

d·i·e

Deutsches Institut für
Entwicklungspolitik



German Development
Institute

Discussion Paper

19/2020

How the G7 Reviews its Work on Development

A Case Study of Internal Accountability

Roger A. Fischer

How the G7 reviews its work on development

A case study of internal accountability

Roger A. Fischer

Bonn 2020

Discussion Paper / Deutsches Institut für Entwicklungspolitik
ISSN (Print) 1860-0441
ISSN (Online) 2512-8698



Except as otherwise noted this publication is licensed under Creative Commons Attribution (CC BY 4.0). You are free to copy, communicate and adapt this work, as long as you attribute the German Development Institute / Deutsches Institut für Entwicklungspolitik (DIE) and the author.

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available at <http://dnb.d-nb.de>.

ISBN 978-3-96021-130-3 (printed edition)

DOI: 10.23661/dp19.2020

Printed on eco-friendly, certified paper

Dr Roger Fischer has led the G7/G20 division in Germany's Federal Ministry for Economic Cooperation and Development (BMZ). He is an Executive Director of the Asian Development Bank (ADB). The contents of this paper cannot be construed as offering any indication of opinions or positions of Germany's Federal Government or its parts.

Email: otradek24@t-online.de

Published with financial support from the Federal Ministry for Economic Cooperation and Development (BMZ)

© Deutsches Institut für Entwicklungspolitik gGmbH
Tulpenfeld 6, 53113 Bonn
☎ +49 (0)228 94927-0
☎ +49 (0)228 94927-130
Email: die@die-gdi.de
www.die-gdi.de



Abstract

The G7 practices forms of external accountability to answer for its behaviour to the outside world and internal accountability to steer the implementation of what it has decided. To follow up on its development-related commitments, it has set up a permanent framework to produce annual public reports on how G7 national administrations have worked together to implement them.

Reports under this framework draw on implementation experience, although G7 Leaders never use them to make decisions on how to carry implementation forward or design new commitments. This is because the G7 process is generally discontinuous and because its accountability process is currently not targeted at channelling feedback from implementation experience back into policymaking. The learning potential inherent in *internal accountability* is thus not fully used. As a result, the G7 is less effective than it could be in implementing its commitments under changing circumstances. In addition, G7 commitments and methodology do not always make it easy for outside stakeholders to check the G7's words against its behaviour, even though this is important for *external accountability*. Consequently, this makes it harder to have a rational debate about the legitimacy of the G7.

This paper suggests ways to improve the accountability practice of the G7 so that it systematically produces learning effects and thereby better supports G7 legitimacy. The G7 can achieve *learning* effects by underpinning every commitment with an explicit notion of "how it wants to achieve what" and making sure that this notion gets tested regularly against implementation experience. Closing this feedback loop could be a job for G7 portfolio ministers who could make decisions on further implementation, based on the experience set out in an accountability report. Better designed commitments and improved follow-up would also support G7 *legitimacy* because this would make it easier for external stakeholders to verify G7 actions against its words.

Keywords: G7, accountability, legitimacy

Contents

Abstract

Abbreviations

Executive summary	1
1 Introduction	3
2 Three starting points	5
2.1 What is the G7?	5
2.2 What is accountability and what is it good for?	8
2.2.1 Four forms of accountability	9
2.2.2 Elements of legitimacy	11
2.2.3 Managing delegated behaviour	13
2.3 How the G7 has used accountability instruments	15
2.3.1 Reporting mandates to the group of G7 national administrations	15
2.3.2 Reporting mandates to international organisations	16
2.3.3 Dialogue with third countries	16
2.3.4 Dialogue with civil society	17
3 The accountability framework for G7 commitments on sustainable development	18
3.1 Scope and limitations	18
3.2 Features	19
3.3 Stated purposes	20
3.4 History	21
4 How the G7 reviews its commitments on development (process)	22
4.1 The life of a G7 commitment	22
4.1.1 How the G7 comes to its commitments	22
4.1.2 When is a summit outcome fit for accountability?	23
4.1.3 How the AWG processes a commitment	25
4.1.4 Reasons to stop reporting (sunset)	26
4.2 Rules of reporting (methodology)	27
4.2.1 Understanding a commitment (interpretation)	27
4.2.2 Base line	28
4.2.3 Indicators	28
4.2.4 Data sources	29
4.3 How does this structure perform?	30
4.3.1 Quality of the commitments	30

4.3.2	Quality of the process	32
5	Ways to improve accountability	33
6	Summary and conclusions	34
	References	37

Tables

Table 1:	Forms of accountability	9
----------	-------------------------	---

Abbreviations

AML/CFT	Anti-money laundering/Combatting the financing of terrorism
APR	annual progress report
AWG	Accountability Working Group (of the G7)
B7	Business 7 (engagement group)
BEPS	base erosion and profit shifting (international tax agenda)
BRICS	Brazil, Russia, India, China and South Africa
C7	Civil 7 (engagement group)
CAR	comprehensive accountability report
DAC	Development Assistance Committee (of the OECD)
EBRD	European Bank for Reconstruction and Development
FATF	Financial Action Task Force
FSWG	Food Security Working Group (of the G7)
G7	Group of Seven
G8	Group of Eight
G20	Group of Twenty
GDP	gross domestic product
GFATM	Global Fund to Fight against Aids, Tuberculosis and Malaria
IBRD	International Bank for Reconstruction and Development
IDA	International Development Association
IFC	International Finance Corporation
IMF	International Monetary Fund
L7	Labour 7 (engagement group)
MDB	multilateral development bank
NEPAD	New Partnership for Africa's Development
ODA	Official Development Assistance
OECD	Organisation for Economic Cooperation and Development
RDB	regional development bank
S7	Science 7 (engagement group)
T7	think tank 7 (engagement group)
TFA	Trade Facilitation Agreement
UNCAC	United Nations Convention against Corruption
W7	Women 7 (engagement group)
Y7	Youth 7 (engagement group)

Executive summary

The G7 must answer for its behaviour to the outside world, while its national administrations are collectively responsible to the group of their Leaders. These are two distinct issues playing out in two different relationships: one between the G7 and external stakeholders; and the other between the G7 and its national administrations. The G7 has developed a number of practices to ensure it is accountable to stakeholders (*external accountability*) and able to lead national administrations in the implementation of G7 mandates (*internal accountability*).

While most accountability practices in the G7 are informal and temporary, it has established a permanent framework to follow up on its development-related summit commitments. Under this framework, G7 national administrations collectively report on how they have implemented these commitments. This practice can serve *external accountability* by showing outside stakeholders what the G7 has done to live up to its commitments and it can serve *internal accountability* by allowing Leaders to control how commitments are carried out and to decide on a way forward. But actual G7 performance is less than perfect on both counts. Outside stakeholders do not always have a chance to effectively verify G7 actions against its words because some commitments lack clarity and reporting is not as focused as it could be. The process inside the G7 does not make it easy for Leaders to use accountability reports as a basis for their decisions: when Leaders design new policies, they never use them. As a result, the feedback loop built into internal accountability does not work and opportunities to use implementation experience for better policymaking are lost.

This paper develops recommendations to improve the accountability practice of the G7 so that it can make a stronger contribution to G7 legitimacy and better support institutional learning. These recommendations are based on a close description of current G7 practice, a review of two bodies of literature (one on legitimacy in international institutions and another on contract design in agency relationships), and a system of four forms of accountability. Together, these elements show the leverage points within G7 accountability practice that can be used to improve learning and legitimacy.

The *description* of G7 practice serves to highlight the informal and discontinuous nature of the G7. This is important because it sets limitations to what the G7 can do in terms of accountability. For example, it cannot produce a formal procedure to remedy grievances.

The literature on *legitimacy* underlines that, in order for *external accountability* to be effective, stakeholders must be able to check an international institution's action against its words. That is why every G7 commitment should state a purpose and define a clear standard for how the G7, as a group, is supposed to behave. Accountability reports must focus on whether the purpose has been achieved and the standard met. This paper develops rules for commitment design and methodology to achieve that.

The findings of *contract theory* show how a system of *internal accountability* must work in order to deliver learning effects and to enable informed decision-making. This point will help G7 decision-makers see what they can do about the ongoing implementation of a previous commitment (for instance, increase the effort, change the approach, or adjust the goal) and help them make robust choices among the options. That is not easy because G7 Leaders generally delegate implementation to agents who are better informed and have

different incentives. Contract theory offers practical ways for Leaders to extract actionable information and neutralise adverse incentives. In the G7 context, this would lead to changes in process, commitment design, and methodology. I make specific recommendations on all of those points. In particular, G7 portfolio ministers can play an important role in translating implementation experience into better decision-making.

A systematic look at different *forms of accountability* shows that every accountability practice must be measured against what it is trying to achieve. G7 practice is primarily designed to serve purposes of internal accountability; this is why improving its learning function is so important. While the G7 uses its practice for external accountability as well, this purpose does not shape its basic design. There are many ways to improve G7 external accountability; however this paper does not expand on them because they largely depend on what the G7 does in terms of policies, as opposed to how it does this in terms of procedure. The recommendations I develop here are strictly limited to the procedural way the G7 approaches accountability. Whether it does the *right* thing is a matter for further research and is ultimately political decision.

1 Introduction

At their summit in 2015, G7 Leaders made this agreement:

As part of a broad effort involving our partner countries and international actors, and as a significant contribution to the Post-2015 Development Agenda, we aim to lift 500 million people in developing countries out of hunger and malnutrition by 2030. (G7, 2015a, p. 19)

The main significance here is in the action the G7 takes against hunger and malnutrition. However, around this action there is communication: *Outside*, the agreement raises expectations about what the G7 can achieve, so there is pressure to prove that actions are consistent with words and that the G7 makes an impact on the ambitious goal it has stated. *Inside*, it forces G7 Leaders to make sure their national administrations do their part to reach that goal. They must lead implementation to make it effective.

