A. Hoogerwerf, ‘Succes en falen van overheidsbeleid in Nederland’ in: A. Hoogerwerf (ed.), *Succes en falen van overheidsbeleid* (Alphen aan den Rijn 1983) 17-40.

¹⁵⁸ A. Sorber, J.R. Bax, A.J.M. Hendriks, ‘Inleiding’ in: A. Sorber, J.R. Bax, A.J.M. Hendriks, *Beleidsvaluatie: wisselwerking tussen cultuur en structuur* (The Hague 1995) 9.

¹⁵⁹ M. Herweijer, ‘Van herwaardering tot heroverweging’ in: Hoogerwerf (ed.), *Succes en falen van overheidsbeleid*, 41-54; F.L. Leeuw, ‘Evaluation Policy in the Netherlands’ *New Directions for Evaluation* 3 (2009) 87-102, 88-90.

¹⁶⁰ Herweijer, ‘Van herwaardering tot heroverweging’, 44.

Biesheuvel Cabinets, but did not follow up on the *Herwaardering* with concrete and durable practices.¹⁶¹ The COBA lacked support from both cabinet members and senior government officials, particularly because enhanced policy analysis required clear *a priori* statement of policy goals: 'explicitly stating policy goals is something of a problem in a multiparty, coalition government system like the Dutch political system. Clarified goals make one vulnerable: politicians are not inclined to show much interest in their own failures.'¹⁶² The same applied to civil servants: 'their careers are also supposed to flourish more by innovation than by reflection on a possibly unsuccessful past.'¹⁶³ Thus, '[b]y the end of the 1970s, the use of policy evaluation ex post with central government was limited.'¹⁶⁴

Based on this analysis, it can be argued that scholars overlooked the evaluation practices within the Dutch MFA. We do know, after all, that some form of evaluation occurred from the beginning of Dutch aid. Dierikx studied the existence of a feedback loop in the Dutch aid department in a number of case studies on aid projects. One of his studies concerns the interplay between NEDECO, a corporate partner of the Dutch aid community, and the MFA. NEDECO was founded by the Dutch government in 1952 as a consultancy firm sending experts abroad to discover potential development projects. The experts were responsible for assessing the feasibility of the projects, as well as their potential fit into the range of Dutch fields of expertise.¹⁶⁵

Dierikx analysed NEDECO's first project, the Saurashtra water management and land reclamation scheme in Gujarat, India. The project turned out to be a failure. Agricultural knowledge of Dutch experts proved to be rather unfit for the particular circumstances of the region. Antonie Warmenhoven, Director of DTH, had already notified that land reclamation in that specific region would not be as easy as planned, but other officials in The Hague apparently did not listen. Moreover, when the Dutch Ambassador in India informed The Hague that the Indian government was incapable of handling all its received foreign aid, the Dutch ministries involved again did not pay attention.¹⁶⁶ Despite this seeming lack of interest, NEDECO and the MFA did evaluate the project. While NEDECO valued the project at least for its learning experience, the MFA was much more critical: such projects should not be replicated on a larger scale in the future.¹⁶⁷

During another project, a feasibility study in Tunisia between 1959 and 1962, a high official from the MFA allegedly concluded that the Netherlands should not embark on such

¹⁶¹ Herweijer, 'Van herwaardering tot heroverweging', 44-45.

¹⁶² M.L. Bemelmans-Vidéc, R. Elte, E. Koolhaas, 'Policy Evaluation in the Netherlands: Institutional Context and State of Affairs' in: Rist (ed.), *Program Evaluation and the Management of Government*, 110.

¹⁶³ Ibidem, 111.

¹⁶⁴ Ibidem, 111.

¹⁶⁵ Dierikx, 'Like a fish out of water', 7.

¹⁶⁶ Ibidem, 6-8.

¹⁶⁷ Dierikx, 'Policy versus Practice', 642-643.

projects because of the failures in the past. The conclusion was not heard by colleagues. Another example analysed by Dierikx was a plan to increase cotton production on the island of Lombok, Indonesia, starting in 1964. Experts warned that the irrigation available on the island fell short of the requirements. Nevertheless, the decision makers at the MFA neglected the experts' advice and pushed through. They perceived the project to be too important for restoring the relationship with Indonesia. An evaluation in 1965 'revealed that none of the local counterparts really knew what the effort was supposed to bring about.'¹⁶⁸

These case studies show that, at least in the very early stages of Dutch aid, MFA staff in The Hague made little use of the information provided by executive staff on the ground. However, as mentioned earlier, Dierikx did not study initiatives aimed to improve feedback. In contrast, the research of Jos van Beurden and Jan-Bart Gewald offers some insight into the establishment of the IOV in 1977. Firstly, the authors explicitly note that Pronk, not his civil servants or parliament, was the main initiator. Secondly, they link the IOV to negative media reports in the years preceding its establishment. Two incidents in particular seemed to have influenced Pronk. The NRC published an article written by Brugman in response to the first report of the Werkgroep Janssen in 1969. Brugman argued 'that to spend so much money "on something rather dubious is dangerous folly."¹⁶⁹ The other media attention that allegedly moved Pronk in the direction of aid evaluation was the crisis of the SNV. De Telegraaf reported on this crisis several times during the early 1970s. Thirdly, concerning the question on who should conduct evaluations, Pronk did not want his own operational staff members to conduct the evaluations. He felt they had been too positive about the effectiveness of aid, primarily because they had been too involved in the projects themselves. Therefore, evaluation had to be conducted by independent researchers that played no part in managing aid projects. In sum, Pronk wanted objective information about the effectiveness of projects and, favourably, he wanted that information before Parliament or the media got hold of it. In other words, he wanted 'to remain one step ahead of his critical opponents'.¹⁷⁰

The IOV was not the only form of Dutch aid evaluation institutionalized during Pronk's tenure. The Dutch evaluation was a dual system, which included both internal evaluation conducted by staff of operational departments and external evaluation conducted by independent evaluators of IOV. The latter answered to the Minister for Development Cooperation directly and

¹⁶⁸ Dierikx, 'Policy versus Practice', 644-646.

¹⁶⁹ J. Brugman, 'Ontwikkelingshulp is een lapmiddel', Nieuwe Rotterdamsche Courant, 11 January 1969. Cited in: Van Beurden, Gewald, *From Output to Outcome*, 22.

¹⁷⁰ Van Beurden, Gewald, *From output to outcome*, 13-28, quote on 23.

performed a ‘warning function’ to prevent potential political crises.¹⁷¹ Until now, the existing literature lacks historical research into the establishment of internal evaluation at the Dutch MFA.

CONCLUSION

This chapter has shown that the literature first of all lacks an overview of the early history of aid evaluation. The literature is not up-to-date and often written from a political science perspective. This leaves room for historians to provide in-depth analyses of the actual debates on the institutionalization of aid evaluation, including all the difficult hindrances that proponents went through before they reached their goal. Chapters II and III illustrate how this process played out in the back rooms of international conferences and in Dutch policymaking.

Nevertheless, the literature on the history of development aid offers a number of take-aways. First of all, development aid was a learning process, ‘a field of continuous refinement, in which existing knowledge could be used to formulate even better solutions and more efficient approaches.’¹⁷² Foreign aid did not follow a timeline with paradigm shifts occurring when modernization theory was replaced all at once by new ideas. It is ‘misleading to speak of a concerted attack on mainstream development ideas’ occurring in the 1970s.¹⁷³ Instead, various critiques on very different components of development theory and practice emerged. Moreover, the particular context of the 1940s until roughly 1970, created by the experience of the Great Depression, the Second World War, the Cold War and decolonization, made development in these years a rather peculiar, state-centred process. This period ‘was the moment in time when public excitement about the needs for and the possibilities of development was at its highest (...) and when the field experienced a rapid professionalization process.’¹⁷⁴

Within such a context it was the state that initiated public policy evaluation, according to the literature. Section 1.1 has demonstrated that evaluation formed an essential part of the new central policy cycle, which included planning, execution, evaluation and policy alteration. The American PPBS, internalized first at the Department of Defense by Robert McNamara, gained popularity in other western countries. The experience of public management reform spilled over directly to the World Bank, when McNamara became its president in 1968. It can be assumed that reforms at the World Bank influenced the UN to speed up its introduction of new policy planning and evaluation. These reforms would take place in the context of the first crisis in development, in which great hesitation resurfaced regarding the feasibility of existing development thought – colloquially named modernization theory.

¹⁷¹ His sample included evaluation systems in Britain, Denmark, the FRG, the Netherlands, Norway and Sweden. Stokke, ‘Policies, Performance, Trends and Challenges in Aid Evaluation’, 11-13.

¹⁷² Unger, *International Development*, 155.

¹⁷³ *Ibidem*, 127.

¹⁷⁴ *Ibidem*, 154.

The analysis in Section 1.2 shows three main conclusions. Firstly, on the organization behind Dutch aid, policy-making was intertwined with academia. Along with Van Dam, Tinbergen and De Vries, Pronk, a member of the academic Werkgroep Janssen and later Minister for Development Cooperation in 1973, also played a significant role in Dutch aid. Pronk's experience in academia and his personal relation to Tinbergen had a serious impact on his policies. Secondly, the transitions in development thought and practice across the globe at the end of the 1960s also rooted in the Netherlands. Van Dam, being a major figure at the MFA, already noticed the complexity of successfully delivering aid in 1964. It was in this influential context that aid evaluation emerged. Thirdly, the literature demonstrates that aid evaluation did exist in the Netherlands to a certain extent. The MFA studied the results of aid, but it remains uncertain whether this occurred on a regular basis and if not, why. Chapter III will answer these questions.

II. AID EVALUATION ACROSS THE BORDER, 1966-1975

So far[,] satisfactory methods for evaluation do not exist. There is a choice of at least 300 definitions of factors leading to development to serve for evaluation purposes!¹⁷⁵

– Professor Johan Galtung, International Peace Research Institute, 1966

In a speech to ECOSOC in 1966, UNESCO's Director General René Maheu voiced the UN's desire to explore the field of aid evaluation: '[H]alf-way through the Development Decade, the Secretary-General [of the UN] invited (...) [UN agencies and member states] to make a collective examination of conscience regarding the effectiveness of all our efforts in the war – the only war that man should know – against poverty and the iniquity of hunger, disease and ignorance'.¹⁷⁶ Behind the scenes, the UN had been involved in experiments of aid evaluation for more than a decade. UNESCO already hosted a 'meeting on criteria and techniques of evaluation of technical assistance for economic development' in 1954.¹⁷⁷ Most evaluation activity in the UN system remained on an exploratory level, but the process accelerated quickly in 1966. Several UN agencies convened that year and founded an Inter-Agency Study Group on Evaluation.¹⁷⁸ Many European countries founded evaluation departments during the 1970s and early 1980s. Their shapes varied, 'reflect[ing] evaluation concerns, priorities and norms within the various political systems and administrative cultures.'¹⁷⁹

What we do not know, however, is the history of administrative and academic debates on aid evaluation before its institutionalization. These debates had a serious impact on the variety of aid evaluation systems. Academics and representatives of national governments and multilateral organizations met several times between 1954 and 1973 to discuss aid evaluation. Moreover, academics wrote numerous works on aid evaluation. This chapter is split up in three

¹⁷⁵ German Foundation for Developing Countries, *Methods and Procedures of Evaluation in Development Aid. Conference Report, 18-22 November 1966 in Berlin* (Berlin 1967) 93. Underlining from original text.

¹⁷⁶ The Director General is the highest ranking officer in UNESCO. Ibidem, 157.

¹⁷⁷ The Technical Assistance Board was the governing body of EPTA. UNESCO Special Summit 12, *Meeting on Criteria and Techniques of Evaluation of Technical Assistance for Economic Development Sponsored by UNESCO in Collaboration with the United Nations Technical Assistance Board, Geneva, 28-30 July, 1954: final report*, SS/12, available from unesdoc.unesco.org. UNESCO organized two more seminars on aid evaluation in 1961 and 1962 in Sirs-el-Layyan, United Arab Republic.

¹⁷⁸ UNITAR, *UN Development Aid. Criteria and Methods of Evaluation* (New York 1971) 14; UN Joint Inspection Unit, *Evaluation System of the United Nations Development Programme*, March 1983, JIU/REP/83/5, accessible online through https://www.unjiu.org/sites/www.unjiu.org/files/jiu_document_files/products/en/reports-notes/JIU%20Products/JIU_REP_1983_5_English.pdf.

¹⁷⁹ Stokke, 'Policies, Performance, Trends and Challenges in Aid Evaluation', 8.

parts to distinguish how the conference participants and the authors of studies on aid evaluation discussed the three subthemes relevant to the question why aid evaluation developed the way it did. First of all, what purpose did they connect to aid evaluation – lesson-learning or accountability? Secondly, who should conduct evaluation within national public administration? Thirdly, did actors discuss options of international cooperation in aid evaluation and if so, how?

Not all of the conference reports were official records drafted by the conference host. The first meeting, held by UNESCO in 1954, led to the publication of official minutes, as did the Berlin Conference. The Berlin Conference was hosted by the *Deutsche Stiftung für Entwicklungsländer* and attended by representatives of national governments and multilateral organizations.¹⁸⁰ The foundation still serves as forum for discussion on training in development aid. Its conference report consists of a summary of the conference proceedings, reports of six working groups that discussed particular topics in detail, and statements of the participating representatives. The Paris Conference was held under the auspices of the OECD's Technical Co-operation Committee in November 1968. Its source material consists of published statements of several attendees, ranging from multilateral organizations to OECD member states.¹⁸¹ The Wassenaar and Amsterdam Conference were organised by the DAC and hosted by the Dutch MFA in respectively 1970 and 1973. Their proceedings were also published officially by the OECD.¹⁸² Moreover, staff members of DFO and DTH wrote memos for their superiors, complementing the official records published by the OECD.¹⁸³

Concerning the studies on aid evaluation, the NAR and a key memo from DTH emphasized the relevance of three international studies: the Jackson Study on the capacity of the UN development system, the Pearson Report on the World Bank activities, and the report of the UN Committee of Development Planning, colloquially named the Tinbergen Report after its chair.¹⁸⁴ The studies were published in 1969 or 1970, during the first crisis of development and transition from the First to the Second Development Decade. Furthermore, the OECD Secretariat conducted

¹⁸⁰ Founded in 1959 and renamed as *Deutsche Stiftung für Internationale Entwicklung* in 1973.

¹⁸¹ OECD, *The Evaluation of Technical Assistance* (Paris 1969). Present were representatives of the UN Inter-Agency Study Group on Evaluation, France, the US, the EEC, the UK, Sweden, the Netherlands, UNDP and the OECD.

¹⁸² OECD, *Aid Evaluation. The Experience of Members of the Development Assistance Committee and of International Organisations* (Paris 1975).

¹⁸³ Scheltema was far from content with the conference: 'Na lezing van dit verslag zou men geneigd kunnen zijn om met Faust te verzuchten: "hier steh ich nun, ich armer Tor, und bin so klug wie je zuvor".' NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, memo 26 april 1971, drs. Scheltema (DFO/MZ) to Chef DFO; NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, letter 12 May 1970, Permanent Representative of the Kingdom of the Netherlands to the OECD, Paris, J. Kaufmann, to E. van Lennep, Secretary-General to the OECD.

¹⁸⁴ Nationale Raad van Advies inzake Hulpverlening aan Minder Ontwikkelde Landen, *Advies Technische Hulp* (The Hague 1972) 7; *Nederlandse ontwikkelingssamenwerking IV*, 214-215.

two studies on aid evaluation, in 1969 and 1975. The study also contained memoranda with national statements from the governments of Sweden, the FRG and the US.¹⁸⁵ Lastly, two handbooks on aid evaluation were published. Firstly, consultant Samuel Hayes published a his handbook in 1966, under auspices of UNESCO.¹⁸⁶ In 1971, the UN Institute for Training and Research (UNITAR) followed up with a handbook on UN development aid evaluation.¹⁸⁷ Combining input from the UN, World Bank, OECD, DAC member states and academics, the conference reports and studies reflect the viewpoints of the most important international actors in aid evaluation between 1966 and 1975.

2.1. THE PURPOSE OF AID EVALUATION

The minutes of the UNESCO seminar note that member states requested the Technical Assistance Committee, the supervisory body of EPTA, to make 'a critical examination of the activities undertaken and the results achieved'.¹⁸⁸ The seminar attendants agreed on two purposes of evaluation:

In the first place, for the efficient administration of large integrated programmes, there is a need for the kind of knowledge of how the process is working which can only come from a critical examination of the results. The second purpose is to provide an assurance to the governments supporting the programme that their money is well spent.¹⁸⁹

According to this, aid evaluation served the dual objective of both lesson-learning and accountability. Lessons could be learnt to enhance the efficiency of a programme. Moreover, evaluation enabled member states to hold the Technical Assistance Committee accountable. Accountability in a multilateral setting, however, differs from accountability in a bilateral setting. After all, evaluators of multilateral activities somehow report back to member states. Evaluation of bilateral projects and programmes has a different reporting mechanism since donor governments have no member states, of course. They do have a parliament, though, to whom evaluators can disclose the conclusions of evaluation.

The conference reports did not always clearly define the role of parliaments in the question 'why evaluate'. Most reports mainly outline the need for evaluation from the perspective of lesson-learning. The Berlin Conference report of 1966 reads that 'the aim of this international

¹⁸⁵ OECD, *The Evaluation of Technical Assistance*, 9-72.

¹⁸⁶ Hayes, Jr., *Evaluating development projects*.

¹⁸⁷ UNITAR, *UN Development Aid*.

¹⁸⁸ UNESCO Special Summit 12, *Meeting on Criteria and Techniques of Evaluation of Technical Assistance for Economic Development Sponsored by UNESCO in Collaboration with the United Nations Technical Assistance Board, Geneva, 28-30 July, 1954: final report*, SS/12, available from unesdoc.unesco.org.

¹⁸⁹ *Ibidem*.

effort is to be able to provide the scarce means with the utmost efficiency and to map out socio economic programmes most beneficial to developing nations.’¹⁹⁰ The purpose of aid evaluation, then, lies in its ‘basic function to help ensure the effectiveness of the aid provided and to develop policies and procedures governing the provision and utilization of such aid.’¹⁹¹ Most strongly, the purpose of efficiency and effectiveness was voiced in the first of six working groups convened during the conferences. This working group listed four purposes for evaluation, all of which corresponded to lesson-learning and none to accountability. Accountability was only mentioned in the margins as an advantage of evaluation conducted by unbiased outsiders.¹⁹²

Three countries presented their view on aid evaluation as a tool for more than just the opportunity to learn lessons: the Netherlands, Germany and Sweden. They contributed to the Berlin Conference report by a written statement. Their delegates argued that aid evaluation could serve as a tool for accountability and public relations. The most obvious example was the Dutch statement, written by Jan Meijer, in which he noted that

evaluation is not only an interesting scientific but also a domestic political problem. As the volume of aid to developing countries is increasing, the responsibility of the administration to Parliament and to the public grows and the drafting of budget proposals becomes more difficult. (...) The increasing proportion of public funds allocated to development needs (...) requires better arguments for directing public opinion towards accepting the impelling necessity to provide development assistance.¹⁹³

The Berlin Conference report also noted that in the Dutch opinion, ‘public relations have to be improved regarding the function of an outside-evaluator.’¹⁹⁴ Without further explanation in the report, this probably implied that the Netherlands supported evaluation as an instrument in the public debate on development aid. Evaluation could serve as propaganda in favour of development aid. This could be seen in light of Dutch interest in acting on the UN’s call to raise public awareness of development aid through the establishment of the NCO, some years later. All in all, it is clear that Meijer valued aid evaluation for more reasons than just lesson-learning.

The Netherlands did not stand alone in 1966. Otto von Schott from the BMZ outlined the German motives: ‘[a]ims of evaluation are equal to the Dutch demand; improvement of the efficacy of aid given and improvement of reasoning for Parliament and the public.’¹⁹⁵ The Swedish

¹⁹⁰ German Foundation for Developing Countries, *Methods and Procedures*, i.