Proving your worth to the outside world and managing delegated behaviour are different challenges but both have to do with accountability (Keohane, 2002, p. 20). While the first plays out in the *external* relationship between the G7 and stakeholders who question whether it is doing the right thing, the second occurs within the *internal* relationship between G7 Leaders who want results and their national administrations who are closer to the action but have different incentives. The former case is about *justifying behaviour* with reference to a stable benchmark (“doing the right thing”), the latter about *extracting information* to use it for dynamic guidance (“to deliver results in a changing environment”). External accountability is more political, because it ultimately asks whether the agent has essentially done the right thing, while whoever exerts external accountability (the principal) has no direct control over the agent’s operations. In some cases, external accountability only serves to express the claim to control the agent (Keohane, 2002, p. 13). In external accountability, the principal, having no operative control, is not engaged in managing changing circumstances. This is why external accountability often relies on rigid normative standards. On the other hand, internal accountability is more managerial, because it asks how the agent can best serve whatever purpose the principal has set. The principal here has leverage over the agent’s operations and can use it to change his/her guidance according to circumstances. For this reason, internal accountability often relies on flexible standards meant to ensure effectiveness. There are variants of external and internal accountability (two of each, see subsection 2.2.1), but the basic difference is whether accountability is owed to an external or to an internal stakeholder.

The G7 has to do both: It practices various ways to answer for its behaviour (*external accountability*) and lead implementation (*internal accountability*). This paper is meant to develop options to improve G7 accountability for its development-related summit commitments. As background, I offer an overview of the different ways the G7 practices accountability but ultimately focus on the framework G7 Leaders use to monitor the implementation of development-related Leader commitments such as the one I have quoted above. I show that this practice, while useful, could do better in producing actionable information and learning effects, and suggest how the G7 can make these gains without fundamentally changing its framework. I argue that it is trying to achieve external *and* internal accountability purposes in a hybrid process that is as rigid as external accountability (because Leaders never use it to change course) and as uncritical as internal accountability

(because national administrations together report about their own work). My suggestions try instead to combine the best of both worlds.

Research on the G7 (Kirton & Larionova, 2018; Brandi, 2019) and international institutions in general (Schedler, 1999; Keohane, 2002; Buchanan & Keohane, 2006) has made important progress in understanding how they can become accountable to outside stakeholders and how this contributes to their legitimacy. Contract theory (Schmidt, 2017) has shown how a principal can best reach his/her goals working with an agent who has better information and different incentives. Both strands of thought are relevant to practices of accountability, but they have never been systematically used to look at the G7. To make their findings useful for G7 practice, I take external and internal categories of accountability and make a distinction within each of them: Does the agent report on his own work or do reports come from a third party? This distinction leads to a *systematic conceptual framework* of four forms of accountability (see subsection 2.2.1) that are closely related but different. This framework prevents the confusion that comes with applying a uniform standard to different forms of accountability and helps one understand the kinds of problem each form is meant to solve. I highlight the *specific nature* of the G7 and explain why some accountability practices found in other institutions do not work in it. I stress the *limitations* of the G7 accountability framework for development-related commitments, to make clear what it is *not* designed to do. On this basis, I translate the findings of legitimacy research and contract theory into *seven criteria* to test how the practice under the framework performs against the goals that the G7 has set for it. The results show strengths and weaknesses that I then use to develop my suggestions. I draw from a selective review of the literature and my professional experience as a member of the G7 Accountability Working Group (AWG), which may account for any bias in my paper.

This paper is intended for *researchers* who might want to process it for further work; *practitioners*, particularly my colleagues in the AWG, who are welcome to use it as a reference even if they may disagree with the points I make; *decision-makers* who are invited to consider my suggestions; *external stakeholders* who want to hold the G7 to account; and *every reader* who wants to know how internal accountability in the G7 works and why it so often disappoints.

My paper is structured as follows:

Section 2 lays the ground for understanding accountability in the G7. There, I first explain the informal and discontinuous *nature* of the G7 (subsection 2.1). Subsection 2.2 then deals with *theory*. Here I develop a systematic conceptual framework of four different forms of accountability (2.2.1) and present selected findings from the literature on legitimacy in international institutions (2.2.2) and contract theory (2.2.3). These findings lead to the *criteria* I use in subsection 4.3 to test the G7 practice under the accountability framework for development-related commitments against its stated goals. In subsection 2.3, I show the various ways the G7 has used accountability instruments *in practice*.

Section 3 explains a special case of this practice, the *G7 accountability framework for development-related commitments*. It presents what this mechanism is designed to do and what its limitations are (3.1), it traces its overall features (3.2), the purposes the G7 want it to serve (3.3), and a bit of its history (3.4).

The next section (4) shows how this framework operates *as a process* and how the AWG works. The section follows the life of a commitment from inception to sunset and offers details about how the AWG decides which summit outcomes are followed up (4.1); furthermore it presents how the AWG sets the rules for reporting (4.2) and then uses the criteria developed in subsection 2.3 to test G7 practice against its stated goals (4.3).

The results of this test lead to *suggestions to improve G7 practice* (Section 5). Their main point is to improve the frameworks' learning function.

Section 6 is a summary, setting out my main messages and suggesting issues for further research.

2 Three starting points

2.1 What is the G7?

The G7 is a group consisting of the heads of state and government (*Leaders*) representing Canada, France, Germany, Italy, Japan, the United Kingdom and the United States, joined by the President of the European Council and the President of the European Commission. From 1998 to 2014, the President of Russia was a member of what was then the G8. The group meets once a year for a summit at the invitation of the presiding member. The G7 presidency rotates annually among member countries, so that each of them gets to chair the group over a seven-year cycle. The presiding country has a number of privileges. It sets the issues for discussion and the goals for the summit. It designs the structure and sequence of meetings that lead up to the summit. It shapes the way the G7 engages with civil society and business organisations. It decides which countries outside the G7 are invited to the summit and other meetings (*outreach*). The entire process is *informal* in the sense that it is not constrained by a fixed mandate or a permanent structure such as a secretariat. It also does not result in legally binding outcomes. While the original focus was on coordinating G7 countries' economic policies to respond to the oil price shock of the 1970s and the breakdown of the Bretton Woods system of fixed currency exchange rates, the G7 has made various impacts on the world at large. Enhanced debt relief for highly indebted poor countries, the establishment of the Global Fund to Fight Aids, Tuberculosis and Malaria (GFATM) and support for the containment of the reactor meltdown in Chernobyl all lead back to G7 initiatives.

While Leaders appreciate the informality, they sometimes wish they could concentrate on fewer, more acute topics. This statement shows how they would prefer to seize the moment and resist bureaucratic process:

We have reflected on how Summits could best focus our attention on the most significant issues of the time. We value Summits for the opportunity they provide to exchange views, build consensus and deepen understanding among us. But we believe Summits should be less ceremonial, with fewer people, documents and declarations, and with more time devoted to informal discussion among us, so that together we may better respond to major issues of common concern. (Tokyo, 1993, 16)¹

Whatever the G7 decides at a summit is published in at least one document. This is either a text agreed by all G7 members (*Leaders' statement*) or a declaration of the Chair reflecting its own sense of the summit outcomes (*Chair's summary*). G7 presidencies prefer Chair's summaries if they are certain about what Leaders agree on and want to avoid lengthy text negotiations. Summit documents often come with more detailed issue-specific *annexes* that Leaders either endorse or just acknowledge. Summit documents are important reference points for internal and external accountability because they show what the G7 Leaders collectively want (see subsection 4.2.1) and what they have promised to the public.

The informal nature of the G7 makes it flexible. But it has five stable features that describe the way it generally works. In and around the G7 there are *discipline*, *demands*, *delegation*, *discontinuity* and a *disconnect*:

- One can only speak of G7 action if members agree to *act as a group*, each member operating in a way that complements the others' behaviour. This can range from coordinating policies (to make sure one does not counteract the other) to engaging in collective action (where cooperation achieves more than the sum of each part). Whatever distinguishes G7 group action from each member operating independently can be called *discipline*.
- As top representatives of seven powerful countries, the G7 *draw certain expectations* about what they can achieve together as well as concerns about what happens if the G7 cannot agree, do not solve problems or even engage in harmful policies. As Leaders of democracies, they are formally accountable to their people. The outside world, non-G7-countries and civil society everywhere confront the G7 with very specific *demands* that would not be equally directed at another group. For example, the G20 has to do with a different set of expectations.
- To get things done, the G7 must turn either to their respective national administrations or to international organisations because there is no G7 implementation structure. The G7 relies on *delegation*.
- The G7 normally engages in single discrete initiatives rather than long-term programmes. This is because, every year, a new G7 presidency sets its own agenda. There is *discontinuity*.

1 The University of Toronto's Munk School of Global Affairs and Public Policy operates a useful website where G7 summit documents are found: <http://www.g8.utoronto.ca> . I quote them using the name of the city where the summit was held, the year of the summit, and the number of the paragraph I refer to.

- G7 Leaders do not often revisit an issue once they have delegated it. Delegated tasks originate in previous summits. G7 Leaders focus on the current year's agenda and the impact they can make in the future. This results in a *disconnect* between backward-looking implementation reports and forward-looking summit agendas. This split is underpinned by a distance in hierarchy: Civil servants implement, Leaders decide.

Many of the largest economies today (China, India, Russia, Indonesia) do not belong to the G7. This is why the relevance of the G7 has been questioned (O'Neill & Terzi, 2018). But when the G7 acts as a single unit it can still project considerable power. Together, G7 member states have voting shares of over 41 per cent in the International Monetary Fund (IMF), the International Bank for Reconstruction and Development (IBRD) and the International Development Association (IDA), over 48 per cent in the International Finance Corporation (IFC) and over 56 per cent in the European Bank for Reconstruction and Development (EBRD). G7 member states' voting shares in the regional development banks (RDBs) are all closely above or under 30 per cent. Representing 58 per cent of the world's net wealth and 46 per cent of global gross domestic product (GDP), G7 countries have globally important markets and can exercise leverage by conditioning access to them. The G7 provides over 75 per cent of all Official Development Assistance (ODA) (G7, 2019a, p. 8). All this means that G7 priorities can make a difference to global sustainable development, especially if it coherently brings its leverage points to bear in different policy areas.

The G7 can only use its power if it acts as a group (*discipline*). Therefore, all G7 procedures are meant to build *consensus*. Deliberations in the G7 proceed from the bottom up and culminate in the summit where Leaders engage in final negotiations and agreement.

- *G7 working groups* (such as the Food Security Working Group (FSWG) or the Accountability Working Group (AWG); see subsection 3.2) cover specific subject areas and bring together mid-level civil servants from relevant sector ministries. Each G7 member government appoints its representative in a working group. Working groups report their agreed outcomes either to the relevant ministers meeting or directly to sherpas or finance deputies.
- *G7 ministers' meetings* (such as the 2019 G7 Environment Ministers meeting in Metz) cover subjects within the remit of their portfolio. Most G7 presidencies convene at least meetings of Finance Ministers and Central Bank Governors as well as Foreign Ministers from each G7 member. Depending on its priorities, a presidency can also arrange meetings of other portfolio ministers. G7 ministers often issue a public outcome document after their meeting. Leaders can use the summit to go beyond what ministers have agreed. Apart from that, ministerial outcomes represent G7 consensus.
- *G7 finance deputies* (high-level civil servants appointed by each G7 Finance Minister) negotiate the outcomes the presidency wants to achieve on issues of finance. Depending on the presidency's goals, this can result in Finance Ministers' and/or summit outcomes. For part of their work, finance deputies rely on the outcomes of working groups.
- *G7 sherpas* (high-level civil servants appointed by each G7 Leader) negotiate the outcomes the presidency wants to achieve in a summit. They cover what working groups and ministers have produced if the presidency wants to raise it to the Leader level. The work of the finance deputies is an exception because they negotiate the finance part of

the summit outcome. In practice, sherpas and finance deputies draft, negotiate and agree on most of the text that Leaders eventually endorse and issue at the end of their summit.

- *G7 Leaders* negotiate outcomes on the issues sherpas or finance deputies have not already agreed on. For the rest, they normally endorse the text their civil servants have previously finalised. Against this background, the G7 declaration issued in Biarritz on 26 August 2019 (G7, 2019b) is an exception because it says that Leaders drafted and negotiated it themselves (see *consensus*, under subsection 4.1.2).
- *Outreach partners* are representatives of non-G7 countries that the presidency invites to the summit or other G7 meetings. G7 dialogue with them is temporary and limited to the issues the presidency has set (see subsection 2.3.3).
- *Engagement groups* are representatives of civil society and business organisations from G7 countries. Currently the G7 recognises the groups of their countries' organisations representing business (B7), civil society (C7), trade unions (L7), the science community (S7), think tanks (T7), women (W7) and youth (Y7). Each G7 presidency appoints organisations in its country to lead the engagement groups during its tenure. Sherpas or working groups invite them to a segment of their meetings so they can discuss demands and expectations (see subsection 2.3.4).

Almost all G7 summit documents mention only what members have agreed (see *consensus*, under subsection 4.1.2). Cases of *open disagreement* in the G7 first emerged in the early 2000s when there were different views on the Kyoto Protocol. The documents produced at the Genoa summit in 2001 and the Evian summit in 2003 illustrate that disagreement. A similar disagreement appeared on the Paris Climate Agreement. The Leaders' statements produced at the Taormina summit in 2017 and at the Charlevoix summit in 2018 make that clear. But this remains the exception. Apart from climate issues, the G7 does not document disagreements.

2.2 What is accountability and what is it good for?

Before we consider who in the G7 is accountable to whom, as well as what must be answered for, and how, we must first understand the general concept and the uses of accountability. The general concept developed here applies to all forms of accountability, internal and external.

Accountability is part of every *principal-agent relationship*. The agent owes the principal a certain behaviour. Reports on the agent's behaviour can help the principal make sure that the agent does what s/he is supposed to do. That is why accountability is sometimes described as a relationship containing a duty to report and a right to impose sanctions (Keohane, 2002, p. 12). But this description is minimal: it covers different forms of accountability (Schedler, 1999, p. 15) without saying how they relate to each other. In one group of cases (typically external accountability), the main point of reporting is to *give reasons* for a behaviour the principal already knows. In another (typically internal accountability), it is to *generate information* the principal would otherwise miss because he has left the work to the agent. Some of the debate on the G7 and accountability is confused because the outside world expects the G7 to give reasons, while G7 processes are designed to extract information. This confusion extends to the question of who the audience of G7 accountability reports is supposed to be? Outside stakeholders want material to make a

judgment on G7 legitimacy, but G7 insiders produce information about implementation experience. Both speak of accountability but may mean different things.

2.2.1 Four forms of accountability

To sort this out, it is useful to forget for a moment about why the agent’s behaviour must be reported on and focus on two formal questions instead: Who works for whom? And who reports to whom? Because agency relationships use *delegation*, work and reports generally pass from one person to another. This leaves us with four possibilities:

Table 1: Forms of accountability		
All cases of accountability concern reporting (giving information) about working		
A	Recipient of reports	Assumption 1: Delegation of reporting: A is never B (that is, A receives reports from another person or institution (B or D)).
B	Provider of reports	
C	Recipient of work	Assumption 2: Delegation of work: C is never D (that is, C receives work from another person or institution (B or D)).
D	Provider of work	
	A = C (A receives both work and reports) – the principal manages the agent’s work	A ≠ C (A receives reports and C receives work) – the principal judges the agent’s work
B = D (B provides work and report)	<u>Internal accountability</u> – purpose of report: to overcome the agent’s behaviour bias. The agent produces information that the principal can use to manage the agent’s behaviour. Question: How can the agent align his/her behaviour to the principal’s will?	<u>External accountability</u> (Legitimacy model) – purpose of report: transparency, “give reasons”. The agent produces information to explain and justify his/her work to a third party. Question: Is the agent (still) entitled to his/her role?
B ≠ D (B provides reports and D provides work)	<u>Expert assessment</u> – purpose of report: to overcome the agent’s judgment bias. A third party produces “independent” information that the principal can use to check the agent’s behaviour against general standards of knowledge and expertise. Question: Is the agent’s performance “state of the art”?	<u>Peer review</u> – purpose of report: to compare performance. An agent produces information about the work of another agent, so third parties can place this work in relation to other work on the same issue. Question: How does this agent perform compared to others?
Source: Author		

The two vertical blocks differ in whom accountability is owed to. The left block shows forms of *internal accountability*, the right block forms of *external accountability*. While internal accountability is managerial in nature, external accountability is more political. This is because the principal’s role is different. In internal accountability s/he has leverage over

the agent's operations and can decide what is to be done, so the agency relationship only has to determine how to do it. In external accountability, the principal has weak or no leverage over the agent's operations, so the agency relationship must also grapple with what to do. The substance of what an agent does is more easily accessible to normative and political judgment than the procedure s/he adheres to.

The two horizontal blocks distinguish whether the agent reports on his/her own work or reporting is done by a third party. While the first relies on the agent's own knowledge, the second builds on independent information. Wherever the principal wants to extract the special experience the agent has gained from implementing the specific task s/he has set, the principal will turn to her/his agent's reports. If the principal needs more general knowledge or wants to protect her-/himself from the bias of an agent reporting about his own work, s/he will turn to independent sources. Each form of accountability is designed to solve a specific type of problem:

- In the upper lefthand case, the agent has privileged access to information about her/his work and incentives that might not align with the principal's. Here, accountability serves to give the principal the information s/he needs to effectively guide the agent's work and control for adverse incentives. G7 accountability for development-related commitments (see subsections 2.3.1, 4) falls in this category.
- In the lower lefthand case, neither the agent nor the principal has all the information needed to make a judgment about the agent's work. This is why the principal relies on information provided by a third party. The G7 often uses international organisations for that purpose (see subsection 2.3.2).
- In the upper righthand case, the agent works for one principal and reports to another. The reason for this asymmetry is that the work of the agent affects not only its recipient but also a whole range of outside stakeholders. The agent here reports in order to give reasons for those external effects. G7 dialogue with outside stakeholders (compare subsections 2.3.3 and 2.3.4) serves that purpose.
- In the lower righthand case, an independent party reports on work an agent does for someone else. The point here is typically to compare how different agents perform. A third party is needed if the recipient of the work does not have the information required to judge the performance of all relevant agents.

The standards against which the agent's work is judged can be very different because they depend on the principal's needs. In consequence, the text examples shown in Table 1 only represent possibilities that can (but do not have to) emerge within each category. There is no necessity to them.

International cooperation has practical examples of every form:

- G7 accountability processes are largely designed as a case of *internal accountability* where the agent reports on his own work (upper lefthand case): G7 national administrations work together to carry out mandates set by G7 Leaders and report to them what they have done. Accountability practice in the G20 (Hilbrich & Schwab, 2018) roughly follows the same pattern.

- The work of the Financial Action Task Force (FATF) is essentially an *expert assessment* (lower lefthand case). The 39 FATF member countries ask the FATF secretariat how they perform against their agreed disciplines against money laundering and terrorist financing (AML/CFT). While member states work to comply, the FATF secretariat reports. Subsection 2.3.2 presents cases where the G7 has mandated expert assessments.
- The Annual Reports published by the multilateral development banks (MDBs) are examples of external accountability, where the agent reports to someone he does not work for (upper righthand case). The MDB's work is mandated by shareholders and directed at developing and emerging economies. But their Annual Reports are intended for a larger audience, regardless of whether it is located in a shareholder or recipient country.
- The OECD DAC (Development Assistance Committee of the Organisation for Economic Cooperation and Development) peer review is a prominent example of a peer review where an independent party reports on the work of an agent for a principal (lower righthand case). It looks at how a donor country engages in development cooperation. This work review is intended to benefit developing countries. But other donors assess it.

Knowing who plays what role in the game of giving and receiving work and reports tells us much about the concrete job accountability has to do. In every case, the principal uses accountability to make sure the agent does what he is supposed to do. But the *benchmarks* to decide whether this is the case and the *kinds of information* the principal needs to make that judgment are very different. That is why every analytical approach to a concrete practice of accountability must be specific. It must identify the benchmark that the principal applies as well as the information s/he has requested. Once these questions are resolved, a critical assessment becomes possible: one can then ask whether a certain accountability process delivers on its stated goals (see subsection 4.3) and whether the stated goals are appropriate when outside stakeholders expect different things.

The remainder of this section looks at bodies of research about what makes accountability work. Literature on *legitimacy* (see subsection 2.2.2) shows how international groups and organisations must *justify* their work to the outside world. Literature on *contract theory* (that is, subsection 2.2.3) shows how the principal can get the information s/he needs to manage the agent's behaviour. Each strand of thought leads to requirements that the G7 must meet to protect its legitimacy and achieve its goals.