¹⁹¹ Original underlining from source. Ibidem, 1.

¹⁹² Ibidem, 122-124.

¹⁹³ Original underlining from source. Ibidem, 79-80.

¹⁹⁴ Ibidem, 126.

¹⁹⁵ Original underlining from source. Ibidem, 75.

delegate further explained that accountability was the secondary purpose of aid evaluation. Aid evaluation helped 'in such a way as to facilitate granting of additional funds for expanding development aid.'¹⁹⁶ He also noted that aid evaluation should 'prove to *government authorities* that effective use has been made of financial resources'.¹⁹⁷ It is uncertain whether parliament is also implied, or whether he referred to the accountability of SIDA towards the Swedish MFA. In conclusion, already in 1966 there were several countries who wanted aid evaluation for reasons of accountability and public relations.

After 1966, the aspect of lesson-learning was a steady element of commentators writing about aid evaluation. The Tinbergen Report explicitly prioritized lesson-learning above other purposes. It contained recommendations for regular appraisal of progress during the Second Development Decade, and these 'appraisals should be designed in the main to bring about necessary adjustments and modifications in policies (...) and provide an opportunity to introduce new aims and policies as needed.'¹⁹⁸ The same argument can be found in the study of the OECD Secretariat in 1969: '[t]he aim of evaluation is to check the effectiveness (impact or incidence) of technical assistance and feed back the findings so that future activities can be improved.'¹⁹⁹ Six years later, the OECD held on to this contention by arguing that '[t]he essential purpose of undertaking an evaluation of an aid activity is not to re-write history but to improve aid operations in the future.'²⁰⁰ Jackson also notes that

early attempts at evaluation (...) were an expression of a growing conviction that technical co-operation for economic and social development was one of the most difficult tasks that had been undertaken (...). In other words, technical co-operation was an activity in which people learnt by doing.²⁰¹

Thus, it is clear that commentators valued evaluation for a large part for its instrumentality in lesson-learning. Nonetheless, it is crucial to also see the context of the first crisis of development from a different angle. Accountability gained more prominence these years. The Pearson Commission clearly noted that its own evaluation was relevant particularly because of the

¹⁹⁶ German Foundation for Developing Countries, *Methods and Procedures of Evaluation in Development Aid. Conference Report, 18-22 November 1966 in Berlin* (Berlin 1967) 102.

¹⁹⁷ Italics added. Ibidem, 102.

¹⁹⁸ United Nations Economic and Social Council, *Committee for Development Planning, Report on the Sixth Session (5-15 January 1970)* (New York 1970), accessible via https://documents-dds-ny.un.org/doc/UNDOC/GEN/NL1/493/10/pdf/NL149310.pdf?OpenElement&_ga=2.266739590.967101429.1551973053-87578162.1551973053, page 34. The review and appraisal section of the Tinbergen Report was noted in the Netherlands by NRC Handelsblad, 'Ecosoc-conferentie: "veelbelovende nieuwe start?"', 15 August 1973.

¹⁹⁹ OECD, *The Evaluation of Technical Assistance*, 16.

²⁰⁰ OECD, *Aid Evaluation*, 1.

²⁰¹ Jackson, *A study of the capacity of the United Nations development system II*, 453.

context of increasing concern about the future of international cooperation for economic development (...). [I]nternational support for development is now flagging. In some of the rich countries its feasibility, even its very purpose, is in question. The climate surrounding foreign aid programs is heavy with disillusion and distrust. (...) Nevertheless, we have reached a point of crisis.²⁰²

First of all, this is a key publication explicitly stating that there was a crisis in development aid. The Pearson Commission did not argue that recipient countries had not experienced economic development during the 1950s and 1960s; in fact, they recognized that economic growth in many of these countries reached higher percentages than had ever been the case in the developed world. Criticism, then, was partly fuelled by ‘misconceptions and unrealistic expectations of “instant development”’ when we should have known that development was a long-term progress.’ For some commentators, aid had gone to waste or inherently led to ‘entanglement in political conflict and military hostilities’. In addition, ‘[i]t is not only among the developed countries that the climate has deteriorated. On the developing side too there are signs of frustration and impatience. (...) On all sides we sense a weariness and a search for new directions.’²⁰³ Such criticism should be countered, the Pearson Commission argued. The Commission therefore conducted the study to answer the question what aid had delivered in the twenty years before 1969, in order to better inform national governments, politicians and the general public, and to prevent a potential stop of financial and technical assistance: ‘for any country to transfer public funds abroad without being able to satisfy its citizens that these funds are being effectively used’.²⁰⁴ The evaluation conducted by the Pearson Commission is therefore not an exercise of lesson-learning, but an exercise of accountability and public relations.

The same line of reasoning can be found in the UNITAR study in 1971. UNITAR also connected evaluation to the first crisis in development. In their eyes, the crisis arose as a consequence of the ‘very disappointing’ accomplishments of the First Development Decade. Evaluation would increase the aid sector’s public accountability: ‘[i]n general, appropriations for international aid are becoming harder and harder to obtain from the national legislative bodies. This applies equally to funds intended for multilateral programmes and for bilateral programmes. A number of donor countries are now (...) insisting that the multilateral agencies produce convincing evidence of the wisdom of their expenditures of contributed resources.’²⁰⁵ UNITAR

²⁰² L.B. Pearson et al., *Partners in Development. Report of the Commission on International Development* (New York 1969) vii, 4.

²⁰³ *Ibidem*, 5.

²⁰⁴ *Ibidem*, 6.

²⁰⁵ UNITAR, *UN Development Aid*, 25.

thus explained that accountability was primarily desired by leading donors, not the UN itself. Donors initiated discussions on aid evaluation during many of the UN agencies' executive board meetings. They did so allegedly because 'there is an insistence among these contributors that they and their legislative bodies need assurance that the resources made available are expended only for the most useful projects'.²⁰⁶ Evaluation served public accountability at home.

Jackson made the same observation: 'The recent increase in governments' interest in evaluation has been primarily connected with the basic question as to whether the major contributors to the various voluntarily-financed development co-operation activities have obtained "value for their money"'.²⁰⁷ He further referred to the fact that improperly executed evaluation could even worsen the public image of the UN. He did see evaluation as a crucial tool for lesson-learning too, as long as it was carried out in organized manner by qualified personnel. Nonetheless, Jackson clearly focuses more on the damage done to the UN's accountability and reputation than to potential policy improvement.

The Tinbergen Report introduced a rather interesting type of accountability. It read that appraisals must also 'induce all participants in the Decade to carry out their obligations in respect of past agreements and commitments and provide.'²⁰⁸ It is unclear who was included as participants in the Decade. At the bare minimum, we could conclude that member states and UN organs took part, but whether other parties such as NGOs were included is not sure. Nonetheless, we can deduct from this quote that evaluation primarily served to enable member states and UN organs to hold other member states to account. This option was the clear opposite of multilateral evaluation as described by Jackson, that served to enable accountability of UN agencies vis-à-vis member states and donor countries in particular.

The first crisis in development also impacted the discussion on aid evaluation outside the framework of the UN and the World Bank. At the OECD Technical Co-operation Committee meeting in Paris, November 1968, the American Representative to the DAC Mr. S.H. van Dyke connected aid evaluation to the increase of public scrutiny. Van Dyke mentioned that in 1960, the US Congress debated the work and results of USAID and concluded by establishing the position of Inspector General for Foreign Assistance. But USAID itself responded too:

the American Aid Programme has been increasingly criticised by people who think that the results achieved have been marginal, or have been too costly, and that changes ought

²⁰⁶ UNITAR, *UN Development Aid*, 9.

²⁰⁷ Jackson, *A study of the capacity of the United Nations development system II*, 188.

²⁰⁸ United Nations Economic and Social Council, *Committee for Development Planning, Report on the Sixth Session (5-15 January 1970)* (New York 1970), accessible via https://documents-dds-ny.un.org/doc/UNDOC/GEN/NL1/493/10/pdf/NL149310.pdf?OpenElement&_ga=2.266739590.967101429.1551973053-87578162.1551973053, page 34.

to be made or the Programme should be reduced. In response to these pressures, [US]AID has conducted over a period of 10 to 15 years some rather searching self-examinations. (...) In response to Congressional pressure and from public criticism in general, [US]AID has now established a Director for Programme Evaluation.²⁰⁹

Seven years later, in 1975, the OECD explicitly stated that the crisis of development left a significant mark on aid evaluation: '[u]ndoubtedly, the growing importance attaching to evaluation work has been due very largely to a gathering sense of disappointment with the results of aid and to the realisation that the problems to be solved and the means to be used to promote development are becoming increasingly complex.'²¹⁰ But the OECD did not link the emergence of evaluation to *public* disappointment with the results of aid. The summary of the discussion does not take into account the role of parliament or public opinion.

The worlds of the OECD and the UN and the World Bank were not that far apart. The OECD Secretariat found it imperative to assess 'how development assistance can be made most effective, a concern shared by the governments of the recipient countries and donor countries alike and one in which the general public is beginning to take an interest.'²¹¹ More strongly, the OECD argued that '[t]here seems to be an obvious link between the evaluation of aid and the attitude of public opinion towards aid. The United Nations and the Development Assistance Committee (DAC) of the OECD are now addressing themselves to this latter problem.'²¹² It is unclear if this initiative led to a particular statement or report, but the mere existence of the UN-DAC cooperation at least confirms that public relations were on the radars of the organizations. All in all, multilateral agencies and national governments clearly supported aid evaluation to publicly promote public spending on foreign aid. Evaluation would display to parliaments and taxpayers that aid delivered.

Interestingly, not all national governments promoted aid evaluation. The French representative at the OECD meeting in Paris in 1968, J.C. Guisset,²¹³ viewed aid evaluation in a different light. Instead of supporting transparency on aid results, he concluded his statement 'by drawing attention to a danger (...): any improvement in the efficiency of projects achieved through a more effective method of evaluation must not lead to a reduction of aid.'²¹⁴ Transparency and accountability could have two consequences – they either increase or decrease

²⁰⁹ OECD, *The Evaluation of Technical Assistance*, 81.

²¹⁰ *Ibidem*, 2.

²¹¹ *Ibidem*, 9.

²¹² *Ibidem*, 9n.

²¹³ Chef de la Section Multilaterale, Direction de la Coopération Technique, Secrétaire des Affaires Etrangères.

²¹⁴ OECD, *The Evaluation of Technical Assistance*, 79.

public support. Nonetheless, Guisset was the only delegate in all sources analysed this chapter to use this argument.

It is true, however, that delegates of multilateral agencies and national governments not always emphasized the need for accountability. Mr. Kramer and Mrs. Von Metzch, staff members of DTH and authors of the Amsterdam Conference report in 1973, described six purposes of evaluation discussed during the conference, four of which entailed aspects of lesson-learning. Two of them, 'verantwoorde allocatie van fondsen' and 'beoordeling realisatie doelstellingen', imply some measure of accountability. The authors did not further specify the details of these two purposes. Throughout the rest of the report, however, they never referred to accountability.²¹⁵ Of course, we do not know whether the participants in Wassenaar were the same individuals as in Berlin. Maybe the discussions of accountability simply might not have reached the final report. But all in all, the Netherlands, Germany and Sweden seemingly did not raise the same arguments in Wassenaar as they raised in Berlin seven years earlier.

A third purpose which has not yet come to light in this thesis, is the expansion of scientific knowledge. This purpose resembles lesson-learning, but there is a crucial difference. Lesson-learning refers directly to the alteration of public policy on the basis of evaluation. The expansion of scientific knowledge does not necessarily cause policy changes. Moreover, while lesson-learning is often a practice where new knowledge stays inside the multilateral or national development agency, scientific knowledge is much more transnational due to the nature of academia. Lastly, the sources mentioning the expansion of scientific knowledge clearly distinguish it from the purpose of lesson-learning. This is the case, for example, with the Jackson Study, which argues that '[a]ssessment of results (...) provides additions to the corpus of world experience on successful and unsuccessful techniques'.²¹⁶

To conclude, it must be said that it is very hard to draw a pattern. We can, however, distil from the sources that lesson-learning has been a key priority for evaluation among international donors at least since 1954. The role of accountability, on the other hand, is more debatable. Concerning multilateral aid, accountability seems to have been an important objective because evaluation enabled member states to hold multilateral institutions accountable. Evaluation could also be a tool for multilateral agencies to prove to member states the success of their activities, on the one hand, and their professional management of activities on the other. After all, evaluation would ensure that lessons will be learned and efficiency and effectiveness of aid will be improved. Concerning bilateral aid, accountability might not have been the key purpose of evaluation for

²¹⁵ NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, memo 20 July 1973, Evaluatie-symposium DAC, Director DTH to Minister Pronk via plv. DGIS and SA.

²¹⁶ Jackson, *A study of the capacity of the United Nations development system II*, 194.

most national governments, but it is clear that accountability emerged in the minds of public servants earlier than the 1970s. This emergence was not necessarily a response to budget cuts, but rather because of rising doubts about aid's effectiveness and efficiency in the public debate. Further research of national sources should illuminate what priority the MFAs and national aid agencies accredited to accountability. Chapter III intends to answer that question concerning Dutch aid policy.

2.2. AID EVALUATION IN NATIONAL PUBLIC ADMINISTRATION

The second debate on aid evaluation in the 1960s and 1970s concerned the question who should conduct evaluation. Firstly, it is important to ask who was at the table discussing possibilities for aid evaluation. The attendance lists of the UNESCO, Berlin and Amsterdam conferences illuminate a great deal about who influenced institutionalization of new ideas in the development community. Secondly, participants also discussed who should actually conduct the evaluations. The two feasible options were either internal or external evaluation, the former being conducted by staff from operational departments and the latter by full-time evaluators, independent from the project or programme.

Starting with the first question, participants at the UNESCO conference included consultants such as Gaston Leduc, a professor of international economics at the University of Paris. Most consultants came from Western countries. Only two consultants from recipient countries attended, from Pakistan and India respectively. It is not specified whether the attendants were representatives of UN member states or simply experts in aid evaluation. Other attendants were western members of the Technical Assistance Board residing in Afghanistan, Bolivia, Lebanon, Libya, Pakistan and Turkey, UN representatives among whom renowned economist Hans W. Singer, and a delegate from the United States Foreign Operations Administration. Alva Myrdal chaired the meeting as the Director of UNESCO's Department of Social Sciences. The presence of Myrdal, Leduc and Singer exemplifies at least some academic engagement in the development of evaluation procedures.²¹⁷ The composition of the participants in Berlin was quite similar. In total, 53 delegates from Western MFAs, development agencies, multilateral institutions and academia attended. Both the developing world and the communist bloc were unrepresented. Why they were absent is unknown.²¹⁸

We also know quite a lot about the attendance at the Amsterdam conference. 41 delegates convened for three days in the Royal Tropical Institute. Among them were representatives of DAC

²¹⁷ UNESCO Special Summit 12, *Meeting on Criteria and Techniques of Evaluation of Technical Assistance for Economic Development Sponsored by UNESCO in Collaboration with the United Nations Technical Assistance Board, Geneva, 28-30 July, 1954: final report, SS/12*, available from unesdoc.unesco.org.

²¹⁸ German Foundation for Developing Countries, *Methods and Procedures of Evaluation in Development Aid*.

member states, UN agencies and international financial institutions. John Kaufmann, the Dutch Permanent Representative to the OECD and later Permanent Representative to the UN in New York, played an important part in organizing the conference. He wrote to colleagues at the MFA that France and the UK should not necessarily be invited since they allegedly had little experience with evaluation. Moreover, he noted that it would be useful to invite some delegates from recipient countries.²¹⁹ Ethiopia, Turkey, Argentina and India attended, while delegates from Thailand, Venezuela, Pakistan and Senegal were also invited but were not present.²²⁰ This time no academics attended the conference, although Leon Janssen, chair of the Werkgroep Janssen, was invited.²²¹ The reason of his absence is unknown, but we can assume that no other academics were there because the purpose of the meeting was specifically ‘to review the experience of Members with evaluation work and to exchange views on appropriate evaluation techniques and criteria.’²²² The DAC simply wished for an exchange of best practices within public administration, without opinions and experiences of academics.²²³

The question as to who should conduct aid evaluation inside the national public administration was discussed on several occasions. As mentioned before, the discussion circled around two options: either built-in evaluation, conducted by staff from operational or programme and budget departments, or evaluation conducted by full-time evaluators. In Berlin, Dieter Breitenbach, a researcher from the University of Saarbrücken, argued that ‘bureaucratization and the potential interest of administrators to report “successes” within their programmes are

²¹⁹ NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, letter 16 December 1969, Permanent Representative of the Kingdom of the Netherlands to the OECD, Paris, J. Kaufmann, to Minister for Development Cooperation Udink.

²²⁰ NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, letter 27 August 1970, S.M. Berline, Dutch Ministry of Finance, to M. Dembele, Directeur de Cabinet, Ministère du Développement Industriel, Dakar; NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, letter 27 August 1970, S.M. Berline, Dutch Ministry of Finance, to A. Casas Gonzales, Vice Minister of Development, Caracas; NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, letter 27 August 1970, S.M. Berline, Dutch Ministry of Finance, to K. Panpiemras, Senior Economist Fiscal and Monetary Section National Economic Development Board, Bangkok; NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, letter 27 August 1970, S.M. Berline, Dutch Ministry of Finance, to E. Ahmad, Deputy Economic Advisor Economic Affairs Division, Planning Commission, Islamabad.

²²¹ NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, letter 7 September 1970, S.M. Berline, Dutch Ministry of Finance, to dr. L.H. Janssen, Tilburg.

²²² NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, note 25 February 1970, OECD Secretariat, Paris.

²²³ NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, telex message with unknown date, Permanent Mission of the Kingdom of the Netherlands to the OECD, Paris, to ms. Leemans, Netherlands Ministry of Foreign Affairs.

dangers that are inherent to the method of built-in evaluation'.²²⁴ Therefore, evaluation conducted by research institutes should therefore complement built-in evaluation. Other delegates generally agreed that such mixed evaluation would suit the objectivity of evaluation best. Evaluation by MFA staff could either be done by operational departments or by a permanent evaluation unit. Such a unit would be particularly useful in short-term evaluations, while MFAs and agencies could hire external researchers from consultancy firms and academia for long-term assessments.

At the Paris Conference in 1968, the Deputy Director of USAID's Programme Evaluation Office Robert Hubbell shared his experience with external inspectors. He was very sceptical for two reasons. External inspectors had difficulties interviewing local USAID personnel, the latter being reluctant to disclose all information. Moreover, the external inspectors lacked sufficient understanding of local circumstances. These hindrances led to evaluation reports with no further use – policy improvement – for USAID. Hubbell therefore stated that evaluation should be part of the regular project management of operational staff. Of course, he acknowledged the risk of subjectivity, but 'if self-evaluation, even if not completely objective, leads to some improvements, we think it is better than an evaluation that was completely objective and which led to no action.'²²⁵ In order to enhance objectivity, Hubbell proposed that field workers hire consultants to guide the evaluation process. Joint evaluation with the host government could also contribute.²²⁶

In their report of the Amsterdam Conference in 1973, Kramer and Von Metzch also dealt with the question how evaluation should be institutionalized, and what parts of the aid activities evaluation should focus on. They noted that USAID presented their experiences with the Logical Framework Approach (LFA), already mentioned in Chapter I. USAID had held several seminars to promote the use of LFA between 1969 and 1973, visited by a number of DAC members, recipient countries and international organizations. Based on this information, the LFA was one of the best practices shared internationally.²²⁷

Furthermore, member states answered to a questionnaire in preparation of the Amsterdam Conference. The results showed that many countries preferred an evaluation unit inside programme and budget or policymaking departments. Other member states preferred a division of evaluation tasks between policymaking departments and operational departments. During the conference, delegates from DAC member states with evaluation institutionalized

²²⁴ German Foundation for Developing Countries, *Methods and Procedures*, 70.