2.2.2 Elements of legitimacy

An important body of literature describes external accountability as a condition of legitimacy (Schedler, 1999; Keohane, 2002; Buchanan & Keohane, 2006). Why is this literature relevant to the G7? The G7 must show it is legitimate because it projects power and attracts expectations (*demands*). It also depends on stakeholders for the effectiveness of many of its actions, such as the pursuit of a tax agenda that the G7 wants as many countries as possible to sign up to (for instance, the OECD-implemented project against base erosion and profit shifting, BEPS). Those stakeholders will only go along with the idea if they believe that G7 action is basically legitimate. In this perspective, outside stakeholders are the principal and the G7 is the agent. Accountability can serve legitimacy even if the outside stakeholders that the G7 is accountable to have no real leverage over it. In these cases, the

G7 can still answer for the external effects of its action (see subsection 2.2.1) and it has in fact engaged in dialogues to that effect (see subsections 2.3.3 and 2.3.4).

How are legitimacy and accountability related? The literature makes the case that external accountability serves to check whether the agent deserves his power. The principal will test the agent's behaviour against a normative standard in order to find that out. This test works because the agent must *justify* her/his behaviour in reports to the principal. Knowing the agent's reasons, the principal can then make a judgment on whether s/he still deserves her/his power. The principal's judgment carries sanctions (public exposure, pressure, punishment, removal) intended to reinforce the normative standard.

This concept is visibly rooted in the democratic accountability found in nation states. It is not easily transferred to international institutions because member countries disagree on normative standards (Buchanan & Keohane, 2006, p. 13) and because outside stakeholders cannot effectively sanction international entities (Keohane, 2002, p. 13). Literature therefore concentrates on minimum requirements (such as responsiveness and compliance with basic normative standards) that make them legitimate.

There is another caveat, which is specific to the G7: *discontinuity* in the G7 means that it cannot build legitimacy the way a formal international organisation can. The G7 has no permanent process to engage with civil society or third countries. Instead, each presidency is free to design its own kind of G7 dialogue with non-governmental stakeholders (*engagement groups*) and other external actors. Formats for dialogue with non-G7 countries such as the Heiligendamm Process (Leininger, 2009) or the Africa Partnership Forum (Hayford & Kloke-Lesch, 2013) have all been temporary. Permanent opportunities for affected people to raise grievances and obtain relief (another way of strengthening legitimacy) are only available in formal international organisations.

But there are other sources of legitimacy that the G7 can access more easily: clarity of purpose, consistency between words and behaviour, transparency in action, and effectiveness in solving problems (Buchanan & Keohane, 2006, p. 422). Here are some questions to test whether G7 procedures meet these standards:

- Are public G7 statements clear to the point that an outside stakeholder can judge for her- or himself whether G7 behaviour is consistent with its words (*credibility, integrity*)?
- Does the G7 define the *purpose* it wants to achieve with an announced action?
- Does the G7 disclose what it has done as a group to comply with previous statements (*transparency*)?

These requirements are minimal. They do not include G7 *responsiveness* to outside demands or compliance with material *normative standards* (such as human rights) that clearly bind G7 behaviour. To be sure, G7 legitimacy depends on much more than the three questions above. External accountability also means holding the G7 to account for its compliance with normative standards. The reason why I still apply only three minimal legitimacy criteria has to do with the limited purpose of this paper. It is to see how aspects of legitimacy can be integrated into the current G7 accountability practice. Subsection 3.1 shows why this practice can meet some – but not all – requirements of G7 legitimacy. The G7 has to answer the larger part of the legitimacy question outside its formal accountability

process, essentially by engaging with outside stakeholders and leading the implementation of the 2030 Agenda.

The three criteria are meant to offer a simpler test: Does the G7 open itself to external probes about whether it does what it says and achieves what it wants? All three criteria can be met within the current G7 accountability framework. The following subsection (2.2.3) presents more criteria. Subsection 4.3 looks at whether the G7 meets them in practice.

2.2.3 Managing delegated behaviour

Contract theory is about how the principal can reach his/her goals although the agent has different incentives and better information. It offers ways to maximise the benefit or minimise the cost of delegation. This literature (Schmidt, 2017) is relevant to the G7 because the G7 has no implementation structure and must *delegate* to actors outside the G7 process. The group of G7 Leaders can delegate a task to an international organisation or to G7 national administrations. I focus on the latter case because G7 accountability is about what the G7 has done as a group. In this case, the group of G7 Leaders asks G7 national administrations to work together and carry out a task. In this view, the G7 Leaders group is the principal and G7 national administrations, acting in cooperation, are the agent. G7 national administrations do most of their work separately. G7 commitments ask them to implement a mandate together (*discipline*). This creates the group of G7 national administrations that acts as the agent for the G7.

The findings of contract theory offer guidance on how to design the relationship between principal and agent and on how to manage it once it exists. Here is some advice on relationship design:

- If the principal is already committed to an agent, s/he does not have the leverage that comes with being able to choose among different ones. The principal cannot use it to make sure that the agent has the right incentives and bears a fair amount of risk. As a result, useful measures may not be taken (*holdup problem*, Schmidt, 2017, p. 501). For the G7, this means that Leaders should design a task *before* they decide who will implement it. As long as they can choose to delegate either to their national administrations or to an international organisation, they can use that leverage to set the right incentives. This is necessary because administrations want to save effort. They may present as implementation of a G7 task that they would actually have done anyway. What looks like implementation in a report might not in fact be an additional effort. This can lead to the G7 over-reporting its action. Subsections 4.2.2 and 4.2.3 show how G7 practice imperfectly controls that risk.
- The agent may have an incentive to withhold information from the principal in order to make the agency relationship more favourable to him/her. This is relevant to the G7 because national administrations and international organisations are in a better position to tell easy tasks from difficult ones (see subsection 4.1.1). They might also be tempted to deliver only the minimum performance needed for political messaging without making a substantial difference to the purpose Leaders have in mind. This is why G7 Leaders should inform themselves about the relationship between efforts and outcomes before they design a task (See Section 5). They should also make sure that success on the

substance of that task will pay off for the administrations or international organisations who implement it, even if this success is hard to communicate politically. At the very least, G7 tasks should not be unfunded mandates.

Two further points relate to the way the principal should manage his/her relationship with the agent. Both have to do with the constraints that every principal faces. S/he must operate through the agent and leave her/him discretion on how to approach the task (*delegation*). The agent almost never has full control over whether s/he reaches the objective set by the principal (*outside interference*). And the principal has only *imperfect information* on the drivers and obstacles on the way to his/her objective. Contract theory offers some useful workarounds:

- When managing the agent's behaviour, the principal should focus on what the agent can do to achieve the principal's objective, not on what is beyond the agent's control. To put it more precisely, the principal should aim his controlling measures only at things that the agent can influence (as opposed to outside interference) and have an effect on the objective (Schmidt, 2017, p. 493). The agent should not be able to manipulate the instruments that the principal uses to manage his or her behaviour. What does this mean in G7 terms? Leaders should make sure they get information that tells them to what extent an outcome is due to their administrations' efforts (*attribution*) and what is a result of outside effects (*interference*). This distinction is important because Leaders only have leverage over the actions of their national administrations, not over outside interference. One way to make that distinction is comparing the agent's performance to what others have achieved in similar situations (*relative performance evaluation*). Similarities in performance are probably due to a situation everyone faces; differences may signal the extra effort a particular agent has made. Leaders should also make sure that their administrations are not free to decide what set of facts they want to report (prevent manipulation). Subsection 4.2 shows how far the G7 meets these requirements.
- The principal should have an idea about how the agent's actions will achieve the objective (*theory of change*). But with imperfect information, this idea can be wrong. For that reason, the principal should generate information that enables him/her to test and (if needed) correct his or her original theory of change. Depending on this information, s/he can see what the best way forward is: doing more of the same; trying different actions; or even giving up on the objective. This means that the G7 should define a theory of change whenever it sets a task. It should make administrations report on whether its theory of change actually works. The learning function inherent in internal accountability can only operate if the theory of change underlying a commitment is regularly tested against implementation experience. Subsection 4.3.1 shows how G7 practice performs on that count.

This leads us to some more test questions for the G7 process:

- Do G7 Leaders target actions that have an effect on their stated objective? Do they discount outside interference?
- Do G7 Leaders make requirements on the quality and effectiveness of their administrations' actions and do they make sure compliance with these requirements is reported on?
- Do G7 Leaders define a theory of change?

- Do G7 Leaders make sure they get the information necessary to test and (if needed) adjust their theory of change?

These questions are formal in the sense that they leave open what substantive task the G7 should set, who it chooses to be its agent, and how it designs the agency relationship. This has to do with the limited purpose of my paper. I want to develop options to change G7 procedure in a way that enhances legitimacy and effectiveness. Accordingly, the recommendations offered in Section 5 are not about what the G7 should do but about *how* it can do it. This limitation makes sure that my recommendations can be implemented without changing the current G7 accountability framework.

2.3 How the G7 has used accountability instruments

This subsection looks at the various ways the G7 has applied instruments of accountability to its work. While it has used accountability mainly to manage tasks it has *delegated* to national administrations (internal accountability), outside *demands* have also played a role (external accountability). Due to its *discontinuous* nature, the G7 has never developed a permanent accountability framework covering all of its work but rather set up separate, temporary accountability processes for each task, normally a one-off reporting mandate. Leaders often receive a number of separate, issue-specific reports at one summit (see subsection 3.4). The permanent accountability framework for all development-related G7 commitments is an exception. This framework can deliver a *continuous* view of how the G7 performs on the goals it has set itself in this area. But it does not cover policy areas not deemed to be development-related. The operation of this framework suffers from the *disconnect* between backward-looking reports and forward-looking Leaders (compare subsection 2.1). Leaders are so concentrated on creating new policies that little attention is paid to how previous commitments have been implemented. This bias affects all G7 proceedings, from the working group level to the sherpas because the whole G7 system is Leader-driven in nature (see subsection 2.1). That is why the G7 learns less than it could from its accountability products (see subsection 4.3).

In the remainder of this section (2.3.1), I will show some examples of the different task-specific accountability processes the G7 has set up. Then I will describe the permanent accountability framework the G7 has established for its work on development-related commitments (2.3.2).

2.3.1 Reporting mandates to the group of G7 national administrations

From early on, G7 Leaders have asked their administrations to collectively report to them. Normally, the idea is to monitor how they have *implemented* a certain decision that Leaders have made. Here are two examples taken from summit documents:

As a follow up to our discussions, we agree to ask ministers to meet in France before our next summit to review the progress made in job creation and consider how best to increase employment in all of our countries. (Halifax, 1995, p. 8)

We urge the Lyon Group [...] to intensify its ongoing work [on drugs and crime prevention] and ask our ministers to report back to our next summit on progress on the action plan on high tech crime, the steps taken against money laundering and the joint action on trafficking in human beings. (Birmingham, 1998, p. 22)

Less often, the idea is to *prepare* a decision Leaders want to make: “We ask [our personal representatives] to report to us by the next summit on possible ways to improve [our approach on food security]” (Cologne, 1999, p. 43). In these cases, G7 Leaders are aiming at internal accountability: the national administrations of G7 countries are supposed to work for, and report to, them. The group of G7 Leaders is the principal and G7 national administrations acting collectively are the agent. The findings of contract theory (see subsection 2.2.3) apply to this kind of situation.

2.3.2 Reporting mandates to international organisations

In another group of cases, the G7 asks international organisations to implement a task or assess member states’ performances. In the latter case, the idea is to generate an independent review of how G7 members perform and/or whether they comply with an agreed discipline. This approach is better positioned to produce objective assessments on individual G7 member countries than reports drafted and agreed upon by G7 administrations. One example is a mandate issued at the L’Aquila summit in 2009: “We ask the WTO together with other international organizations, within their respective mandates, to continue to monitor the situation and to report publicly on the adherence to these commitments [to refrain from protectionist measures] on a quarterly basis” (L’Aquila, 2009, p. 46). At the Deauville summit in 2011, the G8 asked the OECD to develop a framework to measure countries’ innovation performance and announced that every G8 country would let itself be measured:

In order to enhance policy efficiency and effectiveness, we also invite the OECD to develop, in a fully inclusive, open and accountable way, in cooperation with relevant international organizations, measurements of innovation performance, focusing on concrete impacts on growth and jobs rather than inputs and investigating the systemic relationship between indicators. All G8 countries will participate. (Deauville, 2011, p. 27)

The G7 has also asked international organisations to provide it with a factual basis to prepare a decision it wants to make: “The relevant international organizations will be asked to complete reports on the state of the world’s forests by 1990” (Paris, 1989, p. 42). All these cases fall into the category of expert assessments (compare subsection 2.2.1). The G7 is the principal and wants an independent agent with expert knowledge to provide it with an objective assessment. The G7 then intends to use this to judge how its members perform or what steps to take.

2.3.3 Dialogue with third countries

The G7 has used different forms of accountability for its relationship with the outside world, notably non-G7 countries and civil society. To understand these forms, it is best to turn to the literature on external accountability and legitimacy (see subsection 2.2.2).

Since 1976, the G7 has acknowledged its interdependence with third countries. It says that reaching G7 goals also depends on them: “The formulation of such policies [on growth, employment, inflation control], in the context of growing interdependence, is not possible without taking into account the course of economic activity in other countries” (San Juan, 1976). And it recognises its responsibility to third countries, developing countries in particular: “We consider it essential to take into account the interests of other nations. And this is most particularly true with respect to the developing countries of the world” (San Juan, 1976); “We shall pay regard to the effects of our policies on the developing countries” (Munich, 1992, p. 18).

The interest of reaching goals it cannot secure by itself and a sense of responsibility towards the world have led the G7 to engage since 2000 in a growing number of outreach exercises and theme-specific dialogues with third countries. The Heiligendamm Process and the G8 Africa Partnership are the most prominent examples (Leininger, 2009; Hayford & Kloke-Lesch, 2013). Such processes can build external accountability if non-G7 countries can use them to voice expectations and effectively question G7 policies. In fact, the G7 offered only limited opportunities to do this. Dialogues were focused on themes set by the G7 presidency. It was always uncertain whether the next G7 presidency would take up issues raised by non-G7 countries (*discontinuity*) and there was reluctance on both sides to disclose failings and discuss ways to improve.

2.3.4 Dialogue with civil society

The G7 also conducts dialogues with civil society, mainly through the G7 engagement groups (see subsection 2.1), for practical and normative reasons. The G7 acknowledges that at least some of its policies will only work with public support: “We will now seek parliamentary and public support for these measures” (Bonn, 1978, p. 31). Sometimes building trust is a G7 goal in itself: “All these meetings demonstrate our commitment to a process of dialogue aimed at strengthening public confidence in food safety” (Genoa, 2001, p. 31). Another motive is to respond to concerns that societies of G7 countries have at the time of the summit, most markedly on the effects of globalisation and the ability to control them: “The challenge is to seize the opportunities globalization affords while addressing its risks to respond to concerns about a lack of control over its effects” (Cologne, 1999, p. 3); “We will act to restore confidence in the fairness and effectiveness of our international tax rules and practices” (Lough Erne, 2013, p. 23).

Depending on how the presidency designs it, G7 dialogue with engagement groups can serve external accountability. Civil society can use it to voice concerns and make suggestions. The G7 presidency is then forced to justify its programme and gets a picture of how it is perceived. Having said that, this process also has limitations: it cannot be a dialogue between the sum of all societies in G7 countries and the G7 as an international group. Collective G7 engagement groups are less representative than the peoples and parliaments who hold each G7 government to account. That is why justifying policies on globalisation largely remains the job of national governments. Often it is only the G7 presidency, and not the G7 as a group, which participates in dialogues with engagement groups while these groups sometimes use dialogues to lobby for pet issues rather than to hold the G7 to account. Often engagement groups and the G7 prefer to discuss ideas for the next summit rather than G7 performance on long-term issues (*discontinuity*).

3 The accountability framework for G7 commitments on sustainable development

Every year, G7 national administrations collectively issue a report about how the G7 has delivered on its development-related commitments. This routine, a case of a reporting mandate to the group of G7 national administrations (see subsection 2.3.1), is based on the *accountability framework* that the G7 established in 2010. This framework governs the accountability reports that the G7 issues on the implementation of development-related summit commitments. According to it, a G7 summit commitment will automatically be reported on if every member agrees that it is related to development. So, whenever G7 Leaders make a “development-related” commitment, they do not have to make an additional explicit decision on how to follow it up. They rely on the process laid down in the permanent accountability framework. Reporting under this framework can cover various different themes or sectors, as long as the underlying benchmarks are development-related G7 commitments.

3.1 Scope and limitations

The permanence established in this framework is an exception to the usual G7 practice. If the G7 wants to follow up on a task, it normally sets up a procedure that fits with that task and ends with it (*discontinuity*, see subsection 2.3). That is why the permanent G7 accountability framework cannot automatically extend to every G7 commitment, although Leaders are free to apply it wherever they want. For the time being, a G7 summit commitment is only followed up through the permanent accountability framework if it is deemed “development-related” (this is the wording of the framework) and if the AWG agrees that it qualifies as eligible (compare *How the AWG recognises commitments*, under subsection 4.1.2).

Why has the G7 limited its permanent accountability framework to “development-related” summit commitments? One reason is probably historical: The permanent accountability framework was introduced to systematise the follow-up of a growing number of thematically related summit commitments and this particular problem first emerged with development-related commitments (see subsection 3.4). Another reason may have to do with incentives: One can make the case that, while the G7 can be relied upon to do what is in its own interest, a special level of control is required whenever it promises to do something for third countries. The growing importance of G7 dialogue with developing countries and civil society organisations engaged in development (see subsections 2.3.3 and 2.3.4) may also have played a role because development issues played a dominant role in those dialogues.

Given the comprehensive reach of the 2030 Agenda, there are strong reasons to extend the scope of the permanent accountability framework beyond “development-related” commitments. However, for the time being, there is no consensus in the G7 to do that. Discussion in the AWG concentrates instead on whether certain summit commitments that go beyond traditional development cooperation can still be “development-related”. The Elmau commitment on sustainable global value chains has cleared that test, even though it requires important changes inside G7 markets.

The limited scope explains what the G7 accountability framework cannot resolve:

- It cannot produce reports about how the G7 has performed generally against the goals of the 2030 Agenda, because it is explicitly limited to what the G7 has agreed on at one of its own summits. The question whether the G7 has made the “right” commitments is beyond the framework’s remit.
- It also cannot make sure that reports say much about policy coherence in the G7. Reporting practice allows for discretion (see subsection 4.2), but the framework only asks for reports on the implementation of “development-related” (not other) commitments. Part of the AWG’s job is to make sure reporting does not exceed this mandate.
- The framework does not produce reports about the independent performance of each G7 member, because it is limited to what the G7 does as a group (*discipline*). If G7 accountability reports sometimes present what each G7 member has done, it is to show contributions to a *joint* effort rather than individual behaviour. For specific assessments of individual country performance, the G7 usually turns to international organisations (see subsection 2.3.2).
- Finally, the framework makes G7 administrations report about their own work, so its potential to produce critical assessments is limited. While there are rules to control this bias (see subsection 4.2), reports under the accountability framework can never be as disinterested as assessments from outside the G7. Reports by academic institutions (such as the University of Toronto) or civil society may have their own limitations but are at least free from the temptation to present G7 work in the best light. To the extent that their governance allows them to voice views that differ from their member states, this is also true for international organisations.

3.2 Features

The framework is designed for self-control. G7 Leaders ask their national administrations to collectively report how they have implemented a specified set of G7 commitments. Every three years, a *comprehensive accountability report (CAR)* covers *all* development-related G7 commitments. In the intervening years, the *annual progress report (APR)* is on one or a number of these commitments, according to the presidency’s choice. The framework essentially creates a case of *internal accountability* (see subsection 2.2.3). This is most visible in the AWG. Leaders established this group in 2009 and asked it to design and carry out the accountability framework. The AWG is a group of mid-level civil servants appointed by each G7 member government. As in all G7 working groups (compare subsection 2.1), the member from the country who has the G7 presidency chairs the AWG. This member produces a draft of the accountability report and leads discussions that result in a version of the report every AWG member agrees to. The AWG chairperson then sends this text to his or her sherpa, who submits it to the group of G7 sherpas. Once agreed by G7 sherpas, the accountability report is published and submitted to Leaders for their endorsement at the summit. Another task for the AWG chairperson is to build AWG consensus on which of the outcomes of the previous summit should be followed up according to the accountability framework and what methods should guide the reporting on these outcomes (see *How the AWG recognises commitments*, under subsections 4.1.2, and 4.1.3). In this process, three features stand out:

- The AWG draws much of the knowledge and skills it needs to produce accountability reports from its members. The *collective implementation experience* of G7 national administrations is a source no other form of reporting on G7 work can equal. This is an advantage of internal accountability.
- G7 accountability reports have to clear three levels of consensus: in the AWG, in the sherpa group and among Leaders. This process is designed to ensure that *G7 members own the contents* of accountability reports.
- The process is *hierarchical* in nature. Reporting is limited to what Leaders have asked for, produced by their subordinates and specifically directed to Leaders.