²²⁵ OECD, *The Evaluation of Technical Assistance*, 83.

²²⁶ *Ibidem*, 83-84.

²²⁷ NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, memo 20 July 1973, Evaluatie-symposium DAC, Director DTH to Minister Pronk via plv. DGIS and SA.

outside operational departments noted that operational staff often complained about evaluators. Operational staff felt like they were monitored. On the other hand, some delegates argued that participation of operational staff would endanger the objectivity of evaluation studies. To some extent, this also applied for evaluation by policymaking staff. Some delegates therefore promoted evaluation by parties outside the MFA, such as consultants or research institutes. Others complained that this option was too costly, too scientific and therefore too little applicable for policy adaptation, and potentially impossible due to the outsiders' unfamiliarity with the respective projects or programmes. All in all, as of 1973 member states could still not consent to which of the options they deemed most suited.²²⁸ Hayes also supported built-in evaluation: 'Uncovering the results of individual projects is simpler and less costly and can often be carried on by the persons directly involved in the project.'²²⁹ The OECD Secretariat too doubted who should conduct the evaluation inside national aid agencies: '[s]o far there seems to be no generally accepted answer to this question, and the theories and practices advanced vary considerably.'²³⁰ The OECD Secretariat acknowledged UN practices that separated evaluation from the project operational units due to reasons of impartiality. Nonetheless, it argued that 'donor agencies should especially concentrate on (...) permanent evaluation, the only method which allows the objectives to be gradually narrowed down and to follow (...) the "continuous and progressive reformulation and adaptation" to which such objectives are subjected throughout the life of the project.'²³¹

In hindsight, it may sound unsurprising that academics, donor governments and multilateral organizations disagreed on the position of evaluation within national public administration. After all, the literature discussing the various aid evaluation procedures in the 1980s explain how these procedures differed among these actors. However, the fact that there was disagreement about the position of evaluators within development agencies might suggest why international cooperation regarding evaluation of bilateral programmes was so difficult. Furthermore, it can be said that recipient countries played only a minor role in the conferences analysed. Historical research could discover how many more such conferences were organised during the given period of time, and whether recipient countries participated to a larger extent than they did in the conferences analysed here. Nevertheless, the selection of conferences captures the most important events among DAC member states and therefore offers valuable insight into the discussions on aid evaluation in the near context of Dutch foreign aid. The next

²²⁸ NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, memo 20 July 1973, Evaluatie-symposium DAC, Director DTH to Minister Pronk via plv. DGIS and SA.

²²⁹ Hayes, Jr., *Evaluating development projects*, 17.

²³⁰ OECD, *The Evaluation of Technical Assistance*, 26.

²³¹ *Ibidem*, 32.

paragraph will shed light about how international cooperation was discussed as an instrument for improved aid evaluation.

2.3. INTERNATIONAL COOPERATION IN AID EVALUATION

First, it is important to address the perceived scientific feasibility of aid evaluation itself. After all, academics and other conference participants often discussed international cooperation in aid evaluation in combination with the complexity of the matter. Throughout all the conferences, it is clear that both staff members of multilateral institutions and governments struggled with the complexity of development and the complexity of *measuring* the effects of aid. For instance, although delegates at the UNESCO meeting had a straightforward idea why aid evaluation was necessary, they recognized that they were only at the beginning of a process. The proceedings mentioned that '[i]t is commonly said by field workers and Resident Representatives that the full complexity of any local situation in which they have to operate only becomes apparent after some years' work.'²³² Maybe the best characterisation of the discussion on the complexity was voiced in Berlin. One of the conference working groups focused on the evaluation of financial aid. Its summary noted that '[t]he Conference has shown that a wide variety of methods and procedures are available'.²³³ Other delegates disagreed, however. Professor Johan Galtung from the International Peace Research Institute in Oslo began his presentation at the Berlin Conference by saying that '[s]o far[,] satisfactory methods for evaluation do not exist. There is a choice of at least 300 definitions of factors leading to development to serve for evaluation purposes!'²³⁴

Nonetheless, conference participants were generally very supportive of joint evaluation. For instance, delegates at the Berlin Conference in 1966 assigned a crucial role in evaluation to recipient countries, 'since in many cases the recipient is in a better position to evaluate the overall impact in its economic, social and political development.'²³⁵ As mentioned earlier, USAID's Deputy Director of Evaluation Robert Hubbell argued in Paris in 1968 that joint evaluation could increase the objectivity of assessments conducted by field personnel. Mr. L.W. Norwood, Assistant Secretary at the UK's Ministry of Overseas Development, agreed and stated that joint evaluation 'is probably the most profitable and the most relevant form in which the evaluation of technical assistance (...) should take place.'²³⁶ The chairman of the Paris meeting, the Deputy Director-General of Economic Co-operation at the Spanish MFA Mr. F.J. Vallauré, added that 'nearly every speaker made [the point that evaluation] (...) is a joint exercise between donor and receiving

²³² UNESCO Special Summit 12, *Meeting on Criteria and Techniques of Evaluation of Technical Assistance for Economic Development Sponsored by UNESCO in Collaboration with the United Nations Technical Assistance Board, Geneva, 28-30 July, 1954: final report, SS/12*, available from unesdoc.unesco.org.

²³³ German Foundation for Developing Countries, *Methods and Procedures*, 133.

²³⁴ *Ibidem*, 93. Underlining from original source.

²³⁵ *Ibidem*, 8.

²³⁶ OECD, *The Evaluation of Technical Assistance*, 93.

countries, between donor countries and any international institutions which may participate in the project'.²³⁷ Thus, donors shared the opinion that joint evaluation was a highly valuable option for international cooperation.

Nevertheless, there were several existing practical hindrances to joint evaluation. Norwood saw in joint evaluation a huge challenge: 'It seems to me that the skills and techniques involved are of such a complex nature that it would be unreasonable to suppose that these skills and techniques are present in the host country'.²³⁸ Five years later, at the Amsterdam Conference in 1973, delegates argued that recipient countries often lacked qualified personnel. This was recognized in the official OECD summary: '[i]n practice (...) the participation of developing countries in evaluation activities has been relatively small, partly because of the difficulty of finding the necessary qualified people available at the time required.'²³⁹ Moreover, the Indian delegate allegedly mentioned that 'de kritische mentaliteit vereist voor evaluatie veelal tegenstrijdig is met het cultuurpatroon van ontwikkelingslanden'.²⁴⁰ Training of evaluators, offered by donor countries, would solve both the lack of personnel and the absence of a critical mindset.

The sources from the viewpoint of the UN listed similar counterarguments. UNITAR claimed that evaluation was a shared responsibility of donors, international organizations and recipient countries. However, participation of the latter complicated evaluation as governments of recipient countries often lacked the necessary institutional machinery, personnel or expertise.²⁴¹ According to Jackson, recipient countries had already 'asked for evaluations to be carried out, or have requested assistance in setting up efficient evaluation units staffed with qualified people.'²⁴² But according to the Pearson Commission, recipient countries had very little say in existing evaluation practices. Its report, itself based on consultations with representatives of the Global North and South,²⁴³ noted that the absence of joint monitoring and review was a

²³⁷ OECD, *The Evaluation of Technical Assistance*, 116.

²³⁸ Ibidem, 93.

²³⁹ OECD, *Aid Evaluation. The Experience of Members of the Development Assistance Committee and of International Organisations* (Paris 1975) 10.

²⁴⁰ NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, memo 20 July 1973, Evaluatie-symposium DAC, Director DTH to Minister Pronk via plv. DGIS and SA.

²⁴¹ UNITAR, *UN Development Aid*.

²⁴² Jackson, *A study of the capacity of the United Nations development system II*, 188.

²⁴³ Pearson et al., *Partners in Development*, viii-ix. The Commission organized meetings in Latin America, Africa, Asia and the Middle East where seventy governments of the Global South could voice their opinions on issues concerning development aid. The Commission also held written consultations with governments of the Global South who did not attend the regional meetings. Moreover, the Commission discussed development aid with the representatives of western governments and multilateral institutions. Interestingly, there had been no consultation between the Commission and communist countries. The authors do not state why, but they do express their regrets since cooperation between communist and non-communist countries 'would, of course, be most helpful to the cause of international development'.

fundamental shortcoming of development community. Donors did review, through the various processes of the DAC peer reviews, operations of various organizations such as the World Bank and the European Development Fund. According to the Pearson Commission, however, recipient countries did not participate in these processes.²⁴⁴ It did not discuss how such a review should be structured. Instead, the Pearson Commission simply called on the DAC to host a meeting with both donors and recipients to discuss the major obstacles recognized in their joint experiences.²⁴⁵

The OECD Secretariat claimed to already involve recipient authorities in their own evaluation activities. In its opinion, 'this method can best turn evaluation into an exercise most nearly approaching "constructive self-criticism and mutual criticism". The method is however subject to certain particular limitations and constraints'.²⁴⁶ Moreover, the OECD Secretariat criticized the international community for the lack of coordination:

virtually all agencies evaluate or attempt to evaluate their activities as if they were or had been the only ones to provide the country concerned with technical assistance. This introverted type of evaluation, while doubtless of some worth if the primary aim is to enable the donor agency to evaluate its own operational effectiveness, can however but lead to distortions and errors of judgment (...). The main responsibility for co-ordination of belongs to the authorities of the recipient country, but the obstacles towards its achievement are less matters of method than of expediency.²⁴⁷

More issues concerning joint evaluation arose during the Amsterdam Conference. Kramer and Von Metzsch mentioned that recipient countries feared political infringement. Some donors, on the other hand, claimed that African countries in particular resisted evaluation because evaluation would possibly delay the swift implementation of aid.²⁴⁸ According to the OECD, '[a] more serious difficulty is that evaluation results are often critical of the policies of the recipient country.'²⁴⁹ These points are in line with the remark of Stokke that 'demands for evaluation have met with little enthusiasm from the recipients of aid, though accepted as part of the conditions involved.'²⁵⁰ Nevertheless, delegates agreed that joint evaluation could play an essential role in improving the feedback loop to authorities of especially recipient countries. Although

²⁴⁴ Pearson et al., *Partners in development*, 227-228.

²⁴⁵ Ibidem, 19.

²⁴⁶ OECD, *The Evaluation of Technical Assistance*, 27-28.

²⁴⁷ Ibidem, 33.

²⁴⁸ NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, memo 20 July 1973, Evaluatie-symposium DAC, Director DTH to Minister Pronk via plv. DGIS and SA.

²⁴⁹ OECD, *Aid Evaluation*, 10.

²⁵⁰ Stokke, 'Policies, Performance, Trends and Challenges in Aid Evaluation', 21.

participants reached no concrete agreements, they at least agreed to return to their MFAs with the objective of expanding joint evaluation.

The conferences proved to be fertile ground for a new, permanent DAC working group. As mentioned in the introduction to this thesis, in 1979 DAC members pledged to revitalize their efforts towards aid evaluation. Two years later, the DAC Group of Evaluation Correspondents met for the first time in Paris. The terms of reference for this group were

to establish and report on the existing evaluation findings concerning the effectiveness of aid; to consider the question of feedback into policy-making; to report on ways in which evaluation could be supportive of public information; and to consider ways in which donors could support evaluation work in developing countries.²⁵¹

Thus, eventually, both lesson-learning, accountability, and participation of recipient countries became enshrined in joint international efforts towards improved aid evaluation. On the topic of multilateral evaluation, however, the conference reports have been very brief. Exceptions are the reports that included statements of multilateral organizations, who informed other conference participants about the evaluation of their own multilateral aid activities. Delegates of national governments debated multilateral evaluation of all activities – bilateral and multilateral – only to a limited extent.

Firstly, the Berlin Conference report notes that delegates discussed the idea of an ‘information retrieval system’ in recipient countries, coordinated by the UN resident representative, where donor countries and international organizations could share their evaluation reports.²⁵² Secondly, in his Wassenaar Conference report Scheltema, staff member of DFO, focused specifically on the discussions on the subject of evaluation. He noted that participants debated the difficulty of assessing indirect effects of individual projects. Such an assessment would be costly and overly complex, and he concluded that project evaluation was futile. Instead, evaluation should focus on overall assessment of a country’s socioeconomic progress. By mentioning that such studies were already being conducted in the context of the Second Development Decade, he seemed to imply that the Netherlands should no longer focus on national evaluation but should rather leave evaluation to the UN.

More information on multilateral evaluation can be found in the studies. Especially the Jackson Study is very vocal about existing evaluation practices at the UN. He complained about the existing practices of evaluation within the UN development system:

²⁵¹ OECD, *A History of the DAC Expert Group on Aid Evaluation*, 7-8.

²⁵² German Foundation for Developing Countries, *Methods and Procedures*, 2.

Quantitatively, so much evaluation is now being attempted that it almost amounts to international hypochondria. It is a definite brake on the capacity of the system. Qualitatively, the position is more disturbing for very few people have the necessary experience and understanding to take this exacting function successfully.²⁵³

Some pages later, Jackson reiterated his criticism and argued that continuing the existing evaluation practice could even have public repercussions:

Evaluation is not always balanced; almost invariably criticism far outweighs praise for a job well done. The present state of uncontrolled and uncoordinated evaluations, not always carried out by capable people, could, if taken to an extreme, greatly damage the image of UNDP and the UN development system, destroy public confidence, and endanger the most promising enterprise of the United Nations.²⁵⁴

A similar point was raised by the Pearson Commission, arguing that 'the multiplicity of reports and assessments originating in the numerous organizations (...) do not project enough unity of purpose to make them a rallying point for public support in the industrialized countries.'²⁵⁵ Abundant and uncoordinated evaluation caused ambiguous, confusing judgements and massive overhead costs deleterious to multilateral aid programmes.²⁵⁶ But this is not to say that both studies denied the necessity of strong multilateral evaluation. The Pearson Commission agreed that 'monitoring and assessment of performance is best done in a multilateral context in which donors and aid-receivers jointly review the past and plan for the future.'²⁵⁷

On a more practical level, UNITAR added to the criticism that 'it is very difficult to distinguish that part of benefits derived which can be attributed to the United Nations programmes as a whole.'²⁵⁸ After all, the UN financial efforts to development in a particular country or region only added up minimally compared to national and bilateral budgets. Therefore, UNITAR suggested that the UN would only evaluate at the project level. Moreover, the UN overemphasized the use of systematic, scientific, and uniform evaluation procedures. UNITAR denounced such procedures for three reasons. Firstly, projects covered all thinkable aspects of social, economic and cultural development. Secondly, the nature, executors, and recipients of assistance varied greatly. Thirdly, the UN offers technical assistance through thousands of

²⁵³ Jackson, *A study of the capacity of the United Nations development system I*, 11.

²⁵⁴ *Ibidem*, 52.

²⁵⁵ Pearson et al., *Partners in development*, 227-228.

²⁵⁶ Jackson, *A study of the capacity of the United Nations development system II*, 190.

²⁵⁷ Pearson et al., *Partners in development*, 17.

²⁵⁸ UNITAR, *UN Development Aid*, 16.

projects ranging from small-scale scholarships to projects worth millions of dollars. ‘Taking all these factors into account (...) it is clear that no single “scientific” method could apply uniformly.’²⁵⁹ Other obstacles to evaluation were ‘the question of quantification’, meaning that some development projects produced outcomes that were hardly quantifiable, and ‘cost considerations’, meaning that investments in evaluation immediately decreased the number of projects executed worldwide. Regarding the latter, the authors quoted the OECD that ‘[evaluation] studies should [also] include a cost benefit analysis of evaluation itself.’²⁶⁰ All in all, it seems from these sources that multilateral evaluation was far from ideal during the late 1960s and early 1970s.

CONCLUSION

We can first distil from the conference reports the continuous presence of academics. The first steps towards professionalization of evaluation were clearly a joint academic and administrative effort. For instance, nearly half the attendees of the Berlin Conference were from universities and other research institutes. Moreover, Kramer also notified his supervisor that DTH was working on a strategy how to include evaluation in Dutch development aid, and mentioned that Wageningen University also worked on such a document, specifically for agricultural projects.²⁶¹ This shows that in the Netherlands, the cooperation between the MFA and academia was far from over in 1973 even though academics did not attend the Wassenaar Conference.

Secondly, actors generally recognized the complexity of aid evaluation. The OECD Secretariat was sceptical of the state of the art in aid evaluation in the late 1960s: ‘[i]t would certainly be rash, in view of our present limited knowledge of such a highly complex process, to claim that evaluation can provide a “measurement of technical assistance productivity”, unless “measurement” is interpreted in an unusually broad sense.’²⁶² The fact that the development community did not fully understand development, had two implications for aid evaluation. First of all, it was hard to evaluate something without the necessary knowledge to do so. Secondly, because development was so complex, actors decided that evaluation was in fact highly necessary to improve their knowledge.²⁶³ The complexity of aid evaluation was especially recognized in relation to both joint and multilateral evaluation. Despite the interest in such methods, the conference discussions lacked concrete plans for developing methods of joint evaluation.

²⁵⁹ UNITAR, *UN Development Aid*, 18.

²⁶⁰ *Ibidem*, 19-20.

²⁶¹ NL-HaNA, 2.05.313, Evaluatie van de seminars, gehouden op 28 – 30 oktober te Wassenaar en van 27 – 29 Juni 1973 te Amsterdam, inventory number 21727, memo 20 July 1973, Evaluatie-symposium DAC, Director DTH to Minister Pronk via plv. DGIS and SA.

²⁶² OECD, *The Evaluation of Technical Assistance*, 9.

²⁶³ Jackson, *A study of the capacity of the United Nations development system II*, 453.

Moreover, recipient countries participated in the discussions analysed by this chapter to only a limited extent.

Thirdly, it seemed as if multilateral agents were more concerned with accountability vis-à-vis national governments than the national governments vis-à-vis their legislative bodies. This could be a question of legitimacy: the UN was still a young organization experimenting with all forms of policies and being dependent of its existence on the consent of its member states. These factors made the UN a unique actor in the development community. Moreover, while national governments held the UN accountable, it is uncertain to which extent parliaments held their governments accountable for successful aid delivery. Chapter III will answer this question concerning Dutch aid, but it can already be said that further research can be valuable to understand this dynamic in other donor countries.

Finally and most importantly, this chapter revealed that aid evaluation was an important international topic long before the establishment and expansion of aid evaluation practices. This chapter demonstrates a lively interplay between ideas and practice in an age of experimentation, in which international cooperation can be witnessed in several forms. In fact, evaluation was abundant within the UN system. In describing the jungle of evaluation activities there, Jackson argued that 'it is (...) certain that evaluation will defeat its own purpose unless it is carried out in an orderly and co-ordinated fashion, with clearly-defined objectives and consistent procedures applied by competent evaluators.'²⁶⁴ This is a clear sign that aid evaluation was still immature at the end of the 1960s, and that UN agencies all experimented in their own ways. The concept of policy intimacy, then, is again very applicable to this context.

²⁶⁴ Jackson, *A study of the capacity of the United Nations development system II*, 189.