3.3 Stated purposes

While the basic design of the G7 accountability framework follows patterns of *internal accountability*, G7 reports also serve purposes of *external accountability* (see subsection 2.2.2). A G7 accountability report is supposed to be published before each summit and the presidency usually tries to reach a large audience for it. G7 summit documents and accountability reports themselves have stated how they want the process to contribute to external accountability. Here are some of the motives:

- Transparency (“To improve transparency and effectiveness, we decide to strengthen our accountability [...]” (L’Aquila, 2009, p. 98));
- the ability of outside stakeholders to hold the G7 to account (“[...] the report seeks to [...] provide information which stakeholders can use to hold G8 Leaders to account for their promises” (G8, 2013, p. 4));
- credibility (“Accountability and transparency remain core principles for the G7 in order to maintain the credibility and effectiveness of decisions of G7 Leaders” (G7, 2016, 1)); and
- the communication of results (“In its third year, the annual G8 Accountability Report continues to evolve as a means for capturing the progress that G8 members have made toward their commitments and for communicating the results of these efforts” (G8, 2012, p. 65)).

But the same set of documents also stresses *internal accountability* (see also subsection 2.2.3). According to them, G7 accountability reports are instruments to:

- Monitor progress (“We are determined to [...] adopt] a full and comprehensive accountability mechanism by 2010 to monitor progress and strengthen the effectiveness of our action” (L’Aquila, 2009, p. 3));
- draw lessons (“This report offers the opportunity to capture important lessons about G8 collective work [...]” (G8, 2011, p. 2)); and
- make informed decisions about changing course. (For example, the Lough Erne Accountability Report (G8, 2013, p. 24) considers how the G8 should adapt the

implementation of its commitments on economic development to the fact that its partner countries' "economies and trade structures are growing and changing".)

Apparently the G7 does not ask itself how far the same process can serve external *and* internal accountability. While testable commitments and consistency between actions and words are important for both, there can be trade-offs between the two forms of accountability (Keohane, 2002, p. 20). Leaders can use internal accountability to learn from implementation practice. But naturally they will hesitate to disclose bad practice out of concern for their reputation. In the process of producing a public report fit for external accountability, the purpose of learning from mistakes is easily lost. Section 5 suggests a way to solve that problem.

3.4 History

The G7 accountability framework emerged because, by 2007, many reporting mandates on different tasks had accumulated to the point of confusion while Leaders were still asking for more reports. In its discontinuous, task-specific way (see subsection 2.3), the G7 had requested more internal accountability products than it could easily process at one summit. That is why the first ("preliminary") G8 accountability report published in 2009 (G8, 2009) was mainly an attempt at reducing complexity. It compiled and synthesised previously mandated reports on the implementation of health, water, education and food security commitments. These separate reports had been drafted and agreed upon by four G7 working groups (see subsection 2.1) in charge of the respective sectors. But the preliminary accountability report also offered an overview covering all four sectors. To describe how commitments were implemented, the preliminary report largely enumerated individual actions taken by G8 members, so it was hard to see what the G8 had done as a group (see subsection 3.1).

At their 2009 summit, Leaders established the AWG (see subsection 3.2) and asked it to draft an accountability framework. Leaders endorsed the draft framework (G8, 2010, p. 76) along with the first comprehensive accountability report in 2010. The new process made sure that reports followed a uniform set of methods (see subsection 4.2), even if they covered commitments in various different sectors or themes. Drafting and negotiating reports was now in the hands of the AWG rather than sector working groups (see subsection 4.1.3). A more formal reporting methodology helped to identify what the G7 did as a group as opposed to what each member did independently (although this still does not always work; refer also to subsection 4.2.3). The annual reporting cycle and the stability of methods allowed following G7 performance in time.

Since 2018 the United States has made retroactive reservations on a number of development-related G7 commitments that were originally made in consensus (G7, 2018, footnotes to pp. 80-93). The United States has not endorsed the 2019 Biarritz Progress Report (G7, 2019a, p. 2). Between 2017 and 2019 the AWG was never able to agree on the accountability report in time for it to be published before the G7 summit, as provided for in the accountability framework. All of this shows that the G7 accountability process must first be safeguarded before it can be improved.

4 How the G7 reviews its commitments on development (process)

This part presents how the G7 practices accountability for development-related commitments, starting with commitments because they are the main reference points. Subsection 4.1 follows the life of a commitment from its inception to the moment when it is not reported on anymore (“sunset” in AWG parlance). Subsection 4.2 considers the interpretation of commitments and explains how the AWG sets rules to guide reporting (“methodology”). Subsection 4.3 uses the test questions developed in subsections 2.2.2 and 2.2.3 to ask whether this structure uses the potential inherent in internal and external accountability. The answers will be the basis for the recommendations offered in Section 5.

4.1 The life of a G7 commitment

This section will follow the stages in the life of a G7 commitment. How does a commitment come about (4.1.1)? What has to happen for it to be reported on (4.1.2)? What kind of groundwork must be done to make sure reporting on development-related G7 commitments follows transparent rules (4.1.3)? When is reporting on a commitment no longer necessary (4.1.4)?

4.1.1 How the G7 comes to its commitments

Like most summit outcomes, commitments go back to something the G7 presidency has proposed. Before assuming the presidency, every G7 member country prepares a programme of what it wants to achieve at the summit it will be hosting. This programme is rooted in things that are specific to the presiding G7 country. The current policy agenda of its government, the strengths and weaknesses of its administration and domestic pressures all play a role in the formation of a G7 presidency programme. But the presidency can only be successful if it reaches consensus by convincing other members. The structures described in subsection 2.1 are one way to do that, because they are all designed to gradually build consensus. Another way is to host public events on topics of the presidency programme in order to mobilise lobbies and raise pressure. Informal meetings among G7 representatives at every level help to understand hesitation, provide arguments and develop compromises. I have made a case study of such a process in an earlier paper (Fischer, 2018).

Seen from the angle of the G7 as the principal to an agent implementing a commitment, this process contains risks. We know some of them from contract theory (see subsection 2.2.3):

- If a G7 presidency uses its programme mainly to pursue domestic goals, other G7 members might just agree to go along with this at the summit, without actually being willing to implement it (*thin consensus*).
- If the administration of the presiding G7 member gets to suggest a commitment, it might be tempted to propose actions it would have undertaken anyway, even without the G7 (*no additionality*).
- If G7 administrations together draft a commitment they know they will be asked to implement, chances are it will not be ambitious (*adverse selection*). However, this same

fact can also strengthen compliance because administrations feel a special responsibility to implement what they have drafted and negotiated themselves (*drafters' ownership*, Chayes & Chayes, 1993). In the G7, this positive effect is visible in places like the FSWG.

- G7 administrations know that, the following year, the G7 will focus on different goals (*discontinuity*). This could lead them to design commitments and rules that do not require a strong follow-up.

Administrations and Leaders do not draw on accountability products when they design and negotiate new commitments. There is a *disconnect* between backward-looking implementation reports and forward-looking summit agendas (see subsection 2.1). The accountability process provides options to close that gap. According to its Terms of Reference, the AWG can offer sherpas advice on the design of commitments. Some G7 accountability reports (G8, 2010, p. 74; G8, 2013, p. 120) contain specific advice on how to make commitments more measurable. But these options have never been acted upon.

4.1.2 When is a summit outcome fit for accountability?

The G7 only reports on the follow-up of a development-related summit outcome if all Leaders have agreed on it (*consensus*, see below) and reporting adds value to its implementation (*mandate*, see below). The AWG makes sure that only commitments that fit these tests can enter the accountability process (*How the AWG recognises commitments*, see below).

Consensus

Leaders normally signal their consensus through words like “we agree”, “we decide” or “we will” in the summit document. If they agree on the contents of an annex to the summit document, they usually write “we endorse” and make a specific reference to that annex. Leaders’ statements and the Chair’s summaries (refer to subsection 2.1) can equally document consensus. When every Leader is present at the summit and no one voices an objection, any type of document can show agreement. Proceedings at and after the G7 summit in Charlevoix 2018 were special in this regard. President Trump agreed to the Leaders’ statement while attending the summit and then revoked his agreement shortly after leaving it. His reasons were apparently not related to the contents of the Leaders’ statement.

At what point in time must there be consensus to create a viable G7 commitment? This moment has to be the end of the summit. If commitments mean anything, they must be binding. All G7 members who have agreed to a commitment at the summit have a responsibility to comply with it afterwards. If a G7 member changes his mind after a summit (as is bound to happen in democratic systems), this does not relieve his administration of the responsibility of continuing to implement the agreements reached there until they are suspended or changed by consensus. Without such consensus, previous agreements remain binding on all G7 members. This is why the explicit reservations that the United States have made to a number of G7 commitments only mean that they do not want their national administration to comply with them anymore and that the G7 cannot count on its active contribution to every aspect of those commitments. The commitments still bind all G7 administrations.

But the Charlevoix case is different. Almost immediately after the summit, President Trump distanced himself from its entire outcome. This indicates that he never took ownership of it in the first place so that, at the end of the summit, there really was no consensus. Through his immediate and somewhat casual distancing, President Trump sent a signal that was stronger than the existence of a formally agreed document. Probably as a consequence, the 2019 Biarritz Leaders' declaration (G7, 2019b) closed the gap between prefabricated texts and Leaders' ownership: in this document Leaders credibly state that they have drafted and negotiated it themselves.

Mandate

Accountability adds value if there is clarity about who is supposed to do what (see subsections 2.2.2 and 2.2.3). The G7 asks its national administrations to undertake a development-related job together. Accountability reports can only monitor implementation if the G7 has defined what this job is. Otherwise it is impossible to tell good from bad implementation (Chayes & Chayes, 1993, p. 183). This is just as true for mandates that the G7 gives to international organisations. However, this case is irrelevant for G7 accountability because this process must show what the G7 (as opposed to an international organisation) has done. An example of sorting out the exact division of labour between the G7 and the OECD is the G7 tax agenda (compare subsection 4.1.4).