III. FROM IDEA TO PRACTICE: INSTITUTIONALIZING DUTCH AID EVALUATION, 1964-1977

De echte betekenis van evaluatie – de “impact-evaluation” – kan alleen op multilateraal, op globaal niveau worden bereikt. – Dutch Minister Bé Udink for Development Cooperation in 1969²⁶⁵

Now that Chapter II has outlined the transnational development of aid evaluation, Chapter III zooms in on the Dutch experiences with aid evaluation before the establishment of the IOV in 1977. In general, Dutch foreign aid expanded and professionalized substantially during the 1960s and 1970s. Aid evaluation in the 1960s revolved around the Werkgroep Janssen. Minister Udink (1967-1971) acknowledged in an interview with historian Marc Dierikx in 2006 that this four-year evaluation of the national aid apparatus and aid policy, conducted between 1965 and 1969, created a stir at the MFA.²⁶⁶ Remarkably, the literature has neglected the Werkgroep Janssen quite a lot. This was recognized by Rob van den Berg in 2001, former Director of the IOV. Van den Berg believed that the Werkgroep itself was partly to blame, due to its ‘boring’ conclusion: an evaluation of the overall Dutch aid policy was allegedly impossible due to the complexity of development and lack of subject knowledge. The conclusion was simply not exciting for historians.²⁶⁷

Up to now, the existing literature does not reveal what the Werkgroep actually concluded, why its research was relevant for aid evaluation in the Netherlands and happened to experiments with aid evaluation afterwards. This chapter therefore deals with the question how aid evaluation was discussed at the Dutch MFA, in parliament and press between 1964 and 1977. This period was rich in experiments and encapsulated the institutionalization and professionalization of Dutch aid evaluation. These trends were embedded in a context of three main events: the study of the Werkgroep Janssen and the publication of its final report in 1969; the political discussion sparked by this report; and the formal institutionalization of evaluation and inspection during the tenure of Minister Pronk (1973-1977). At these moments, lively discussions surfaced on the purpose, the national institutionalization and the international cooperation concerning aid evaluation. These three categories will make up the structure of the analysis below. The records

²⁶⁵ *Verslag der handelingen van de Tweede Kamer der Staten-Generaal 1968-1969*, 3254.

²⁶⁶ M.L.J. Dierikx, ‘Ontwikkelingshulp is op zijn best *marginal but perhaps in some cases vital*: interview met drs. B.J. Udink’ in: L.J. van Damme, M.G.M. Smits (eds.), *Voor de ontwikkeling van de derde wereld. Politici en ambtenaren over de Nederlandse ontwikkelingssamenwerking, 1949-1989* (Amsterdam 2009) 83.

²⁶⁷ R.D. van den Berg, ‘Leert de hulp van lessen uit het verleden? Doe wel en zie niet om’ *Internationale Spectator* 5 (2001) 253-260.

in the Dutch National Archives show that the Netherlands was extremely eager to gather knowledge from other governments, as well as to participate in international initiatives by the OECD, World Bank and the UN.

The analysis is based on several sources. First of all, a significant number of memos from 1969 have been preserved in the archives of the MFA. In addition, the NAR issued two recommendations on Dutch aid evaluation in February and August 1972. Secondly, the parliamentary proceedings contained a letter of Minister Udink to the Lower House and the lively parliamentary debate that followed in June 1969. Thirdly, Delpher offers editions of newspapers writing about aid evaluation. *De Volkskrant* and *Het Parool* already published articles about the Werkgroep's report and Udink's response before the parliamentary debate. Other newspapers wrote about the debate itself the day after.²⁶⁸ Moreover, *Trouw* and *De Volkskrant* also referred to the NAR recommendations in 1972 and 1973.

This chapter leaves out several matters related to Dutch aid evaluation. For instance, it does not analyse the evaluation note of Dutch aid to Indonesia published by Pronk in 1977. Although the evaluation became a bombshell at the MFA, the evaluation was not part of the discussion on the actual institutionalization of evaluation. It is therefore largely irrelevant for this thesis. This chapter further leaves aside the study conducted by the Algemene Rekenkamer in the same year, nor Pronk's written response to that document. This study can be typified as an audit, not as an evaluation.

3.1 SETTING THE AGENDA: THE WERKGROEP JANSSEN

Being the first large-scale review of Dutch foreign aid, the case of the Werkgroep offers great insight into the development of Dutch aid policy and evaluation practices. Although Minister of Foreign Affairs Luns announced the onset of bilateral aid in 1962, the Dutch government preferred coordination of aid activities via multilateral organizations throughout the 1960s. The preference for multilateralism also applied for aid evaluation. Luns shared in his 1962 policy note that the first step towards international coordination was to share information on aid activities, 'waarvan evaluaties van de ondernomen hulpactiviteiten een belangrijk onderdeel zullen uitmaken'.²⁶⁹ Moreover, he recognized the importance of aid evaluation:

Aan het onderwerp van de evaluatie van het hulpverleningsresultaat zal ik in ieder geval veel aandacht besteden. In het verband van de Verenigde Naties en van de O.E.C.D. is men bezig aan zeer belangrijke research; dit wil ik graag ook in Nederland zorgvuldig bekijken.

²⁶⁸ *Het Parool, Trouw, Nieuwsblad van het Noorden, Algemeen Handelsblad, Nederlands Dagblad, De Tijd, Het Vrije Volk, Leeuwarder Courant.*

²⁶⁹ Handelingen der Staten-Generaal, 1961-1962, Bijlagen 2e Kamer, no. 6817 – 1.

Het grote belang van de evaluatie is, geloof ik, te weten te komen welke soorten van hulpverlening en welke methodiek het grootste en meest blijvende effect heeft op de ontwikkeling. Op dat punt is onze kennis nog maar zeer betrekkelijk.²⁷⁰

In 1966, Minister Bot (1965-1967) published his policy note and shared with Parliament that multilateral experiments with evaluation were underway. He also explained why he partly deviated from the multilateral approach of Luns and instituted a national evaluation team, the Werkgroep Janssen. The overarching motive behind this decision lay in the perceived complexity of development aid: *'Een afdoende wetenschappelijke basis voor het te voeren beleid ontbreekt.'*²⁷¹ Bot hoped that the Werkgroep would assess which policies and methods did or did not work.

In his first year as DGIS, Jan Meijer requested professor Leon Janssen to explore whether a scientific evaluation of Dutch aid policy would be possible. Janssen taught development economics at the universities of Tilburg and Nijmegen.²⁷² Several preparatory meetings were held at the MFA before the research of the Werkgroep took off. Academics from six Dutch universities and research institutes were invited to join, among them Tinbergen as the most renowned Dutch scholar on development.²⁷³ Van Dam, being Meijer's closest advisor and a prominent academic, was also present. He expected that an evaluation would not lead to big conclusions on the short term, but emphasized during one of the meetings that the MFA simply wanted to know whether Dutch aid in the existing form was efficient. The MFA further requested a review of 'the effectiveness of every guilder'.²⁷⁴

Others also voiced concerns about the feasibility of the study. Janssen discussed the research proposal with professor W. Brand, professor J.A. Ponsioen and H. Bos, all working at the *Nederlands Economisch Instituut*. They first of all agreed that a study of the academic literature was necessary, because 'het probleem van de evaluatie van ontwikkelingswerk houdt reeds vele geesten bezig (...) [en] om te komen tot een verdere uitbouw van de methodologie van evaluatie.'²⁷⁵ But in general, they did not expect 'spectacular results' from an evaluation of aid

²⁷⁰ *Handelingen der Staten-Generaal*, Tweede Kamer, 1963-1964, 11 December 1963.

²⁷¹ Minister Bot, 'Nota hulpverlening aan minder-ontwikkelde landen' (1966) <http://resources.huygens.knaw.nl/ontwikkelingssamenwerking> (21 March 2019), 12.

²⁷² *Nederlandse ontwikkelingssamenwerking. Bronnenuitgave Deel 2 1964-1974*, red. M.L.J. Dierikx (Den Haag 2003) 140-141.

²⁷³ Nederlands Economisch Instituut (N.E.I.), Institute of Social Studies (I.S.S.), Instituut voor sociaal-economisch minder-ontwikkelde gebieden (ISMOG), Instituut voor de Sociologie der niet-westerse volken, Landbouwhogeschool, Instituut voor Ontwikkelingsvraagstukken (IVO). NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6763, letter 4 January 1965, A.J. Piekaar, Ministry of Education, Art and Science, to J. Meijer, Director-General of International Cooperation, Ministry of Foreign Affairs.

²⁷⁴ NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6763, report 10 February 1965, bespreking inzake de evaluatie van de hulpverlening.

²⁷⁵ NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6763, letter 10 December 1964, L.H. Janssen, SJ, to Jan Meijer.

policy, at least not on the short term. Janssen, reflecting on the discussions with fellow academics, wrote to Meijer that most aid projects were still relatively young which made them inapplicable for evaluation.²⁷⁶ Minister Bot himself also shared his doubts about the feasibility of the research. In the margins of a letter to Janssen he asked whether he could expect concrete results? Moreover, he asked whether foreign governments also dealt with aid evaluation. These questions were deleted from the final version, for reasons unknown. One of his staff members possibly gave him the answers, but nonetheless he confirmed to Janssen in the final version that existing evaluation methods were still unsophisticated, making thorough scientific evaluation difficult to conduct.²⁷⁷ All in all, the source material concerning the onset of the Werkgroep fully reflects the perceived complexity of aid evaluation voiced in the international sources analysed in Chapter II.

The Werkgroep published a number of reports of which only the final one, the most comprehensive, was published. Janssen dispatched preliminary reports to the government only. In the final report the authors stated that, due to the complexity and limited knowledge of development, they could not answer the question what type of aid policy generally delivered the best results. Janssen argued in *De Volkskrant* that ‘de voordelen van de hulp vaak in het menselijke, niet-materiële vlak liggen en moeilijk in cijfers of getallen zijn om te zetten. “Op dat gebied wordt van de economie en de sociale wetenschappen teveel verwacht’.²⁷⁸ Science could not yet fix everything. According to Janssen, no other country had ever conducted an evaluation that would answer such a question. He also stated that he had warned the government about the impossibility and the financial costs of such a research from the onset of the project – which is confirmed by the source analysis above.²⁷⁹

Nevertheless, the Werkgroep did make a number of notable conclusions.²⁸⁰ Most relevant for this thesis is the Werkgroep’s conclusion about the necessity of permanent aid evaluation. As a benchmark, they found that recommendations from progress reports sent to The Hague were not sufficiently incorporated in new policies.²⁸¹ This conclusion relates to the point made by Dierikx, outlined in Chapter I, that senior staff at the MFA sometimes neglected the advice of project experts. Subsequently, the Werkgroep explored ideas to improve the feedback between

²⁷⁶ SOURCE

²⁷⁷ ‘Geen enkel land is er tot nu toe in geslaagd dat te doen. (...) Vanaf het begin hebben we tegen de regering gezegd, dat zo’n stuk beleid niet mogelijk is. Alleen al de kosten die daaraan verbonden zijn staan in geen verhouding tot het nut van zo’n analyse.’ In: *Nederlandse ontwikkelingssamenwerking II*, 222-223.

²⁷⁸ *De Volkskrant*, ‘Prof. Janssen over kritiek: Udink verwachtte te veel van beleids-onderzoek’, 6 June 1969.

²⁷⁹ *De Volkskrant*, ‘Prof. Janssen over kritiek: Udink verwachtte te veel van beleids-onderzoek’, 6 June 1969.

²⁸⁰ This paragraph only lists the conclusions relevant for this thesis.

²⁸¹ Werkgroep Evaluatie Nederlandse Ontwikkelingshulp, *Rapport Evaluatie van de Nederlandse Ontwikkelingshulp* (Tilburg 1969) 18.

The Hague and project experts. In doing so, the researchers touched on most of the dilemmas discussed in the international aid community.

Concerning the national institutionalization of evaluation, the Werkgroep offered two options. Firstly, the Werkgroep recommended 'administrative control', that would evaluate the efficiency of a project. This is similar to built-in evaluation and would be conducted by operational departments. The second option would focus on the analysis of a project's long-term and wider significance. Such an evaluation could not be conducted by operational departments only. After all, this type of evaluation required more technical expertise in evaluation. Therefore, the Werkgroep suggested that a permanent evaluation team consisting of insiders and outsiders might be the best solution. Insiders would be staff from operational departments, bringing project expertise to the table, while outsiders could be hired from consultancy firms and academia. As a proof of viability, the Werkgroep explicitly referred to the success of mixed evaluation within USAID.²⁸² According to the Werkgroep, some aid projects also required a long-term evaluation conducted a number years after implementation. Those who executed the project may not be able to serve as evaluators anymore, making evaluation fully dependent on the input of outsiders. These evaluations would be conducted for individual projects and programmes. In addition, evaluations of the entire Dutch aid apparatus should take place, similar to the study conducted by the Werkgroep. Such grand evaluations would encapsulate the evaluation of projects and programmes and ensure that the feedback loop from project executives to senior staff at the MFA would close.²⁸³

Concerning the question on international cooperation, the Werkgroep considered whether recipient countries should participate in evaluation practice. Evaluation could only work when it included the opinions of recipient countries, simply because development aid was a joint activity of donor and recipient. The Werkgroep consulted experts for advice how to incorporate recipient countries, but this offered no clear-cut solution about how to institutionalise joint evaluation. The researchers eventually dismissed the matter because it was 'a matter of international cooperation'.²⁸⁴ The authors did not specify what they meant with this suggestion. Whether they really intended to include evaluation by recipient countries is unclear. After all, the authors wrote that an experienced evaluator is often capable enough to discover the assessments made by recipients. All in all, the Werkgroep did not want to meddle in international developments concerning aid evaluation and therefore only recommended the establishment of

²⁸² Werkgroep Evaluatie Nederlandse Ontwikkelingshulp, *Rapport Evaluatie van de Nederlandse Ontwikkelingshulp* (Tilburg 1969) 267-268.

²⁸³ Ibidem, 267-268.

²⁸⁴ Ibidem, 266-267.

permanent evaluation on the national level. They did not reflect on the possibility of multilateral evaluation.²⁸⁵

3.2. STIRRING UP THE HORNET'S NEST: THE REPORT AT THE MFA, IN PARLIAMENT AND MEDIA

Despite all doubts about the feasibility of the Werkgroep's assignment, there was widespread interest in its final report. This was a clear sign that the Dutch development community, including politicians, public servants and academics, shared the feeling that evaluation was useful.²⁸⁶ However, the report of the Werkgroep received a lot of criticism at the MFA in early 1969. Van Dam even advised against informing parliament about the report because it would create a 'polemical situation'.²⁸⁷ He noted that aid evaluation was still underdeveloped. Publicly discussing the subject would therefore only risk stirring up the hornet's nest. Udink countered that his department was strong enough to welcome the report's critique and refrained from denouncing the report all at once. He set aside Van Dam's advice, and tried to circumvent growing public criticism that aid was a waste of money. He shared his opinion about the report through a separate letter to the Tweede Kamer. The letter sparked considerable parliamentary debate, which ended with two proposed but ultimately rejected motions to establish a permanent evaluation unit.

3.2.1. Lesson-learning or accountability?

The first question to be answered concerns the purpose of aid evaluation. The Werkgroep itself focused mostly on lesson-learning. In emphasizing the necessity of a feedback loop in the aid apparatus, the authors never really thought of evaluation as a means for parliament to hold the government accountable. Memos from the MFA paint a similar picture. As the report was first circulated at the MFA, its staff had considerable time to discuss its findings before MPs got hold of it. A small number of experts within the MFA emerged. Among them was A. van der Goot, staff member of DTH who also seated in the NAR.²⁸⁸ He wrote a seventeen-page-long memo discerning the purpose and types of evaluation. He defined aid evaluation, based on the discussions in Berlin and Paris, as 'het geven van een oordeel over de effectiviteit van de diverse

²⁸⁵ Werkgroep Evaluatie Nederlandse Ontwikkelingshulp, *Rapport Evaluatie van de Nederlandse Ontwikkelingshulp*, 271-274.

²⁸⁶ NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, letter 14 May 1969, ir. F. Deeleman, Department of Agricultural Aid to Developing Countries, Ministry of Agriculture and Fisheries, to the Coördinatie Commissie inzake Hulpverlening aan Minder Ontwikkelde Landen, Ministry of Foreign Affairs.

²⁸⁷ *Nederlandse ontwikkelingssamenwerking* III, 324n; NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memorandum 15 April 1969, F. van Dam to J. Meijer.

²⁸⁸ Leidse Courant, 'Lintjesregen', 27 april 1973.

hulpverleningsactiviteiten, gebaseerd op het gebruik van wetenschappelijk verantwoorde methoden en technieken.’²⁸⁹ In addition, the purpose of evaluation was ‘om het nuttig effect van de verleende hulp zoveel mogelijk te verzekeren, en om bij te dragen tot het ontwikkelen van adequate richtlijnen en procedures voor de verlening van die hulp.’²⁹⁰ According to Van der Goot, delegates in Paris concluded that the key purpose of evaluation was measuring aid’s efficiency, i.e. the useful application of available financial and human resources. Aid evaluation had the same purpose as productivity enhancing tools in corporate management. Just like corporate managers devised measures on the basis of monthly, quarterly or annual productivity reports, policy makers in the aid apparatus should know what policies are inefficient in order to modify them. This comparison is so remarkable because it implies that the public sector should, or at least could, be managed with private sector-like tools. This clearly resembles the NPM thought coined earlier in this thesis. All in all, evaluation primarily would primarily serve feedback from below to managing circles up top. This clearly suggests that MFA staff prioritized lesson-learning over accountability.

Udink himself also implied that lesson-learning was evaluation’s main purpose. He argued in his note to Parliament that improving the quality of Dutch aid required continuous re-evaluation of goals, review of aid practices, and adjustment of policies. Udink wrote that these three components were part of Dutch aid from its inception through government reports sent to parliament annually.²⁹¹ Other sources also demonstrated that the government viewed lesson-learning as the main purpose of aid evaluation. The Ministry of Agriculture and Fisheries (L&V), for example, advised Udink to follow the report’s recommendation and set up an evaluation unit at the MFA. L&V argued that such a unit would help the government in understanding the complex process of agricultural development. In other words, L&V supported the idea that evaluation must above all provide feedback.²⁹²

The NAR used the same line of reasoning in favour of lesson-learning when it offered recommendations to Minister Boertien, Udink’s successor, in February 1972. The NAR first of all wrote that

zelfs zij die per ambassade de technische samenwerking als specifiek terrein van belangenbehartiging toegewezen krijgen, slechts bij uitzondering enige tijd bij de Directie

²⁸⁹ NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memorandum 1 April 1969, A. van der Goot.

²⁹⁰ Ibidem.

²⁹¹ *Handelingen der Staten-Generaal*, Tweede Kamer, 1968-1969, Bijlage 9800, hfst. V, no. 62.

²⁹² NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, letter 14 May 1969, ir. F. Deeleman, Department of Agricultural Aid to Developing Countries, Ministry of Agriculture and Fisheries, to the Coördinatie Commissie inzake Hulpverlening aan Minder Ontwikkelde Landen, Ministry of Foreign Affairs.

Internationale Technische Hulp ter voorbereiding doorbrengen en ook overigens nauwelijks enige speciale scholing ontvangen. Op die manier is het geen wonder, dat tussen “Den Haag” en “de periferie” de communicatie niet optimaal kan functioneren.²⁹³

DTH and its antennae in the field, embassies and consulates, did not sufficiently communicate. The NAR followed up on this in another recommendation, published in August 1972. Here, the NAR again argued that the main purpose of evaluation would be the improvement of aid quality. Relatedly, the NAR framed lesson-learning not just in terms of organizational knowledge among MFA staff. The NAR also preferred evaluation for the purpose of increasing scientific understanding of development.²⁹⁴ This preference related to a message of Minister Bot in his 1966 policy note, in which he acknowledged very clearly that science did not yet understand development in all its complexities. The Dutch government could, in the eyes of Bot and the NAR, help scientists by expanding aid evaluation.

But the NAR favoured aid evaluation for more than just lesson-learning. Its August 1972 recommendations also listed accountability as a purpose of aid evaluation. The NAR admitted that evaluation can serve ‘om parlement, kiezers en belastingbetalers het vertrouwen te geven dat de gelden die hieraan worden uitgegeven, zo doelmatig mogelijk worden besteed.’²⁹⁵ On the other hand, evaluation could also display to parliament and the general public that aid money was spent unsuccessfully, which could result in a decreasing amount of trust. This relates to the point made by Guisset, the French delegate at the Paris Conference in 1968. Nonetheless, the NAR argued

in algemene zin is openbaarheid van de resultaten van ontwikkelingssamenwerking gewenst. De resultaten van evaluatie-onderzoek moeten op echter begrijpelijke wijze worden gepresenteerd, opdat het grote publiek meer inzicht krijgt, ook in de complexiteit en de beperkte mogelijkheden van hulpverlening. (...) Daarbij dient echter bedacht te worden dat publicatie van mislukkingen nadelig kan werken voor het vertrouwen in de ontwikkelingssamenwerking.²⁹⁶

The NAR was not the only actor who promoted accountability. Commentators expressed their support for accountability in the press. Several articles highlighted that that aid policy had expanded and had become more complicated. Therefore, their authors felt that evaluation was a necessary element in fields of complex government policies. The Minister should present his

²⁹³ NAR, *Advies Technische Hulp*, 41.