The mandate must be directed at G7 national administrations acting as a group. Actions that they take independently cannot be attributed to the G7. Given that G7 accountability serves to answer the question of what the G7 has done, independent actions cannot be part of it.

How the AWG recognises commitments

After each summit, the presidency makes a judgment about which summit outcomes need a follow-up through the accountability process. Once it has made that choice, it submits the relevant pieces of the summit document text to the AWG and proposes including them in the *list of active commitments* (G7 2019a, p. 6-7). This list presents all G7 development-related commitments that are currently followed up through the accountability process. The AWG drafts and updates it by consensus. Part of this update is deciding whether the pieces of summit language submitted by the presidency amount to eligible commitments. The AWG checks this against the criteria already mentioned:

- Can the submitted piece of language be attributed to Leaders?
- Have all Leaders agreed to it at the end of the summit?
- Does it ask G7 national administrations to carry out a specified task together?
- Is this task development-related?

If all AWG members agree that a piece of summit language clears these four tests, the AWG decides to add it to the list of commitments. As a consequence, it will continually be reported on through the G7 accountability process. The corresponding part of updating the list of commitments is deciding which commitments to drop (sunset), so that reporting on that commitment ends. Subsection 4.1.4 explains the criteria for that decision.

4.1.3 How the AWG processes a commitment

Once the AWG has recognised a commitment, it agrees on rules to guide reporting on its implementation. These rules are:

- The *base line*, meaning the starting point of the reporting period, usually the year when the commitment was agreed upon;
- a set of *indicators*: criteria to judge whether there is progress in carrying out the commitment; and
- a set of *data sources* to determine where the information used for reporting must come from.

The base line, the set of indicators and the set of data sources together form the *methodology* for a commitment. Methodologies are published along with commitments in every CAR (G7, 2019a). Readers of subsequent accountability reports can compare assessments and follow over time how the G7 has performed on a commitment. To make this possible, the methodology for a commitment is normally stable until reporting on it stops. Subsection 4.3 shows how the methodologies used in G7 practice deliver in terms of internal accountability.

The presidency that has hosted a summit proposes methodologies for the commitments agreed on at that summit. This is not only because the incumbent presidency sets the agenda in every working group (see subsection 2.1). Designing methodology requires interpreting a commitment (see subsection 4.2.1) and the presidency, having led its drafting and negotiation (see subsection 4.1.1), is presumed to understand it better than others. While the AWG relies on the presidency's proposals, it decides on methodology by consensus.

Each commitment gets to be reported on at least every three years, because CARs cover all active commitments and recur in a three-year cycle. In the intervening years, the presidency must issue an APR but can choose which active commitments it wants to cover. The presidency usually opts for a set of commitments that is thematically close to its summit agenda. But it must also be distant enough to make sure members do not see the drafting of the accountability report as prejudging summit outcomes. Germany's way to strike that balance in 2015 was to aim for summit outcomes on climate policy and issue an accountability report about progress on G7 commitments on biodiversity (G7, 2015b).

The presidency drafts a report, often with the help of academic institutions. Depending on the agreed data sources (see subsection 4.2.4), it draws on G7 national administrations, international organisations and others to gather the information it needs. The draft is cleared by the AWG and the sherpas and then published and endorsed by Leaders (see subsection 3.2). The AWG Terms of Reference require that an accountability report be published before the summit, even if this has not been possible from 2017 to 2019 (see subsection 3.4). The idea is that the report should serve as a reference for Leaders and the general public so that they can set the G7's future plans against its past performance. In practice though, the focus on the forward-looking parts of the summit agenda is so strong that accountability reports have no meaningful impact on the debates at and around the summits.

The way that the G7 processes its commitments does not have the features usually found in internal accountability. A principal often asks the agent to report back to her/him if certain

conditions are met. These conditions can be everything the principal should know to make an informed choice about how to proceed (for instance, a budget overrun or an emergency threatening the project goal). If the agent reports, s/he can be sure that the principal will engage and make a decision because otherwise s/he may not reach her/his goal. In the G7 however, the *disconnect* between backward-looking implementation reports and forward-looking summit agendas usually breaks this feedback loop. Accountability reports are delivered but not taken up in a Leaders discussion.

Recent research on G7 work in general seems to show that, given the right circumstances, this feedback can work (Cormier, 2018): If accountability for a G7 commitment coincides with a ministers' meeting on that commitment, compliance improves. G7 portfolio ministers act as the principals who inform themselves about the state of implementation and decide on the way forward. The reason might be that discontinuity does not affect portfolio ministers as much as Leaders. While the former keep to their portfolio, the latter must address issues across all policy areas. Section 5 develops a recommendation on that basis.

4.1.4 Reasons to stop reporting (sunset)

Leaders do not even explicitly decide when their administrations can stop reporting to them. The AWG takes that decision according to the value further reporting can add. Here are the reasons why the AWG decides to “sunset” a commitment:

- If the goal of a commitment has been achieved, reporting on it no longer makes sense, regardless of whether it had been met through G7 action. However, what counts as achievement depends on how the AWG understands the commitment (compare subsection 4.2.1): if the point is to achieve an outcome and then prevent backsliding, there is a case for continued reporting.
- If the G7 has changed the goal of a commitment, reporting on that commitment cannot just continue indefinitely. Implementation must follow what the G7 *currently* wants and reporting has to show whether this is the case. The G7 can change its goal in two ways: it can either stop pursuing it or it can replace it with a new goal that is larger in scope or reaches further in ambition. As a consequence, the G7 commitment to help increasing *health workforce coverage* (G7, 2017, p. 109) was folded into the larger commitment to support the improvement of *health systems* (G7, 2019a, p. 37), so that now there is no separate reporting on the former. These are implicit Leader decisions to sunset a commitment, because Leaders have made the choice to change the underlying goal or make a larger commitment.
- If a goal turns out to be unrealistic or not achievable by the measures that the G7 has committed to, it must suspend or change the commitment. Continued reporting will not do. The ball is back in the Leaders' court.
- If someone outside the G7 follows up on a commitment, the G7 does not need to duplicate that. The G7 has handed over the implementation of the New Alliance for Food Security and Nutrition in Africa, an initiative the G8 started in 2012, to the African Union and the New Partnership for Africa's Development (NEPAD), so there is no need to continue G7 reporting on it (G7, 2019a, p. 5). But the G7 still reports on its commitment to the BEPS

agenda, although the G20 has made a similar commitment and operates an accountability process of its own (Hilbrich & Schwab, 2018). This is because the AWG feels that the G7 as a group continues to add value to this process, even if the duties of G7 members under the BEPS discipline are no different from those of other G20 members. The G7 is supposed to add value through ongoing political leadership rather than mere compliance.

If one of these reasons applies, the G7 does not skirt its responsibility if it stops reporting on a commitment. That is why accountability reports must explain the reasons for sunset decisions. The Biarritz Progress Report does just that (G7, 2019a, p. 5).

4.2 Rules of reporting (methodology)

Well-designed rules can make sure that reports deliver what accountability is supposed to do: enable the public to compare G7 action to its words, prevent G7 national administrations from withholding relevant facts, and give G7 Leaders the information they need to make decisions on the way forward (see subsection 2.2). All this (and more) depends on how a commitment is understood. These subsections develop an approach to the interpretation of commitments (4.2.1) and then looks at how current G7 methodology performs against the criteria developed in subsections 2.2.2 and 2.2.3 (4.2.2 - 4.2.4). Its findings are limited in the sense that the agreed methodology only sets a minimum standard for reporting and the AWG is free to also report useful information that the methodology does not require. Good reporting practice can compensate for flaws in the methodology.

4.2.1 Understanding a commitment (interpretation)

A commitment must be interpreted because its follow-up requires decisions. G7 national *administrations* must decide what to do to implement the commitment. The commitment certainly constrains that discretion, but it does not determine what exactly administrations are supposed to do. They must solve implementation problems that Leaders have not anticipated and have the leeway that comes with delegation. After a period of implementation, the AWG must decide whether G7 administrations have made progress on the commitment. To make these decisions, the implementing part of an administration and the AWG both need a sense of the leeway and the limits of a commitment. Interpretation is meant to give that sense.

Every commitment serves as a benchmark of internal and external accountability. That is why interpretation has to ask whether national administrations have done what Leaders want and outside stakeholders can reasonably expect. The answer is a description of what successful implementation of the commitment must look like. The only way to find that out is using the *commitment text* agreed at the summit, because Leaders normally do not engage again once they have made a commitment (*disconnect*). Here are the ways to extract meaning from a commitment text:

- *Wording* and *logical structure* can reveal the tests that implementation must meet to count as successful. Some commitments define their goal. Others define measures. Subsection 4.2.3 shows why this is not enough.

- When making a commitment, G7 Leaders respond to a *problem* they want to solve and a set of external *expectations* about what the G7 is going to do about it. Why is it useful to know this historical *context* if problems and expectations change: because it tells us that a commitment *generally* applies to problems bearing the same features and situations raising the same expectations as the ones Leaders faced at the summit.
- It is also useful to look at *other commitments*. For example, the G7 commitment on *aid effectiveness* (G7, 2019a, p. 13) sets a general standard for the way national administrations must implement every G7 commitment that requires development cooperation. When implementing the whole set of active G7 commitments, G7 national administrations face *synergies* and *trade-offs* between different commitments (G7, 2015b, pp. 60ff.). Given that they owe compliance with all commitments to one principal, no commitment means that others can be disregarded. Every commitment must be implemented and interpreted in the context of other related commitments.
- What a commitment means also depends on its role in the accountability process. Reports on the implementation of a commitment are meant to tell the principal *whether there is still something to do* in order to reach his/her goal, and *who it is that must take further action*. In some cases, it is G7 Leaders, in others, G7 national administrations. That is why a G7 commitment must be understood from two sides: as an instrument of *internal accountability*, it means that G7 national administrations must do only their part to accomplish the goal Leaders have set. If they have implemented all required measures but the goal is not met, they must ask Leaders how to go forward. As an instrument of *external accountability*, a commitment means that the G7 must do all it can to accomplish its stated goal, regardless of what the Leaders have previously asked their administrations to do, because that may not be enough. The question whether there is still something to do has different answers for G7 Leaders and G7 national administrations.