²⁹⁴ NAR, *Advies Evaluatie van de Nederlandse Ontwikkelingshulp*, 16-17, 21.

²⁹⁵ *Ibidem*, 6.

²⁹⁶ *Ibidem*, 22, 42.

actions and results to the public. An opinion piece by an unknown author in the *Algemeen Handelsblad*, argued that

[j]uist nu de details van het beleid van meer dan één minister met de dag technischer en ingewikkelder worden zouden parlement en pers gebaat zijn met analyses door deskundigen ook van andere onderdelen van het regeringsbeleid.²⁹⁷

Another newspaper, *De Volkskrant*, was frustrated by the reluctance of the NAR about the publication of results of aid evaluation. In the eyes of the newspaper's editors,

[o]ntwikkelingssamenwerking dient een volwassen stuk van de staatstaak te zijn (of te worden), en daarover dienen de burgers volledig en oprecht te worden geïnformeerd. Achterhouden van wezenlijke informatie om de gunst van de burgers voor een beleid te winnen, komt neer op manipulatie. Bovendien zal openheid ook over het falen, de critici in staat stellen hun streven naar meer en kwalitatief betere hulp met feiten te onderbouwen. Gefundeerde kritiek is voor elk beleid onmisbaar.²⁹⁸

Labour Party leader Joop den Uyl also pushed for more public accountability, even though his remark was not necessarily about the evaluation of aid. He argued that in favour of more transparency towards parliament on aid spending in 1969, especially because the 'hundred million guilders' of tied aid were in essence public funding to Dutch companies.²⁹⁹ More generally, the NAR agreed that evaluation serves as a necessary public control in modern public administration: 'De overheid moet in de eerste plaats voorzien in een goede bedrijfsvoering, in die zin dat er zowel een goede controle op de besteding van de gelden als een operationele controle op de werkzaamheden is.'³⁰⁰ Based on this, we can at least conclude that actors in the development community had their eyes on accountability.

3.2.2. National aid evaluation

Concerning the institutionalization of evaluation in national governance, the Werkgroep recommended the MFA to erect a separate, permanent evaluation team. The team would be staffed by outsiders – consultants and academics – and insiders from operational departments such as DTH and DFO. The recommendation led to significant discussion and eventually to the

²⁹⁷ *Algemeen Handelsblad*, 'Symposium', 14 June 1969.

²⁹⁸ *De Volkskrant*, 'Adviesraad over evaluatie-beleid. Hulp aan Derde Wereld toetsen op z'n effect', 19 September 1972.

²⁹⁹ Trouw, 'Udink: Nederlandse hulp is niet commercieel', 13 June 1969.

³⁰⁰ NAR, *Advies Evaluatie van de Nederlandse Ontwikkelingshulp*, 15.

establishment of the IOV in 1977. The first major response within the MFA came from Van der Goot. He summed up five possible methods of aid evaluation, divided in two categories: operational and substantive evaluation. Operational evaluation entailed the assessment of productivity carried out in 1) progress reports, 2) effectiveness studies and 3) efficiency studies. Evaluations of the operational type focused on the relation between means, goals and results. Substantive evaluation, on the other hand, emphasized the nature and quality of goals and results. Such evaluation assessed the wider effectiveness of a project or programme, taking into account development as a holistic national action plan. Substantive evaluation entailed both 4) significance studies – what does a certain project or programme contribute to a national development objective? – and 5) impact studies – how does the entire body of aid activities contribute to the entire national development plan? Substantive evaluations needed macroeconomic data, data that only the host government could acquire. Thus, international cooperation was highly necessary for substantive evaluation.³⁰¹

In 1969, MFA staff debated these categorizations in relation to the question whether evaluation should be permanent. A memo written by an unspecified DTH staff member noted that the first step towards decent and cost-efficient evaluation was above all to recognize that evaluation had to be a permanent and integrated part of the aid apparatus. This recognition would result in a significant change in the entire lifespan of a project, from preparation to implementation and its further impact on future policy. After all, administrators should then ensure that at every stage of the project their actions would serve the eventual evaluation. That meant, for example, that during project preparation all expectations, such as goals and costs, should be clearly stated. Progress reports during the implementation phase should be set up in a manner that would suit the requirements of evaluation. Although the author of the memo clearly argued that evaluation should be an integral element in Dutch aid, the author did not specify who would be responsible to conduct the actual evaluation.³⁰²

The NAR also engaged in the discussion about where to institutionalize evaluation, i.e. who should conduct it. The NAR distinguished between two types of evaluation: internal and external. Internal evaluation could be done either by a separate evaluation unit or by planning or operational departments of the MFA, such as DTH or DFO. There were several objections against evaluation conducted fully or partly by the latter. For instance, their staff members were unqualified and had limited time for conducting evaluation. Staff members also lacked objectivity due to the ‘psychological investments’ they put in their projects.³⁰³

³⁰¹ NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memorandum 1 April 1969, A. van der Goot.

³⁰² NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memo date unknown, DTH.

³⁰³ NAR, *Advies Evaluatie van de Nederlandse ontwikkelingshulp*, 15-19.

3.2.3. International cooperation

The question whether evaluation should be settled in a separate team at the MFA was a major one during the parliamentary debate in June 1969. However, Minister Udink and MPs did not discuss this option in relation to the option of built-in evaluation. Instead, they discussed alternatives that lay across the borders: either multilateral cooperation or joint evaluation. The report of the Werkgroep itself also received international attention. UNDP and the UN Food and Agriculture Organization (FAO) asked for a copy of the report, and UNDP official Dick Papousek complimented the members of Werkgroep for not avoiding to criticize government policy.³⁰⁴ The research on the fellowship programme, one of the appendices to the final report, was also shared with foreign governments and international agencies.³⁰⁵ These are some of the many indicators that the Netherlands intensively participated in the international development of aid evaluation.

The analysis concerning international cooperation should also start with the memo by Van der Goot. He summarized developments in foreign aid policies of other donors from the late 1950s onwards. In his opinion, the rationalization of aid policies in the US, UK, France and the Netherlands in previous years ran parallel with a sudden interest in the efficiency and effectiveness of aid policies. Another MFA staff member, Mr. Erath, acknowledged that the Netherlands were far behind in international perspective. He and Van der Goot often referred to Germany, the US and Sweden for examples of strong evaluation practices.³⁰⁶

Udink seemed personally interested in international practices. He sent a memo with two questions concerning aid evaluation to his senior staff: 'Is het Nederlandse toetsingssysteem voor technische hulp even goed als dat in de V.S./Zweden/Duitsland en zo neen waarom niet en wat doen wij er aan? Zijn de landenevaluaties van de Wereldbank publiek c.q. beschikbaar voor de Tweede Kamer?'³⁰⁷ Van Warmenhoven, Director of DTH, responded to the first question. With regard to German evaluation practices, he answered that their results were not publicly accessible yet. The German aid department gave participating academics the opportunity to

³⁰⁴ NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, letter 30 May 1969, D.A. Papousek, UNDP, to G. Ringnalda, Directorate-General International Cooperation, Netherlands Ministry of Foreign Affairs; NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, letter 15 April 1969, O. van Teutem, Evaluation Branch Economic Analysis Division FAO, Rome, to DGIS.

³⁰⁵ NL-HaNA, 2.05.313, inventory number 27541, letter 14 May 1968, C.A. van Ravenswaaij to R.H.M. Koesoemo Joedo.

³⁰⁶ NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memorandum 17 April 1969, Plv. Chef AJV to Jan Meijer; NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memo date unknown, DTH; NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memo date unknown, author and recipient unknown.

³⁰⁷ NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memorandum 9 June 1969, Minister for Development Cooperation Udink.

publish their results first, as a way of thanking them for their service. Compared to the Dutch system, German evaluation primarily emphasized cost-benefit analyses. The Dutch government instead believed that development could not be measured just in terms of cost and benefit. The Dutch system therefore more resembled the Swedish system, in particular because both practices occurred continuously on the basis of progress reports, effectiveness studies, and debriefing of experts. The Swedes were, just like the Dutch, very much interested in the improvement of aid quality.³⁰⁸

In a different memo, Udink shared his disappointment that the Werkgroep did not study multilateral actions on aid evaluation. Therefore, he requested even more information about evaluation practices of the World Bank to Minister Witteveen of Finance. The latter could have access to this information through his contacts with Piet Lieftinck, former Dutch Minister of Finance and Executive Director of the World Bank and IMF. Udink was primarily interested in answers to two questions: how did the World Bank define aid evaluation? And did World Bank evaluations measure results of specific projects or of general country development strategies?³⁰⁹ In sum, these examples present a clear view that he was highly interested in learning from foreign and multilateral initiatives.

Regarding multilateral evaluation, Van der Goot noted UN country studies in Chile, Thailand and Tunisia and the OECD studies in Greece, Turkey and Spain conducted some years earlier. These were categorized as impact studies, and experienced at least one of the problems of joint and multilateral evaluation described earlier in Chapter II: it was impossible to cover all development activities in a country without having all relevant actors participating in evaluation. The UN could only assess the impact of its own operations, while the OECD was unable to include operations of private aid agencies. The OECD even concluded that impact studies could only succeed when projects of all agencies, private and public, were coordinated better internationally.³¹⁰ Nonetheless, both Van der Goot and Erath both argued that the development community recognized the value of aid evaluation for the improvement of aid. This recognition had just not yet led to intensive and coordinated evaluation practices, and the criteria used for evaluation had not yet been scientifically proven.³¹¹

³⁰⁸ NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memorandum 9 June 1969 no. 292/133, Chef DTH to R via S.

³⁰⁹ NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, letter 10 April 1969, Minister Udink to Minister Witteveen.

³¹⁰ NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memorandum 1 April 1969, A. van der Goot.

³¹¹ NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memorandum 17 April 1969, Plv. Chef AJV to Jan Meijer; NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memo date unknown, DTH; NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memo date unknown, author and recipient unknown.

MFA staff did not only study the international context. They actively engaged with international partners, as seen in Chapter II, by hosting the conferences in Wassenaar and Amsterdam. Van der Goot argued that it would be a waste of money to conduct an impact study of the Dutch bilateral aid programme because it would be 'disproportionately expensive', scientifically very complex, and redundant since proposals for aid evaluation were at an advanced stage in the DD-II negotiations. Van der Goot in fact concluded that impact studies should be a cooperation between recipient and donor countries. He also argued that a central system of 'data storage and retrieval' would be necessary for continuous operational evaluation. He referred to an ongoing UN study that investigated whether a computer would suffice as an international data system.³¹²

Udink informed the Tweede Kamer that international cooperation on aid evaluation was still very limited, which allegedly limited national evaluation itself. Firstly, national evaluation was often deemed incomplete as evaluators lacked information on development programmes of foreign agencies. The reviewers could not assess the effects of Dutch programmes if they were unaware of aid operations of other donors in the same region. Secondly, detailed evaluation studies were costly, and sometimes cost even more than the development programmes themselves. Thirdly, evaluation methodology was still in its infancy, particularly concerning isolated, non-commercial aid activities such as the building of a school. Experts were simply not yet able to scientifically prove what effect such a project had on regional or national development.³¹³

To solve these problems Udink actively promoted improving international cooperation in aid evaluation matters. He assigned a special role to the UN, the OECD and the World Bank in developing evaluation methods. The Dutch government would not seek an expansion of its existing evaluation procedures, but would instead strive to put evaluation on the agenda of the negotiations for the Second Development Decade.³¹⁴ In sum, Udink believed that decent evaluation of development aid could only be conducted through international cooperation: 'De echte betekenis van evaluatie – de "impact-evaluatie" – kan alleen op multilateral, op globaal niveau worden bereikt.'³¹⁵ He therefore planned to organize an international conference on evaluation methods in 1970 including participants from agencies such as the World Bank, UNDP, and other UN bodies.³¹⁶ This idea would result in the Wassenaar and later the Amsterdam conferences.

³¹² NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memorandum 1 April 1969, A. van der Goot.

³¹³ *Handelingen der Staten-Generaal*, Tweede Kamer, 1968-1969, Bijlage 9800, hfst. V, no. 62.

³¹⁴ *Ibidem*.

³¹⁵ *Handelingen der Staten-Generaal*, Tweede Kamer, 1968-1969, 3254.

³¹⁶ *Ibidem*, 3249.

The idea of hosting a conference had been proposed in a memo written by an unknown staff member earlier in 1969. The author of the memo noted that the timing of such a symposium was crucial, since the Second Development Decade would almost commence and the need for scientific methods of aid evaluation was high. The main problem in developing methods, however, was the fact that relevant academic disciplines came up with their own methods. Successful aid evaluation required a single method that would integrate methods of various academic disciplines. Bringing together all relevant expertise at the conference would contribute to the harmonisation of methods.³¹⁷

Not all MFA staff supported cooperation with multilateral organizations to the same extent as Udink. John Kaufmann, the Dutch Permanent Representative to the UN in Geneva, wrote that international aid initiatives often lacked clearly stated objectives. He therefore suggested cooperation between several medium-sized donor countries, such as the Netherlands, Sweden and Canada. These countries would be able to design evaluation methods and criteria, which the UNDP would not be capable of doing. The specialized agencies of the UN, after all, had not left Kaufmann with a very 'optimistic' outlook.³¹⁸

Nearly all parties in the Lower House joined Udink in fostering multilateral initiatives, even though some MPs did have reservations about the timing of such initiatives. VVD politician Johan Schlingemann linked the necessity of national evaluation to the probability of multilateral evaluation. He warned that it could take years before international evaluation practices would be developed. In the meantime, therefore, it seemed reasonable to him to establish national aid evaluation procedures.³¹⁹ Sef Imkamp, MP for D66, also doubted the ability of multilateral organizations to establish aid evaluation, and was not convinced by Udink's argument against a permanent evaluation unit

1) omdat ik niet zie, binnen welke termijn een internationale evaluatie van de grond kan komen, en 2) omdat ik een evaluatie, zoals die nu binnen het kader van het departement geschiedt, echt zie als een soort cuisine interne, gewoon een intern bakkeleien, wat er gebeurt en waar men nooit achter kan komen.³²⁰

³¹⁷ NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, memorandum date unknown, author and recipient unknown.

³¹⁸ 'Tenzij het Jackson-rapport en de daarop te volgen besluiten tot een nieuwe aanpak zullen leiden.' in: NL-HaNA, 2.05.313, Evaluatie-onderzoek betreffende ontwikkelingshulp, 1964-1971, inventory number 6764, telex message 28 April 1969, J. Kaufmann, Permanent Representative of the Kingdom of the Netherlands to the UN, Geneva, to Van Dam.

³¹⁹ *Handelingen der Staten-Generaal*, Tweede Kamer, 1968-1969, 3224-3226, 3261.

³²⁰ *Ibidem*, 3262-2363.

The NAR also recommended to increase cooperation with multilateral organizations, but also recognized that MFA staff would obstruct the follow-up of multilateral aid evaluations. MFA staff would not feel personally committed to multilateral evaluation, because the distance between them and the multilateral evaluators would be too large. Nevertheless, the NAR argued that the Netherlands should join the evaluation initiatives taken in the context of the Second Development Decade as much as possible.³²¹

Compared to its reluctant stance on multilateral evaluation, the NAR staunchly supported joint evaluation: 'in het algemeen is het van veel belang, dat evaluatie waar mogelijk verricht wordt door donor(s) en het ontvangende land samen.'³²² The NAR acknowledged that some recipient countries did not possess the level of expertise necessary to conduct evaluation, although others – such as India – had the capacity to do so. The NAR furthermore realized that setting up joint evaluation would be difficult. Joint evaluation would be dependent on mutual interests between donor and recipient governments, which was not always the case in previous experiences. In addition, joint evaluation could spark animosity with civil servants unwilling to comply with the necessary procedures of joint evaluation.³²³

In 1972, *De Volkskrant* published an article on aid evaluation. The editors noticed that previously, the largest share of critique on foreign aid had often come from those who were opposed to it: these critics referred to corruption and the waste of tax money. This had changed, as more and more critique came from supporters of foreign aid. They did not denounce the idea of foreign aid, but rather the way it was executed. Foreign aid delivery could be better, but how? Evaluation of aid was key, in the opinion of the editors. They mentioned the publication of the report written by Werkgroep Janssen in 1969, 'dat wel veel stof deed opwaaien, maar weinig effect heeft gehad.'³²⁴ A year later, however, one of the Werkgroep members – Jan Pronk – became the Minister for Development Cooperation. How did he deal with the recommendations?

3.3. INSTITUTIONALIZING AID EVALUATION

On the 1st of July 1977, two months after the national elections and almost four months after the Cabinet Den Uyl lost its parliamentary majority, Pronk officially created the IOV. Lodewijk Van Gorkom, Meijer's successor as DGIS between 1974 and 1980, mentioned in an interview with Van Damme and Smits in 2007 that the IOV was Pronk's own idea. Pronk's staff allegedly even opposed it.³²⁵ A year later Van Damme and Smits also interviewed Pronk. He mentioned that the

³²¹ NAR, *Advies Evaluatie van de Nederlandse Ontwikkelingshulp*, 17, 21, 41.

³²² *Ibidem*, 17.

³²³ *Ibidem*, 17.

³²⁴ *De Volkskrant*, 'Adviesraad over evaluatie-beleid. Hulp aan Derde Wereld toetsen op z'n effect', 19 September 1972.

³²⁵ L.J. van Damme, M.G.M. Smits, 'Mijn onmacht heeft met het meest teleurgesteld' in: Van Damme, Smits (eds.), *Voor de ontwikkeling van de derde wereld*, 163.

idea for the IOV arose from his personal judgment that foreign aid would receive intense public criticism in the years to come. After decades of public support for aid, Pronk felt that the tide was turning. Aid expenditures would increasingly be subject to public scrutiny, and the MFA would be held accountable for the efficiency and effectiveness of its policies.³²⁶ Van Beurden and Gewald briefly answer the question why the IOV was erected: '[t]o prevent (...) public relations disasters occurring in the future Pronk believed he needed to [be] better informed of failures before his critics got to hear of them.'³²⁷ But the focus of their book is much more on the work of the IOV than the history of its establishment. So, what do the sources tell about why Pronk created the IOV?

3.3.1. Purpose

An analysis of the memos from the MFA offers a more diverse view. First of all, it was up to internal discussion whether the establishment of the IOV was able to counter public criticism. A.W.F. Roos, staff member of DTH/BL³²⁸, sent a detailed memo to Pronk as a response to the latter's request to prepare the establishment of the IOV. In that memo, Roos argued that the IOV could be useful

voor het wegnemen van gevoelens van wantrouwen bij het publiek over de doelmatigheid van de besteding van middelen uit het plafond. Het blote bestaan van een inspectie kan wat dit betreft reeds als een verzekering gelden. Bij de instelling van een inspectie zou aan dit aspect anderzijds ook aandacht moeten worden besteed, al was het alleen maar om te voorkomen dat de instelling ervan zou kunnen worden opgevat als een bevestiging dat eventuele vermoedens over ondoelmatige besteding van middelen gegrond waren. Het instellen van een inspectie is derhalve een politieke daad, en zal als zodanig ook worden beoordeeld, mede tegen de achtergrond van het te zijner tijd verschijnende rapport over het thans lopende onderzoek van de Algemene Rekenkamer.³²⁹

Some three weeks later, the Head of DTH/UH³³⁰ wrote to the Head of BL, the superior of Roos:

Het is maar de vraag of het beschikken over inspectie in het eigen apparaat er inderdaad toe zal bijdragen om een vermeend wantrouwen bij het publiek ten aanzien van doelmatig

³²⁶ L.J. van Damme, M.G.M. Smits, 'Ik heb getracht er bovenop te zitten' in: Van Damme, Smits (eds.), *Voor de ontwikkeling van de derde wereld*, 183.

³²⁷ Van Beurden, Gewald, *From output to outcome*, 20.

³²⁸ *Bureau Beleidszaken, Directie Internationale Technische Hulp*.

³²⁹ *Nederlandse ontwikkelingssamenwerking IV*, 597.

³³⁰ *Afdeling Uitvoering Technische Hulpverlening, Directie Internationale Technische Hulp*.

aanwenden van O.S.³³¹-fondsen weg te nemen. Ik geloof hier niet zo erg in. Het Nederlandse volk raakt niet zo erg onder de indruk van inspectie-resultaten in eigen huis (men hangt niet zo gauw de vuile was buiten de deur is dan de algemene opinie), wel van een extern onafhankelijk oordeel zoals van de Algemene Rekenkamer.³³²

Deputy Director of DTH Peters also questioned whether the IOV was able to remove distrust. He distinguished between external and internal objectives. If the IOV would exist for external relations, then it would produce studies for the Minister to present to Parliament, the media or at international conferences. He imagined that the establishment of the IOV would give leeway to those who desired more parliamentary grip on aid policy – in other words, the external option would enhance the Minister’s accountability vis-à-vis parliament. The second option, on the other hand, would leave the IOV as an internal policy tool only to inform the Minister and his senior staff whether aid delivers the desired results.³³³

Pronk himself saw inspection as a necessity for any mature government policy. In his memo to DGIS on 20 September 1976, in which he requested the latter to prepare the establishment of the IOV, he argued in favour of inspection because of the ‘considerable size of the aid programme.’³³⁴ The exact meaning of the sentence is hard to define, but it is clear that he now regarded bilateral development policy as mature government policy. He also explicitly referred to the existence of inspectorates at the Ministries of Education, CRM and Finance.³³⁵ It suggests that he specifically aimed for a ‘common’ policy review unit, instead of something new and unique in public administration. Peters confirmed this in a memo from December 1976, in which he wrote that state and size of bilateral aid required an independent inspection, similar to existing practices in other ministries.³³⁶ Nonetheless, it is still difficult to determine whether such a policy review unit would serve solely the purpose of lesson-learning or the purpose of accountability. Pronk desired the IOV especially because he needed independent research to check whether aid reaches its pre-stated objectives.³³⁷

³³¹ *Ontwikkelingssamenwerking*.

³³² NL-HaNA, 2.05.330, Instelling van de Inspectie Ontwikkelingssamenwerking te Velde (IOV), vaststelling van de formatie, taken en bevoegdheden, evaluatie van de uitgevoerde werkzaamheden en onderzoeken door de Algemene Rekenkamer naar de bijdrage die het werk van de inspectie levert aan de kwaliteitsverbetering van de bilaterale ontwikkelingssamenwerking, inventory number 27328, memo Hoofd UH to Hoofd BL, 30 november 1976.

³³³ *Nederlandse ontwikkelingssamenwerking IV*, 615-616.

³³⁴ ‘de aanzienlijke omvang van de begrotingsgelden, welke voor Ontwikkelingssamenwerking zijn uitgetrokken.’ In: *Nederlandse ontwikkelingssamenwerking IV*, 581.

³³⁵ *Nederlandse ontwikkelingssamenwerking IV*, 581.

³³⁶ ‘de huidige aard en omvang van de bilaterale ontwikkelingssamenwerking de behoefte aan een onafhankelijke inspectie oproept, naar analogie van hetgeen terzake bij andere ministeries bestaat’. Ibidem, 614.

³³⁷ ‘vooral in het onafhankelijk onderzoek of de besteding van hulp gelden in overeenstemming is met het beoogde doel.’ Ibidem, 581.

Thus, Pronk saw the IOV as a tool of control. He argued in the margin of a memo from Roos, on the 28th of February 1977, that the main responsibility of the IOV was

de controle op de besteding van ontwikkelingsgelden middels het onderzoeken van de kwaliteit der in uitvoering zijnde activiteiten.³³⁸

Pronk called this the IOV's 'unique responsibility'. Another responsibility would be to 'make recommendations for adaptation of projects and programmes to enhance their efficiency and/or effectiveness.'³³⁹ A memo from DTH/JR³⁴⁰ in 1978, after the establishment of the IOV, contained the same message:

Inspectie is een beleidsinstrument ter verhoging van de kwaliteit van de Nederlandse hulp. De instelling van IOV betekent overigens niet dat een nieuwe verplichting op R³⁴¹ of op de Staat is komen te rusten: steeds aanwezig is en was de plicht de beperkte overheidsgelden zo goed mogelijk te besteden.³⁴²

In addition to this aspect of lesson-learning, Roos explained that the IOV would at least serve the accountability of the Minister vis-à-vis parliament:

Uiteindelijk is en blijft de Minister voor Ontwikkelingssamenwerking de enige (politiek) verantwoordelijke voor het totale beleid. Hij dient daarom in staat te zijn een oordeel te vormen over de beleidsuitvoering door de diverse directies. Een Inspectie Ontwikkelingssamenwerking kan daarbij functioneren als orgaan om de Minister in staat te stellen zelfstandig en onafhankelijk van de betrokken directies, de praktijk van de beleidsuitvoering te beoordelen.³⁴³

³³⁸ *Nederlandse ontwikkelingssamenwerking IV*, 665n.

³³⁹ 'het doen van aanbevelingen t.a.v. wenselijke bijsturing van lopende projecten en/of programma's die kunnen leiden tot een vergroting van de efficiency en/of effectiviteit'. In: *Ibidem*, 665.

³⁴⁰ *Bureau Juridische Zaken, Directie Internationale Technische Hulp*.

³⁴¹ Former Telex codes are still used to refer to the Ministers at the MFA. 'R' is used for the Minister for Development Cooperation.

³⁴² NL-HaNA, 2.05.330, Instelling van de Inspectie Ontwikkelingssamenwerking te Velde (IOV), vaststelling van de formatie, taken en bevoegdheden, evaluatie van de uitgevoerde werkzaamheden en onderzoeken door de Algemene Rekenkamer naar de bijdrage die het werk van de inspectie levert aan de kwaliteitsverbetering van de bilaterale ontwikkelingssamenwerking, inventory number 27328, memo DTH/JR to IOV, inspectieclausule in overeenkomsten, 7 februari 1978.

³⁴³ *Nederlandse ontwikkelingssamenwerking IV*, 667.

In the margin, Pronk responded to this particular remark with 'Inderdaad'.³⁴⁴ DTH/JR reiterated the role of the IOV in the Minister's accountability towards parliament in 1978:

De aard, intensiteit van de inspectie zal variëren met de omvang van de verantwoordelijkheid van Nederland cq Ontwikkelingssamenwerking. Met verantwoordelijkheid wordt hier bedoeld zowel een intern-politieke verantwoordelijkheid van R (en M) t.o.v. de Staten-Generaal, als een juridische verantwoordelijkheid naar buiten (ontwikkelingsland, samenwerkende organisatie).³⁴⁵

These sources reveal that Pronk instituted the IOV primarily to increase his awareness of the successes and failures of his policies. This awareness would enable him to better inform or respond to parliament when he needed to. The value for lesson-learning accredited to evaluation should not be forgotten, however. In fact, we see that civil servants regarded aid evaluation predominantly in the light of increasing feedback and learning lessons. This especially applies for the other type of evaluation at the MFA, the built-in evaluation procedures. The first time we see this, is in a memo written between 1969 and 1974, of which the exact date, author and recipient are unfortunately unknown. It is possibly written by DTH as a review of its own policies including several recommendations for the future. The author acknowledged that Dutch technical assistance had not yet developed a sufficient feedback mechanism between the project experts on the ground and the MFA. Consequently, the effectiveness of many projects turned out lower than it could potentially be. The author therefore recommended 'on the spot' evaluation, conducted in harmony by Dutch project experts, DTH staff and their local counterparts in the project area.³⁴⁶

In 1974, a new policy of built-in evaluation was established. This policy was primarily aimed at improving the quality of project implementation.³⁴⁷ Two years later, however, the Head of DTH/UH summed up how that same built-in evaluation had not led to significant conclusions:

³⁴⁴ *Nederlandse ontwikkelingssamenwerking IV*, 667.

³⁴⁵ NL-HaNA, 2.05.330, Instelling van de Inspectie Ontwikkelingssamenwerking te Velde (IOV), vaststelling van de formatie, taken en bevoegdheden, evaluatie van de uitgevoerde werkzaamheden en onderzoeken door de Algemene Rekenkamer naar de bijdrage die het werk van de inspectie levert aan de kwaliteitsverbetering van de bilaterale ontwikkelingssamenwerking, inventory number 27328, memo DTH/JR to IOV, inspectieclausule in overeenkomsten, 7 februari 1978.

³⁴⁶ NL-HaNA, 2.05.313, Evaluatie van de Nederlandse ontwikkelingshulp, inventory number 27542, note date unknown, 'De uitvoering van Nederlandse technische hulpprogramma's in ontwikkelingslanden', author and recipient unknown.

³⁴⁷ 'in eerste instantie gericht op kwaliteitsverbetering van de projektrealisatie.' In: *Nederlandse ontwikkelingssamenwerking IV*, 209.

Al die onderdelen en fasen van de voorbereiding en uitvoering van projecten in het veld (...), worden veelal van verre gevolgd zonder dat er een duidelijk inzicht is in de gang van zaken. Nog weinig gestructureerd en vaak incidenteel komen voortgangsgegevens via rapportages, verslagen van projectbezoeken, gesprekken met projectdeskundigen en ambassades op de DTH-bureaux terecht. Vanaf 1974 probeert de ingebouwde projectevaluatie hier meer structuur en continuïteit in aan te brengen. (...) Harde noten, althans tot dusver, worden niet gekraakt. Het is niet ongebruikelijk dat dergelijke onderzoeken uitmonden in een bevestiging van reeds veronderstelde standpunten. Een inspectie-eenheid is naar mijn mening beter in staat de zaken waar het omgaat duidelijker te stellen.³⁴⁸

A cause for the insufficient execution of evaluation was thought to be the lack of objectivity among staff members of the operational departments. This lack of feedback was not only recognized in the context of built-in evaluation. In a circular letter announcing the IOV, the MFA's Secretary-General argued that:

In het Nederlandse overheidsapparaat ondervindt de terugkoppeling van veldervaring naar de operationele beleidscentra en verder naar de algemene beleidsvorming moeilijkheden. Het gevolg is, enerzijds, een minder dan optimaal inzicht bij de diverse beleidsinstanties in de mogelijkheden en problemen te velde en, anderzijds, frustratie in het veld. Verwacht mag worden, dat de uitvoering van inspectiewerkzaamheden tevens zal leiden tot verbeteringen in dit vlak.³⁴⁹

Of course, it is impossible to exactly trace how many MFA staff members preferred lesson-learning over accountability. Only oral history research might enable the historian to reveal an answer to that question, as we cannot reconstruct the atmosphere at the MFA from just the preserved written source material. Similarly, it is difficult to trace whether actors outside the MFA engaged in matters of aid evaluation. The high amount of attention for aid evaluation in 1969 did not resurface in the years of Pronk, not even when he established the IOV. As an exception, Labour

³⁴⁸ NL-HaNA, 2.05.330, Instelling van de Inspectie Ontwikkelingssamenwerking te Velde (IOV), vaststelling van de formatie, taken en bevoegdheden, evaluatie van de uitgevoerde werkzaamheden en onderzoeken door de Algemene Rekenkamer naar de bijdrage die het werk van de inspectie levert aan de kwaliteitsverbetering van de bilaterale ontwikkelingssamenwerking, inventory number 27328, memo Hoofd UH to Hoofd BL, 30 november 1976.

³⁴⁹ NL-HaNA, 2.05.330, Instelling van de Inspectie Ontwikkelingssamenwerking te Velde (IOV), vaststelling van de formatie, taken en bevoegdheden, evaluatie van de uitgevoerde werkzaamheden en onderzoeken door de Algemene Rekenkamer naar de bijdrage die het werk van de inspectie levert aan de kwaliteitsverbetering van de bilaterale ontwikkelingssamenwerking, inventory number 27328, circular letter 15 August 1977, Secretaris-Generaal to all missions and departments.

Party Senator Suzanne Steigenga-Kouwe demanded full disclosure of evaluation reports to Parliament early 1978. She also doubted whether Dutch diplomats possessed the necessary skill to contribute to evaluations, and thereby implied that the MFA should increase its budgetary capacity reserved for evaluation. There is no doubt that Steigenga-Kouwe desired parliamentary control and thus some form of accountability on the part of the Minister.³⁵⁰ A second exception was an article by J. van der Deijs, who argued for a more scientific approach to aid in NRC Handelsblad, published just months after the start of Pronk's tenure in 1973:

In de praktijk ontbreekt veelal een geïntegreerde wetenschappelijke begeleiding van het Nederlandse hulpverleningsbeleid, waarin plaats is voor analyse vooraf voor een bepaald project en de effecten daarvan ook op andere terreinen, alsook een evaluatie na afloop.³⁵¹

This opinion corresponds to the thought of Udink and the Werkgroep Janssen some years earlier. The Dutch aid policy lacked scientific footing. Nonetheless, the sources analysed here do not reveal that the IOV was destined to give Pronk this scientific footing. It was merely an instrument that would inform him personally as soon as possible whether his aid reached its pre-stated objectives.

3.3.2. Aid evaluation in national governance

During Pronk's tenure, an important distinction was made between evaluation and inspection. Evaluation, the built-in evaluation conducted by operational departments DTH and DFO, was as a tool within an aid project's life cycle.³⁵² The researchers of the IOV, on the other hand, conducted inspection and stood outside the operational departments of the MFA:

IOV zal haar taken moeten kunnen uitoefenen onafhankelijk, in alle rust en vrij van door IOV ongewenste invloeden. Daar is voor nodig, dat IOV de projektplaats(en) kan bezoeken, kan spreken met het door Nederland ingezette personeel, het door Nederland geleverde materiaal kan zien, en voorts alle informatie met betrekking tot het projekt kan krijgen.³⁵³

³⁵⁰ *Handelingen der Staten-Generaal*, Eerste Kamer 1977-1978, 9 mei 1978.

³⁵¹ NRC Handelsblad, 'Ander wetenschapsorganisatie nodig voor ander beleid van Pronk', 9 November 1973.

³⁵² *Nederlandse ontwikkelingssamenwerking IV*, 206-219.

³⁵³ NL-HaNA, 2.05.330, Instelling van de Inspectie Ontwikkelingssamenwerking te Velde (IOV), vaststelling van de formatie, taken en bevoegdheden, evaluatie van de uitgevoerde werkzaamheden en onderzoeken door de Algemene Rekenkamer naar de bijdrage die het werk van de inspectie levert aan de kwaliteitsverbetering van de bilaterale ontwikkelingssamenwerking, inventory number 27328, memo DTH/JR to IOV, inspectieclausule in overeenkomsten, 7 februari 1978.

The Secretary-General of the MFA reiterated the independence of the IOV and granted its researchers unhindered access to all data and correspondence of the operational departments.³⁵⁴ The Minister also allowed the IOV to request assistance from experts outside the MFA.³⁵⁵ In the circular letter announcing its establishment, the IOV claimed to work differently than the existing built-in evaluation in four ways. Firstly, the inspection concerned the use of resources in particular, while built-in evaluation also concerned the assessment of objectives per se. Secondly, inspection focused solely on Dutch activities, evaluation focused on the interplay between Dutch and other activities, including its coherence with policies of multilateral institutions, other donor governments or recipient countries. Thirdly, the inspection of the IOV was a unilateral exercise, while built-in evaluation always occurred in cooperation with local counterparts. Fourthly, inspection concluded with recommendations, evaluation with joint decisions.

Anders gezegd: evaluatie is een operationeel werktuig voor samenwerking met autoriteiten en instanties in ontwikkelingslanden, inspectie is een intern Nederlands werktuig dat specifiek staat ingesteld op de kwaliteit van het eigen aandeel in de hulp.³⁵⁶

Despite all efforts of the MFA to institutionalize both evaluation and inspection, staff members still experienced difficulties in putting ideas into practice. DTH/BL wrote in 1974 that the Werkgroep Janssen was unable to accomplish its task – to assess whether a certain type of aid delivered better than the other – because of the complexity of the development process, insufficient scientific methods and the lack of necessary data. DTH/BL argued that this problem had not changed. Evaluation of programme aid was still not viable.³⁵⁷ Two years later, Pronk admitted in his policy note that, from a scientific perspective, ex-post evaluation was still difficult. Therefore, he pointed out that his initiatives focused on built-in project evaluation at first. Some

³⁵⁴ NL-HaNA, 2.05.330, Instelling van de Inspectie Ontwikkelingssamenwerking te Velde (IOV), vaststelling van de formatie, taken en bevoegdheden, evaluatie van de uitgevoerde werkzaamheden en onderzoeken door de Algemene Rekenkamer naar de bijdrage die het werk van de inspectie levert aan de kwaliteitsverbetering van de bilaterale ontwikkelingssamenwerking, inventory number 27328, memo 15 August 1977, Secretaris-Generaal to all missions and departments.

³⁵⁵ NL-HaNA, 2.05.330, Instelling van de Inspectie Ontwikkelingssamenwerking te Velde (IOV), vaststelling van de formatie, taken en bevoegdheden, evaluatie van de uitgevoerde werkzaamheden en onderzoeken door de Algemene Rekenkamer naar de bijdrage die het werk van de inspectie levert aan de kwaliteitsverbetering van de bilaterale ontwikkelingssamenwerking, inventory number 27328, memo 15 August 1977, Secretaris-Generaal to all missions and departments.

³⁵⁶ NL-HaNA, 2.05.330, Instelling van de Inspectie Ontwikkelingssamenwerking te Velde (IOV), vaststelling van de formatie, taken en bevoegdheden, evaluatie van de uitgevoerde werkzaamheden en onderzoeken door de Algemene Rekenkamer naar de bijdrage die het werk van de inspectie levert aan de kwaliteitsverbetering van de bilaterale ontwikkelingssamenwerking, inventory number 27328, circular letter 15 August 1977, Secretaris-Generaal to all missions and departments.

³⁵⁷ *Nederlandse ontwikkelingssamenwerking IV*, 213-214.

commentators outside the MFA agreed to this, as we can read in De Volkskrant late 1977.³⁵⁸ In sum, it might be that the complexity of aid evaluation delayed the implementation of the IOV until Pronk's last months as the Minister for Development Cooperation.

3.3.3. International aid evaluation

The previous paragraph already mentioned that Dutch built-in evaluation always occurred with partners of the respective aid project. Although the sources of Chapter II only reach until 1975, the Dutch sources analysed here reveal that aid evaluation strongly remained an international character. Pronk recognized the increasing international popularity of aid evaluation in his *Nota Bilaterale Ontwikkelingssamenwerking* of 1976.³⁵⁹ But already from the beginning of his tenure, Pronk expressed an interest to cooperate internationally. He outlined during his first weekly meeting with DGIS in 1973 that the MFA should find 'connection' with evaluation procedures of the UN.³⁶⁰

DTH/BL acknowledged in 1974 that the new policy of project evaluation was based on the built-in evaluation used by UNDP and USAID, the Logical Framework Approach (LFA). According to DTH/BL, it was essential to monitor the experiences with project evaluation of other donors and recipient countries.³⁶¹ In the same year, DTH/BL followed up on this through a letter to the Dutch Ambassador in Washington. The Head of DTH/BL wanted to obtain information concerning the financial and staff requirements of evaluation within USAID. He noted that he was already working on a concrete proposal for the evaluation of technical assistance projects. This letter clearly shows that the MFA decided to implement some form of evaluation at least in 1974.³⁶² More information about initiatives across the border came in 1976 and 1977, when Roos had several bilateral meetings in Ottawa, London, New York, Washington, Bonn and Rome with MPs, representatives of the World Bank and the UN and staff of foreign MFAs.³⁶³ He gathered all information in a study presented to Pronk. And eventually, after the establishment of the IOV, the Dutch Permanent Mission to the OECD requested English and French copies of the IOV's founding

³⁵⁸ 'Het [ontwikkelingsbudget] is zo'n tweehonderd gulden per man per jaar. Een goed idee dus om eens rustig uit te leggen wat daarmee wordt gedaan. Gemakkelijk is dat niet, want dan zou je moeten kunnen nagaan of de hulp inderdaad de mensen aan de onderkant bereikt en of de juiste projecten zijn gekozen. Dat is niet gemakkelijk na te gaan. Evaluatie van ontwikkelingshulp is moeilijk. Pronk heeft nu inspecties te velde ingesteld.' De Volkskrant, 'Pronk legt uit waar zijn miljoenen heengaan', 26 October 1977.

³⁵⁹ 'Nota bilaterale ontwikkelingssamenwerking. Om de kwaliteit van de Nederlandse hulp', *Handelingen der Staten-Generaal*, Tweede Kamer, 1976-1977.

³⁶⁰ Pronk allegedly used the word 'aansluiting'. In: *Nederlandse ontwikkelingssamenwerking IV*, 1.

³⁶¹ 'Voorts is het voor het uitwerken van het systeem van projekt-evaluatie van belang de ervaring van andere donororganisaties, van de ontwikkelingslanden en van DTH zelf in de evaluatie van projecten nauwgezet te volgen.' In: *Nederlandse ontwikkelingssamenwerking IV*, 209.

³⁶² NL-HaNA, 2.05.313, Evaluatie van de Nederlandse ontwikkelingshulp, inventory number 27542, letter 5 August 1974, Head of Bureau Beleidszaken to Dutch Ambassador in Washington.

³⁶³ NL-HaNA, 2.05.330, Instellen van een Inspectie Ontwikkelingsbeleid 1976-1977, inventory number 32974.

documents, because the inspection was an issue ‘that would strongly arouse the interest of DAC members.’³⁶⁴

However, while Udink and the Tweede Kamer agreed on fostering evaluation through multilateral channels in 1969, the accessible primary sources show that this option was never a topic of discussion during Pronk’s tenure. While Pronk emphasized the need for international standardization of methods and procedures of aid evaluation and ‘connection’ to the evaluation systems of the UNDP and USAID in his policy, he did not mention the possibility of evaluation of bilateral activities by multilateral agencies.³⁶⁵ It would also become clear that the IOV would not cooperate with multilateral evaluations or even focus on the Dutch contribution to multilateral development activities: ‘Buiten haar actieradius liggen die hulpvormen waarbij de bevoegdheid van de Nederlandse overheid opgaat in de autoriteit van multilaterale lichamen (...). In het veldwerk van deze lichamen valt geen afzonderlijke inbreng van de Nederlandse regering te identificeren.’³⁶⁶

Interestingly, Pronk explicitly disliked joint evaluation.³⁶⁷ This is reported by F. Koopman, staff member of DTH/BL, in the minutes of a meeting held on in May 1974. Koopman noted that Pronk and staff members disagreed on this matter. The Minister argued that ex-post evaluation would be difficult in the case of projects where Dutch objectives did not match the objectives of recipient countries. Central authorities would not be willing to cooperate. He mentioned this in his policy note two years later, arguing that recipient countries were generally less interested in ex-post evaluation than donor governments were.³⁶⁸ But Pronk and staff eventually agreed on joint evaluation after staff members successfully convinced him ‘[v]an diverse kanten’.³⁶⁹

In his 1976 policy note, Pronk did highlight a positive aspect to joint evaluation. It could serve as a form of aid in itself:

³⁶⁴ ‘dat dac-leden zeer zal interesseren.’ In: NL-HaNA, 2.05.330, Instelling van de Inspectie Ontwikkelingssamenwerking te Velde (IOV), vaststelling van de formatie, taken en bevoegdheden, evaluatie van de uitgevoerde werkzaamheden en onderzoeken door de Algemene Rekenkamer naar de bijdrage die het werk van de inspectie levert aan de kwaliteitsverbetering van de bilaterale ontwikkelingsamenwerking, inventory number 27328, telex message 25 August 1977, Permanent Mission to OECD to MFA.

³⁶⁵ ‘Nota bilaterale ontwikkelingssamenwerking. Om de kwaliteit van de Nederlandse hulp’, *Handelingen der Staten-Generaal*, Tweede Kamer, 1976-1977.

³⁶⁶ NL-HaNA, 2.05.330, Instelling van de Inspectie Ontwikkelingssamenwerking te Velde (IOV), vaststelling van de formatie, taken en bevoegdheden, evaluatie van de uitgevoerde werkzaamheden en onderzoeken door de Algemene Rekenkamer naar de bijdrage die het werk van de inspectie levert aan de kwaliteitsverbetering van de bilaterale ontwikkelingsamenwerking, inventory number 27328, circular letter 15 August 1977, Secretaris-Generaal to all missions and departments.

³⁶⁷ ‘Bovendien achtte hij joint-evaluation niet noodzakelijk.’ In: *Nederlandse ontwikkelingssamenwerking IV*, 229.

³⁶⁸ ‘Nota bilaterale ontwikkelingssamenwerking. Om de kwaliteit van de Nederlandse hulp’, *Handelingen der Staten-Generaal*, Tweede Kamer, 1976-1977, 74.

³⁶⁹ *Nederlandse ontwikkelingssamenwerking IV*, 228-230.

[h]et gezamenlijk met de betrokken lokale instanties kritisch analyseren van de planning en uitvoering van gerealiseerde samenwerkingsactiviteiten ondersteunt het proces dat erop gericht is de ontwikkelingslanden in staat te stellen zelf de planning en het management van de eigen ontwikkelingsactiviteiten uit te voeren.³⁷⁰

Joint evaluation, then, fit perfectly in Pronk's policy objective of aid as a means to reach self-reliance of recipient countries.³⁷¹ Civil servants also discussed possibilities to train recipient countries in conducting evaluation. According to a 1974 memo from DTH/BL, recipient countries were already expanding their capacity in aid evaluation to ensure their self-reliance and independence vis-à-vis donors. DTH/BL therefore advised Pronk to focus on joint evaluation and invite recipient countries to its existing evaluation activities.³⁷²

CONCLUSION

Despite the establishment of the IOV in 1977, commentators still wondered whether decent aid evaluation was in fact possible. Newspaper Trouw was a fierce sceptic and wrote in 1978:

Op de vraag of de buitenlandse – en met name de Nederlandse – hulp al dan niet sociaal nut afwerpt is geen antwoord te geven. Alles is zo langzamerhand door de wetenschap onderzocht, maar wat betreft de effectiviteit van de ontwikkelingshulp blijft het vaak gissen.³⁷³

Trouw criticized the Dutch government, both Pronk and his successor De Koning, for dealing too little with aid evaluation. The newspaper noted that just one all-encompassing study of Dutch aid had been conducted, referring to a study in Colombia by the Free University of Amsterdam. Trouw recommended more of such studies and emphasized the critical and objective role universities could play. Despite this article, one must conclude that newspapers displayed little interest in aid evaluation.

This observation helps in answering the first of three research questions posed in this chapter, concerning the purpose of aid evaluation. Seeing the lack of interest from the newspapers and parliament, most notably the near absence of aid evaluation in the debates of the Eerste Kamer, the idea for aid evaluation came primarily from within the MFA. It can be argued that the idea of evaluation arose inside the MFA to defend its aid expenditures. Accountability did not

³⁷⁰ 'Nota bilaterale ontwikkelingssamenwerking. Om de kwaliteit van de Nederlandse hulp', *Handelingen der Staten-Generaal*, Tweede Kamer, 1976-1977.

³⁷¹ *Nederlandse ontwikkelingssamenwerking IV*, 206-209.

³⁷² *Ibidem*, 206-207.

³⁷³ Trouw, 'Nederland en Indonesië (1). "Je moet van je schuldgevoel af"', 30 September 1978.

originate in a parliamentary desire to control. Lesson-learning resurfaced during the debates at the MFA, but it is difficult to discern a pattern which of the two purposes were deemed more important at what point in time.

Concerning the domestic institutionalization of aid evaluation, one of the reasons why it took years to form the IOV was financial. Pronk noted in a memo to DGIS that there was little room for additional personnel in the ranks of the MFA.³⁷⁴ Another reason was criticism from DTH. Deputy Director DTH Peters argued in a memo that the establishment of the IOV would demote DTH and DFO from policy departments to departments only responsible for project execution.³⁷⁵

Lastly, it might seem surprising that the MFA's focus shifted from multilateral evaluation to national evaluation. As seen in the memos of the MFA, the discussion on evaluation left out multilateral initiatives completely. Nonetheless, evaluation by multilateral institutions progressed enormously in the 1970s. The World Bank created its OED in 1973 and the UN agencies experimented too. It might have been the case, therefore, that Dutch civil servants paid little attention to multilateral evaluation in their memos as this process had already taken off strongly.

³⁷⁴ *Nederlandse ontwikkelingssamenwerking IV*, 582.

³⁷⁵ *Ibidem*, 612-618.

CONCLUSION

This thesis set out by asking what factors have determined the institutionalization of aid evaluation in the 1960s and 1970s. Evaluation of public policy took off in the 1960s in a number of western countries, the US in particular, to improve the efficiency and effectiveness of welfare state programmes. To some extent, this practice spilled over to US aid policy and (aid) policies of other donor governments and multilateral organizations. In the 1970s evaluation of public policy really expanded, primarily because budget cuts pushed governments to fund only the most efficient and effective programmes. This thesis has aimed to provide more insight in this process, focusing particularly on the development of aid evaluation in the Netherlands until 1977. It has provided answers to the question how actors in the development sector – civil servants of MFAs and multilateral organizations, politicians, academics and newspapers – discussed aid evaluation: what purpose did aid evaluation have, who should conduct evaluation, and how could international cooperation improve the quality of aid evaluations? Here, I will single out the six main findings.

The thesis firstly demonstrates that in a context of input-based management where the main priority of many governments was to reach the 0,7% ODA mark – input-based aid management – the focus was gradually shifting towards outputs or, more broadly, results. This is not to say that the MFA governed aid policy through results-based management, the type of public administration linked to aid policies of the 21st century. However, based on the analysis of Pronk's motives for establishing the IOV, one could argue that he used evaluation – or inspection – as an instrument to increase awareness of his policy's successes. Without a doubt he wanted the IOV to scientifically substantiate the policy decisions he made. His mere insistence on the need for such a check already provides some evidence of an archetype of results-based management.

Secondly, although it was Pronk who established the IOV, his predecessors laid significant groundwork for the institutionalization of aid evaluation during the 1970s. To some extent, this groundwork of the 1960s and early 1970s was a fine example of policy intimacy, the concept of former Dutch Minister Donner and coined in the introduction to this thesis. Behind the scenes, Jan Meijer and Ferdinand van Dam saw the need for policy review and consulted with MFA staff members and many academics on the feasibility of evaluation. Parliament played a very tiny role in this process. Even more, it was an example of *transnational* policy intimacy. The Netherlands played an important role in convening experts and representatives from national aid agencies, multilateral organizations and academia to share best practices, such as the American Logical Framework Approach.

Thirdly, accountability was much more a priority in multilateral evaluation than in Dutch national evaluation. This had to do with the strong supervisory role of member states in

multilateral organizations on the one hand, and the lack of supervision by the Dutch Lower House on national aid governance. Donor governments were particularly interested in knowing how effective and efficient their contributions to multilateral activities were. Furthermore, the purpose of public relations is also a curious one. Meijer voiced the need for evaluation as a tool for public relations during the Berlin conference. It is curious since we know that Meijer had quite some influence on the DD-II proposals through his close connection to Tinbergen, the chair of the CDP. The DD-II proposals encouraged UN member states to develop policies to raise awareness of development aid among the general public. Of course, we cannot conclude on the basis of the Berlin Conference report that this idea came from Meijer. Further research could illuminate whether this was actually the case.

Fourthly, the domestic institutionalization of aid evaluation was a budgetary and administrative question. Concerning the latter, many European aid agencies had little money to spend on the actual aid programmes and projects, which made it even harder to leave spare financial resources for hiring evaluators. According to Arnold, budget constraints played a crucial role in the development of aid policy in the Netherlands.³⁷⁶ This thesis shows that it also had a major impact in establishing evaluation practices. Moreover, it is important to note that MFA staff members generally disliked the idea of evaluation, especially conducted by evaluators outside their own departments. But built-in evaluation was also in some disdain: the obligation to evaluate would increase their workload in project management, in a time when the number of projects grew explosively compared to the growth of MFA personnel.

Fifthly, this thesis has demonstrated that academics, politicians, policymakers and other members of the expert community not just disagreed on certain matters of development. The thesis shows in fact that many actors still debated how they should actually make development work. Some of the experts in the development sectors of the 1960s were aware of the shortcomings of knowledge, as the quote of Ferdinand van Dam at the beginning of Chapter I reveals. A decade later, Pronk argued that development policy had become mature and that mature government policy requires inspection. The IOV provided him the tools to control the execution and success his policy. However, we learn from the sources that aid evaluation itself was barely mature at the time Pronk left the MFA in 1977. MFA staff members even proclaimed more than once that methods of aid evaluation were not yet sufficient. This conclusion opposes postdevelopment, but also more general IR literature that perceives development thought of the 1960s as coherent and hegemonial. For instance, Duco Hellema analysed the 1966 note of Minister Udink, which allegedly 'still assumed a great trust in the operation of the capitalist world economy: recipient countries must above all be taught to adapt themselves to this operation of

³⁷⁶ Arnold, *Implementing Development Assistance*, 176, 179.

the world economy.³⁷⁷ Instead, the sources on aid evaluation show that actors continuously debated and questioned existing development ideas.

Sixthly, we can further remark that 1969 was an important year for aid evaluation in the both in the international and the Dutch aid community. Internationally, we saw the publication of the Jackson Study, the Pearson Report and the Tinbergen Report. In the Netherlands, the Werkgroep Janssen published its report which led to significant debate in parliament. These publications and their attention to aid evaluation to the wider crisis of development aid. Although they amount to only a portion of the entire body of documents published in that year, they represent the emerging thought that evaluation should be a vital part of every aid apparatus. Within the crisis of development aid, the studies presented evaluation as a necessary instrument to improve policies *and* accountability vis-à-vis parliaments and the general public.

Despite these findings, this thesis brings up many other questions that have not been answered here. It has shown that aid evaluation as a topic of historical research offers a fantastic insight in broader discussions on foreign aid. One is able to get a good grasp of the state of the art in development aid in a certain period of time just from reading the discussions on aid evaluation. Let me single out three suggestions for further research. Firstly, the scope of the thesis was not to examine the existence and working of a feedback loop within the Dutch MFA or in the global development sector in general. But the source material used for this thesis contains a lot of information that might actually give an idea of such a mechanism. The archives of the MFA includes many project evaluations conducted during the 1970s, and even more during the 1980s. In addition, research on the feedback loop through analyses of the IOV studies would be valuable too.

Secondly, this thesis lacks in-depth analyses of experiments with and institutionalization of aid evaluation within a number of key actors. Chapter II demonstrated that multilateral organizations, through the studies of Jackson, the CDP and the Pearson Committee, as well as through participation at the conferences were influential in steering the debate. Moreover, historian Martin Rempe has conducted some on EEC evaluation initiatives. But still, we know very little about the policy intimacy of multilateral organizations. In addition, the thesis has focused on aid evaluation only in western public administrations. But they were not the only ones to engage in aid evaluation. Stokke noted that some recipient countries had established evaluation practices early on, in fact earlier than most developed countries. He referred to India, for example, where the government founded the Programme Evaluation Organisation in 1952.³⁷⁸ Some

³⁷⁷ D.A. Hellema, *Dutch Foreign Policy: The Role of the Netherlands in World Politics* (Dordrecht 2009) 221.

³⁷⁸ Stokke, 'Policies, Performance, Trends and Challenges in Aid Evaluation', 5.

evaluations of projects in India have also been published in the late 1950 and 1960s.³⁷⁹ Laporte also recalls that the Global South lacked the necessary capacity to conduct evaluations. Therefore, the UN focused on strengthening such institutions.³⁸⁰ Lastly, this thesis has not studied evaluation practices in NGOs. At a conference in Zandvoort in 1974 Pronk agreed with CEBEMO, ICCO and NOVIB that the MFP would be evaluated by randomized samples. The implementation of this agreement was prepared by a joint working group of the MFA and the three NGOs. The archives of this joint working group have only recently become publicly accessible and might offer valuable information about the further institutionalization of aid evaluation.³⁸¹ This thesis was limited roughly to the archival material on aid evaluation at the Dutch MFA, while a closer look at the archives of other actors could provide historians with a bigger picture of aid evaluation.

Thirdly, this thesis has deliberately excluded pre-war development and colonial policies. We know from the literature two things that raise interesting questions for further research. Firstly, the literature tells us that there were many continuities between pre-war development and/or colonial policies and their post-war successors. Dutch colonial development policies were systematically planned and included severe central government intervention.³⁸² Secondly, we know that evaluation of public policy in general also had roots before the Second World War. But were colonial development policies also evaluated? And if so, why then did it take the national development agencies until the late 1960s and 1970s to develop institutionalized evaluation practices? This is a question closely linked to the questions asked in this thesis, but a question that this thesis cannot answer on the basis of its time selection.

Corinna Unger concluded her book *International Development* by arguing that '[h]istorical perspectives can offer insight into the complexities of past development projects, and they can help us better understand which factors have contributed to making a project likely to fulfil its goals or have made it unlikely to do so.'³⁸³ She further quoted the many historians who have recently used historical research to gain knowledge useful for the contemporary development and humanitarian sector.³⁸⁴ This thesis has aimed to provide some insight into at least the

³⁷⁹ E.g. S.C. Dube, *India's Changing Villages: Human Factors in Community Development* (Ithaca 1958); M.J. Coldwell, R. Dumont, M. Read, *Report of a Community Development Evaluation Mission in India* (New Delhi 1959); U. Singh, *Community Development in India: Evaluation and Statistical Analysis* (Kanpur 1962).

³⁸⁰ Laporte, *L'évaluation, un objet politique*, 59.

³⁸¹ NL-HaNa, 2.05.330, Gemeenschappelijk Overleg Medefinanciering/Directie Internationale Technische Hulpverlening (GOM/DTH)-werkgroep Evaluatie en Procedures; vergaderingen, inventory number 33125, Konklusies van de konferentie te Zandvoort tussen de Minister voor Ontwikkelingssamenwerking en de drie bemiddelende organisaties (CEBEMO, ICCO, NOVIB) inzake het Medefinancieringsprogramma op 20 en 21 november 1974.

³⁸² De Jong, 'Onder ethisch insigne', 76.

³⁸³ Unger, *International Development*, 158.

³⁸⁴ E.g. Bayly, Rao, Szreter, Woolcock (eds.), *History, Historians and Development Policy*; J.N. Bolton, 'Improving the Use of History by the International Humanitarian Sector' *European Review of History* 1-2 (2016) 193-209.

complexity of a small and peculiar project within that sector: the evaluation of its activities. It has contributed to our understanding of its roots, its purposes and the historical interests of the actors involved. Whether current evaluation practices experience the same of measure of complexity as in the 1960s and 1970s, is up for future historians.

Annex I. LIST OF ABBREVIATIONS

AJV	<i>Afdeling Jonge Vrijwilligers</i>
CDP	United Nations Committee on Development Planning
CRM	<i>Ministerie voor Cultuur, Recreatie en Maatschappelijk Werk</i>
DAC	OECD Development Assistance Committee
DD	United Nations Development Decade (-I: 1960s; -II: 1970s)
DGIS	<i>Directoraat-Generaal Internationale Samenwerking</i>
DFID	Department for International Development (United Kingdom)
DFO	<i>Directie Financieel-Economische Ontwikkelingshulp</i>
DTH	<i>Directie Internationale Technische Hulp</i>
EEC	European Economic Community
EPTA	United Nations Expanded Programme for Technical Assistance
FAO	United Nations Food and Agriculture Organization
FMO	<i>Financierings-Maatschappij voor Ontwikkelingslanden</i>
IOV	<i>Inspectie Ontwikkelingssamenwerking te Velde</i>
ISMOG	<i>Instituut voor sociaal-economische studie van minderontwikkelde gebieden</i>
LFA	Logical Framework Approach
L&V	<i>Ministerie voor Landbouw en Visserij</i>
MFA	Ministry of Foreign Affairs
MFP	<i>Medefinancieringsprogramma</i>
NAR	<i>Nationale Adviesraad inzake de Ontwikkelingshulp</i>
NCO	<i>Nationale Commissie Ontwikkelingsstrategie '70-'80</i>
NIEO	New International Economic Order
OECD	Organisation for Economic Co-operation and Development
OED	World Bank Operations Evaluation Department
UNCTAD	United Nations Conference on Trade and Development
UNDP	United Nations Development Programme

Annex II. BIBLIOGRAPHY

Sources

Nationaal Archief:

- 2.05.118 Buitenlandse Zaken / Code-Archief 55-64
2.05.313 Buza / Code-archief 1965-1974
2.05.323 Nat. Adviesraad Ontwikkelingssamenwerking
2.05.330 Buza / Code 1975-1984

Delpher:

Trouw, De Volkskrant, De Telegraaf, Het Parool.

Publications of sources:

Nederlandse ontwikkelingssamenwerking II, 1964-1967, red. M.L.J. Dierikx, RGP (The Hague 2003).

Nederlandse ontwikkelingssamenwerking III, 1967-1973, red. M.L.J. Dierikx, RGP (The Hague 2005).

Nederlandse ontwikkelingssamenwerking IV, 1973-1977, red. M.L.J. Dierikx, RGP (The Hague 2005).

Staten-Generaal Digitaal:

Handelingen der Staten-Generaal, Eerste Kamer 1977-1978.

Handelingen der Staten-Generaal, Tweede Kamer, 1961-1962.

Handelingen der Staten-Generaal, Tweede Kamer, 1963-1964.

Handelingen der Staten-Generaal, Tweede Kamer, 1968-1969.

Handelingen der Staten-Generaal, Tweede Kamer, 1976-1977.

Huygens ING (<http://resources.huygens.knaw.nl/ontwikkelingssamenwerking>):

Ministerie van Buitenlandse Zaken, *Nota hulpverlening aan minder-ontwikkelde landen* (The Hague 1966).

_____, *Nota bilaterale ontwikkelingssamenwerking* (The Hague 1976).

Werkgroep Janssen, *Rapport Evaluatie van de Nederlandse Ontwikkelingshulp* (Tilburg 1969).

Other:

Brugman, J., 'De illusie van ontwikkelingshulp', *Hollands Maandblad*, mei 1968.

https://www.dbnl.org/tekst/hol006196901_01/hol006196901_01_0059.php.

- Coldwell, M.J., R. Dumont, M. Read, *Report of a Community Development Evaluation Mission in India* (New Delhi 1959).
- Dube, S.C., *India's Changing Villages: Human Factors in Community Development* (Ithaca 1958).
- Frank, A.G., *Capitalism and Underdevelopment in Latin America. Historical Studies of Chile and Brazil* (New York 1967).
- German Foundation for Developing Countries, *Methods and Procedures of Evaluation in Development Aid. Conference Report, 18-22 November 1966 in Berlin* (Berlin 1967).
- Hayes, Jr., S.P., *Evaluating development projects* (Louvain 1966 [1959]).
- Hirschman, A.O., *Development Projects Observed* (Washington 1967).
- Jackson, R.G.A., *A study of the capacity of the United Nations development system I & II* (Geneva 1969).
- Myrdal, G., *Asian Drama. An Inquiry into the Poverty of Nations* (New York 1968).
- Nationale Raad van Advies inzake Hulpverlening aan Minder Ontwikkelde Landen, *Advies Technische Hulp* (The Hague 1972).
- _____, *Advies Evaluatie van de Nederlandse ontwikkelingshulp* (The Hague 1972).
- OECD, *Aid Evaluation. The Experience of Members of the Development Assistance Committee and of International Organisations* (Paris 1975).
- _____, *The Evaluation of Technical Assistance* (Paris 1969).
- Pearson, L.B., et al., *Partners in Development. Report of the Commission on International Development* (New York 1969).
- Pronk, J.P., 'Doe wel en zie niet om' (1 February 2007)
<http://www.janpronk.nl/speeches/dutch/doe-wel-en-zie-niet-om.html> (retrieved 12 March 2019).
- Singh, U., *Community Development in India: Evaluation and Statistical Analysis* (Kanpur 1962).
- Taghi Farvar, M., J.P. Milton (eds.), *The Careless Technology: Ecology and International Development* (New York 1969).
- Truman, H., 'Inaugural Address' (edition 20 January 1949)
<http://www.let.rug.nl/usa/presidents/harry-s-truman/inaugural-address-1949.php>
 (15 August 2019).
- UNITAR, *UN Development Aid. Criteria and Methods of Evaluation* (New York 1971).
- United Nations Economic and Social Council, *Committee for Development Planning, Report on the Sixth Session (5-15 January 1970)* (New York 1970).
- UN Joint Inspection Unit, *Evaluation System of the United Nations Development Programme*, (New York 1983).

Literature

- Aberbach, J.D., R.D. Putnam, B.A. Rockman, *Bureaucrats and Politicians in Western Democracies* (Cambridge, MA 1981).
- Arens, E.H., 'Multilateral Institution-Building and National Interest: Dutch development policy in the 1960s' *Contemporary European History* 4 (2003) 457-472.
- Arndt, H.W., 'Economic Development: A Semantic History' *Economic Development and Cultural Change* 3 (1981) 457-466.
- Arnold, S.H., *Implementing Development Assistance: European Approaches To Basic Needs* (Boulder 1982).
- Bayly, C.A., V. Rao, S. Szreter, M. Woolcock (eds.), *History, Historians and Development Policy. A necessary dialogue* (Manchester 2011).
- Berlage, L., O. Stokke (eds.) *Evaluating Development Assistance. Approaches and Methods* (London 1992).
- Berg, R.D., van den, 'Leert de hulp van lessen uit het verleden? Doe wel en zie niet om' *Internationale Spectator* 5 (2001) 253-260.
- Beurden, J. van, J.B. Gewalt, *From Output to Outcome? 25 Years of IOB Evaluations* (Amsterdam 2004).
- Beusekom, M. van, *Negotiating Development: African Farmers and Colonial Experts at the Office du Niger, 1920-1960* (Oxford 2002).
- Bieckmann, F., E. Lammers, *Hivos! 1968-2008. Een onvoltooide geschiedenis: van hulp naar andersglobalisme* (Amsterdam 2008).
- Binnendijk, A.L., 'Donor agency experience with the monitoring and evaluation of development projects' *Evaluation Review* 3 (1989) 206-222.
- Bolton, J.N., 'Improving the Use of History by the International Humanitarian Sector' *European Review of History* 1-2 (2016) 193-209.
- Borowy, I., M. Schmelzer (eds.), *History of the Future of Economic Growth: Historical Roots of Current Debates on Sustainable Degrowth* (Abingdon 2017).
- Born, J., "Effect van miljoenen aan ontwikkelingshulp moet beter worden onderzocht" (9 July 2019) <https://eenvandaag.avrotros.nl/item/effect-van-miljoenen-aan-ontwikkelingshulp-moet-beter-worden-onderzocht/> (consulted 12 August 2019).
- Bovens, M., *De Vierde Macht Revisited. Over ambtelijke macht en publieke verantwoording* (Utrecht 2000).
- Brandsma, M., P. Klein, *Jan Pronk. Rebel met een missie* (Utrecht 1996).
- Brolin, T., 'Conceptualizations of "results" in Swedish policy for development cooperation from the 1960s to the 2000s' *Bulletin of Geography. Socio-economic Series* 33 (2016) 7-22.
- Brüne, S. (ed.), *Erfolgskontrolle in der entwicklungspolitischen Zusammenarbeit* (Hamburg 1998).

- Caspari, A., *Evaluation der Nachhaltigkeit von Entwicklungszusammenarbeit. Zur Notwendigkeit angemessener Konzepte und Methoden* (Wiesbaden 2004).
- Clements, P., T. Chianca, R. Sasaki, 'Reducing world poverty by improving evaluation of development aid' *American Journal of Evaluation* 2 (2008) 195-214.
- Cooper, F., 'Writing the History of Development' *Journal of Modern European History* 2 (2010) 5-23.
- Chianca, T., 'International aid evaluation: An analysis and policy proposals' (edition April 2008) <https://scholarworks.wmich.edu/dissertations/758/> (31 May 2018).
- _____, 'The OECD/DAC Criteria for International Development Evaluations: An Assessment and Ideas for Improvement' *Journal of Multidisciplinary Evaluation* 9 (2008) 41-51.
- Cracknell, B.E., (ed.), *The Evaluation of Aid Projects and Programmes. Proceedings of the Conference organised by the Overseas Development Administration in the Institute of Development Studies, at the University of Sussex 7-8 April 1983* (London 1984).
- Cracknell, B.E., *Evaluating Development Aid: Issues, Problems and Solutions* (London 2000).
- Damme, L.J. van, M.G.M. Smits (eds.), *Voor de ontwikkeling van de derde wereld. Politici en ambtenaren over de Nederlandse ontwikkelingssamenwerking, 1949-1989* (Amsterdam 2009).
- Dahler-Larsen, P., *The Evaluation Society* (Stanford 2012).
- Dierikx, M.L.J., 'Like a fish out of water: experimenting fishery in Lake Victoria, 1964-1974' *Special Working Paper Series on 'Unintended Effects of International Cooperation'* 1 (2017).
- _____, 'Policy versus Practice. Behind the Scenes in Dutch Development Aid, 1949-1989' *The International History Review* 4 (2017) 638-653.
- _____, 'Ontwikkelingssamenwerking, spagaat tussen moraliteit en handel' *Historisch Nieuwsblad* 5 (2002).
- Dr. Kelder en Co, 'Het "rendement van ontwikkelingssamenwerking', (1 December 2018) <https://www.nporadio1.nl/dr-kelder-en-co/onderwerpen/481922-het-rendement-van-hulp> (consulted 12 August 2019).
- Easterly, W., *The White Man's Burden. Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good* (New York 2006).
- Ekbladh, D., *The Great American Mission: Modernization and the Construction of an American World Order* (Princeton 2009).
- Escobar, A., *Encountering Development: the making and unmaking of the Third World* (Princeton 1995).
- Engerman, D.C., 'The Romance of Economic Development and New Histories of the Cold War' *Diplomatic History* 1 (2004) 23-54.

- Fischer, F., G.J. Miller, M.S. Sidney (eds.), *Handbook of Public Policy Analysis: Theory, Politics, and Methods* (Boca Raton 2007).
- Frey, M., S. Kunkel, C.R. Unger (eds.), *International Organizations and Development 1945-1990* (Basingstoke 2014).
- Frissen, P., *Het geheim van de laatste staat. Kritiek van de transparantie* (Amsterdam 2016).
- Geselbracht, R.H., (ed.), *Foreign aid and the legacy of Harry S. Truman* (Kirksville 2014).
- Gilman, N., *Mandarins of the Future: Modernization Theory in Cold War America* (Baltimore 2003).
- Graaff, B. de, D. Hellema, B. van der Zwan (eds.), *De Nederlandse buitenlandse politiek in de twintigste eeuw* (Amsterdam 2003).
- Grasso, P., S.S. Wasty, R. Weaving (eds.), *World Bank Operations Evaluation Department: The First 30 Years* (Washington 2003).
- Haas, P.M., 'Introduction: Epistemic Communities and International Policy Coordination' *International Organization* 1 (1992) 1-35.
- Hatton, M.J., K. Schroeder, 'Results-based management: friend or foe?' *Development in Practice* 3 (2007) 426-432.
- Haveman, R.H., 'Policy Analysis and Evaluation Research After Twenty Years' *Policy Studies Journal* 2 (1987) 191-218.
- Hellema, D.A., *Dutch Foreign Policy: The Role of the Netherlands in World Politics* (Dordrecht 2009).
- _____, *Neutraliteit & Vrijhandel. De geschiedenis van de Nederlandse buitenlandse betrekkingen* (Amsterdam 2002).
- Hilton, M., 'Charity and the End of Empire: British Non-Governmental Organizations, Africa and International Development in the 1960s' *The American Historical Review* 2 (2018) 493-517.
- Hodge, J.M., 'Writing the History of Development (Part 1: The First Wave)' *Humanity* 4 (2015) 429-463.
- _____, 'Writing the History of Development (Part 2: Longer, Deeper, Wider)' *Humanity* 2 (2016) 125-174.
- Hoebink, P., *Geven is nemen. De Nederlandse ontwikkelingshulp aan Tanzania en Sri Lanka* (Nijmegen 1988).
- _____, 'Hoe de dominee de koopman versloeg. Nederlandse ontwikkelingssamenwerking gewogen' *Internationale Spectator* 11 (2006) 578-584.
- _____, 'The Humanitarianisation of the Foreign Aid Programme in the Netherlands' *The European Journal of Development Research* 1 (1999) 176-202.
- Hood, C., 'A public management for all seasons?' *Public Administration* 1 (1991) 3-19.

- Hoogerwerf, A., (ed.), *Succes en falen van overheidsbeleid* (Alphen aan den Rijn 1983).
- Hueting, E., *De permanente herstructurering in het welzijnswerk* (Zutphen 1989).
- Jolly, R., L. Emmerij, T.G. Weiss, *UN Ideas that Changed the World* (Bloomington 2009).
- Kantowsky, D., (eds.), *Evaluierungsforschung und -praxis in der Entwicklungshilfe* (Zürich 1977).
- Keulen, S., *Monumenten van Beleid. De wisselwerking tussen Nederlands rijksoverheidsbeleid, sociale wetenschappen en politieke cultuur, 1945-2002* (Hilversum 2014).
- Kuitenbrouwer, M., *De ontdekking van de Derde Wereld: beeldvorming en beleid in Nederland, 1950-1990* (The Hague 1994).
- _____, 'Een eeuwig durende ereschuld? Ontwikkelingssamenwerking en koloniaal verleden' *Internationale Spectator* 7/8 (2000) 377-382.
- Kulnazarova, A., V. Popovski (eds.), *The Palgrave Handbook of Global Approaches to Peace* (Basingstoke 2019).
- Laporte, C., (Ph.D. Thesis), *L'évaluation, un objet politique: le cas d'étude de l'aide au développement*, Institut d'études politiques de Paris, March 2015.
- Lane, J.E., *The Public Sector* (London 1993).
- Latham, M.E., *Modernization as Ideology: American Social Science and "Nation Building" in the Kennedy Era* (Chapel Hill 2000).
- Lankester, T., "'Asian Drama": the pursuit of modernization in India and Indonesia' *Asian Affairs* 3 (2004) 291-304.
- Leeuw, F.L., 'Evaluation Policy in the Netherlands' *New Directions for Evaluation* 3 (2009) 87-102.
- Macekura, S.J., E. Manela (eds.), *The Development Century. A Global History* (Cambridge MA 2018).
- Morra Imas, L.G., R. Rist, *The road to results: Designing and conducting effective development evaluations* (Washington 2009).
- Nekkers, J.A., P.A.M. Malcontent (eds.), *De geschiedenis van vijftig jaar Nederlandse ontwikkelingssamenwerking 1949-1999* (The Hague 1999).
- Nelkin, D., 'Scientific Knowledge, Public Policy, and Democracy' *Knowledge, Creation, Diffusion, Utilization* 1 (1979) 106-122.
- Nisbet, R., *Social Change and History: Aspects of the Western Theory of Development* (New York 1969).
- Nispen, F. van, P. Scholten, *Policy analysis in the Netherlands* (Bristol 2014).
- Oakley, A., 'Public Policy Experimentation: Lessons from America' *Policy Studies* 2 (1998) 93-114.
- OECD, *A History of the DAC Expert Group on Aid Evaluation* (Paris 1993).
- _____, *DAC in Dates. The History of the OECD's Development Assistance Committee* (Paris 2006).

- _____, *The Story of Official Development Assistance. A History of the Development Assistance Committee and the Development Co-operation Directorate in Dates, Names and Figures* (Paris 1996).
- Ongaro, E., S. van Thiel (eds.), *The Palgrave Handbook of Public Administration and Management in Europe* (London 2018).
- Osborne, S.P., 'The New Public Governance?' *Public Management Review* 3 (2006) 377-387.
- Packenham, R.A., *Liberal America and the Third World: Political Development Ideas in Foreign Aid and Social Science* (Princeton 1973).
- Packard, R., F. Cooper (eds.), *International Development and the Social Sciences: Essays on the History and Politics of Knowledge* (Berkeley 1997).
- Park, A.S., 'Does Development Learn from History?' *World Development* (2017) 52-64.
- Pollitt, C., 'Occasional excursions: A brief history of policy evaluation in the UK' *Parliamentary Affairs* 3 (1993) 353-363.
- Radin, B.A., *Beyond Machiavelli. Policy Analysis Comes of Age* (Washington D.C. 2000).
- Rist, C. (ed.), *Program Evaluation and the Management of Government. Patterns & Prospects across Eight Nations* (New Brunswick 1999).
- Rist, G., *The history of development: from western origins to global faith* (London 2014).
- Rossi, P.H., H.E. Freeman, M.W. Lipsey, *Evaluation. A systematic approach* (Thousand Oaks 1999).
- Sachs, W., *The development dictionary: a guide to knowledge as power* (London 1992).
- Sasaki, R., 'A Review of the History and the Current Practice of Aid Evaluation' *Journal of Multidisciplinary Evaluation* 5 (2006) 55-88.
- Scott, J.C., *Seeing like a state: how certain schemes to improve the human condition have failed* (New Haven 1998).
- Sharma, P.A., *Robert McNamara's Other War: The World Bank and International Development* (Philadelphia 2017).
- Sorber, A., J.R. Bax, A.J.M. Hendriks, *Beleidsevaluatie: wisselwerking tussen cultuur en structuur* (The Hague 1995).
- Stokke, O. (ed.), *Evaluating development assistance: policies and performance* (London 1991).
- _____, *Western Middle Powers and Global Poverty. The Determinants of the Aid Policies of Canada, Denmark, the Netherlands, Norway and Sweden* (Uppsala 1989).
- Stokke, O., *International Development Assistance. Policy Drivers and Performance* (Basingstoke 2019).
- _____, *The UN and Development: From Aid to Cooperation* (Bloomington 2009).
- Tignor, R.L., *W. Arthur Lewis and the Birth of Development* (Princeton 2006).

- Unger, C., *Entwicklungspfade in Indien. Eine internationale Geschichte, 1947-1980* (Göttingen 2015).
- _____, *International Development. A Postwar History* (London 2018).
- Voorhoeve, J.J.C., *Peace, Profits and Principles. A Study of Dutch Foreign Policy* (Leiden 1985).
- Wagner, P. (ed.), *Social sciences and modern states: national experiences and theoretical crossroads* (Cambridge 1991).
- Wisler, C. (ed.), 'Evaluation and Auditing: Prospects for Convergence' *New directions for evaluation* 3 (1996) 1-67.
- Zwan, B. van der, B. de Graaff, D. Hellema (eds.), *In dienst van Buitenlandse Zaken. Achttien portretten van ambtenaren en diplomaten in de twintigste eeuw* (Amsterdam 2008).