4.2.2 Base line

The base line as agreed by the AWG only sets the starting point of the reporting period. It is normally the year when the commitment was made, the idea being that measures taken before the commitment cannot be seen as contributions to implement it. But there is a larger challenge. In the G7 accountability context, national administrations have an incentive to present what they would have done anyway as measures to implement a G7 commitment (see subsection 2.2.3). This can lead to exaggerated accounts of the actions the commitment has caused. To prevent this over-reporting, an accountability framework must control this risk. That is why the base line should also describe the relevant actions national administrations have taken *before* the commitment was made. Although this step cannot entirely isolate the actions caused by the commitment, at least it does discount some that were not. It also helps to build a theory of change because it can serve as a proxy for what would have happened without the commitment.

4.2.3 Indicators

Indicators determine what information is reported. Their job is making sure Leaders get the information they need. Leaders must know whether the G7 is moving towards or away from

its goal and why (see subsection 2.2.3). Joint effective action by G7 administrations, outside interference, or pure luck may all play a role. Leaders need to know the specific difference their administrations have made because this is the only factor they can control. This is salient in commitments like the one quoted in the Introduction (Section 1): hunger and malnutrition are rising although the G7 has increased its effort (G7, 2019a, p. 68). Without knowing the impact of their effort (as opposed to outside interference), the Leaders cannot make an informed choice about how to increase their efforts, change the approach or adjust the goal (see subsection 2.2.3). Having said that, most indicators used in G7 practice do not distinguish between the effects of measures meant to implement the commitment and the effects of outside interference (see subsection 4.3.1). Normally they just ask what G7 administrations have done and which results came about, without considering how the two might be linked. With this kind of indicator, Leaders wanting to accelerate progress against their goal may decide to raise their administrations' effort, even if this does not have an effect on the goal. This can happen because these indicators are not designed to show if and how the supposed causal link between effort and goal (theory of change) actually works.

Leaders and the public also need to know which actions they can attribute to the G7. In the G7 accountability context, national administrations have an incentive to present independent measures as "G7 action", but only things G7 administrations do together can be called that (see subsections: 2.1, 2.2.2, 3.1 and *How the AWG recognises commitments* under subsection 4.1.2). Indicators can and should be designed to tell group action from measures taken independently. There are easy ways to know whether an action is coordinated with others: one test is to ask if G7 administrations exchange information about what they are doing. But most indicators used in G7 practice do not work that way (see subsection 4.3.1), so that G7 accountability products may over-report G7 action.

On a more basic level, the public needs to be able compare G7 actions to its words (see subsection 2.2.2). For that, it needs a sufficiently concrete benchmark of behaviour. If the commitment itself is not clear enough, indicators must do that job as these define what behaviour counts as complying with the commitment. In G7 practice, some indicators provide more definition than others (see subsection 4.3.1).

Causation, attribution and definition are relevant to all commitments. But indicators also have to extract information that is specific to each commitment. There are some rules of thumb for designing indicators to fit commitments. Indicators should cover the *whole scope* of a commitment, so that no aspect of it goes unreported. They should not be biased towards quantitative information if the commitment calls for a *qualitative* assessment. And they must be *neutral* in order to compute progress towards, as well as sliding away from, the goal. On these counts, the indicators used in G7 practice do reasonably well (see subsection 4.3.1).

4.2.4 Data sources

Data sources determine where the information used for reporting must come from. The AWG usually chooses public information produced and regularly updated by an international organisation or "self-reporting", meaning that assessments draw on information each G7 national administration has provided itself. What difference can data sources make? A report can only be meaningful if it draws on information that speaks to the indicators. If a report uses irrelevant data sources, indicators cannot do their job because

there is no information to underpin them. The AWG must choose data sources carefully to fit indicators.

The choice of data sources can also protect G7 Leaders from part of the bias that comes with G7 national administrations reporting about their own work. G7 national administrations (as every agent) have an incentive to present their work in the best light. G7 Leaders (as every principal) must constrain opportunities to manipulate (see subsection 2.2.3). Choosing the information provided by an international organisation as a single mandated data source is one way to do that. While the national administrations represented in the AWG can still use a report to interpret this data in a certain way, they cannot leave it out.

The AWG often chooses “self-reporting” as a data source because international organisations do not normally offer information that sheds light on the *qualitative* side of G7 administrations’ behaviour (see subsection 4.2.3). But this does not constrain the agent’s bias towards reporting in his favour. Expert assessments (see subsection 2.3.2) avoid this bias.

4.3 How does this structure perform?

This section looks at whether the commitments the G7 has made and the accountability process it practices meet the criteria developed in subsections 2.2.2 and 2.2.3 above. The answers will point us to where change may be needed.

4.3.1 Quality of the commitments

Testing G7 commitments against the criteria developed in subsections 2.2.2 and 2.2.3 shows that there are good and bad examples. On some criteria though, no G7 commitment delivers. G7 Leaders never have information to distinguish the effects of their administrations’ joint efforts from outside interference. They never start from an explicit theory of change and have therefore no way to test it against implementation experience. These deficits mean G7 Leaders cannot lead implementation in the way principals in most internal accountability mechanisms can. Normally the principal extracts actionable information about what was effective in the agent’s effort and uses it to guide him/her on the way forward. This does not happen in the G7 because the current summit agenda displaces Leader-level discussion about implementation issues (*disconnect*). Here is how G7 commitments stand up against our criteria:

- Stakeholders should be able to compare G7 action to its words. That is easy with commitments like this: “We reaffirm our [...] commitment to reverse the declining trend of ODA to the Least Developed Countries [...]” (Elmau, 2015, p. 19). The Biarritz Progress Report (G7, 2019a, p. 10, Fig. 1.3) shows with minimal effort that the G7 has not achieved that. By contrast, the commitment “[...] to counter any form of stigma, discrimination and human rights violation [...]” (L’Aquila, 2009, p. 123) makes it hard to tell whether the G7 has complied. There is no clear benchmark for behaviour. The indicator (“G7 support for HIVAIDS”) does not help because it does not require specific information about countering stigma. This example shows why it is important for indicators to cover everything the commitment says (see subsection 4.2.3).

- Stakeholders and G7 Leaders should be able to distinguish what G7 administrations did as a group from what they did independently. Commitments like “We fully support a successful [...] replenishment of the Global Fund [to Fight AIDS, Tuberculosis and Malaria]” (Ise Shima, 2016, p. 12) clearly look at what G7 members do together. Accordingly, the indicator asks for the collective G7 contribution to the latest replenishment. The following commitment is different: “We commit to increase global assistance to meet immediate and longer term needs of refugees and other displaced persons as well as their host communities, via humanitarian, financial and development assistance [...]” (Ise Shima, 2016, p. 18). This text does not ask whether increases in assistance are due to collective G7 action or to measures G7 members take independently. Not making this difference means that the G7 might present as its own action what national administrations would have done anyway (over-reporting).
- Stakeholders should be able to know what the G7 is trying to achieve with a commitment. Commitments like the following: “We stand ready to provide [...] substantial technical assistance [...] to help implement a WTO Trade Facilitation deal, in particular to the benefit of the Least Developed Countries” (Lough Erne, 2013, p. 17) state their purpose in clear terms. The public can easily know that the World Trade Organizations’ Trade Facilitation Agreement (TFA) entered into force on 22 February 2017, that members implement it and that Least Developed Countries can expect support through the Trade Facilitation Agreement Facility. Other commitments make it clear that G7 action is intended to implement a standard set elsewhere, for example the disciplines on *aid effectiveness* or the norms on *asset recovery* contained in the United Nations Convention Against Corruption (UNCAC). Some commitments on the other hand, make a rather broad statement of purpose: “[...] to promote economic growth in developing countries and foster greater equality of opportunity within and between countries” (Charlevoix, 2018, p. 7). It is hard to tell whether the G7 is making progress towards this goal.
- G7 Leaders should be able to know the difference their administrations have made to their purpose, as opposed to the effect of outside interference. No active G7 commitment makes that difference. Most commitments define the measures G7 national administrations must take and state a goal for these measures, but they do not state their assumptions about the causal link between the two.
- G7 Leaders should set standards for the quality of their administrations’ performance. The G7 commitment on aid effectiveness does exactly that. The commitment on Antimicrobial Resistance also says how administrations are supposed to carry it out: “We commit to make collective efforts to strengthening and actively implementing a multi-sectoral One Health approach, taking into account the sectors including human and animal health, agriculture, food and the environment” (Elmau, 2015, p. 13; Ise Shima, 2016, pp. 12-13). Others do not: “We are [...] committed to intensifying our efforts to slow the loss of biodiversity” (Deauville, 2011, p. 54). At least, an indicator for this commitment asks whether biodiversity concerns are “mainstreamed throughout all aid planning and programming operations”.
- When making a commitment, G7 Leaders should have a concept of how their administrations’ measures will achieve their goal (*theory of change*). A plausible theory of change is implicit in some commitments: If multi-stakeholder initiatives promote responsible production and consumption in G7 countries, the application of labour, social

and environmental standards in global supply chains is expected to improve (G7, 2019a, p. 29). If G7 members support WHO reform and increase its resources, WHO capacity is expected to increase (G7, 2019a, p. 44). Most commitments however just seem to assume that increasing ODA to a sector will lead to improvements in that sector. Indicators for these commitments normally measure ODA flows to a specific sector, identified by a certain DAC code. The majority of G7 commitments do not consider how sectors interact or how outside forces can help or hinder their goal. Defining assumptions about how the G7 causes positive change and testing them against implementation experience would generate large gains in learning and effectiveness.

- But: G7 accountability currently does not have the tools to test the theory of change underpinning a commitment against actual implementation experience.

On all these counts, G7 can (and should) improve its practice without changing the accountability framework.

4.3.2 Quality of the process

The accountability process sets in *after* the G7 has made a commitment and mandated national administrations to implement it. For this reason, it cannot make sure the G7 makes the right commitments, chooses the right way to implement them, and sets the right incentives for those who will do it. In a limited way, the accountability process can make the G7 more responsive, for example by inviting civil society representatives to meet and discuss with the AWG, as happened in 2019. But the G7 must tackle most of these aspects *outside* its accountability process. This subsection focuses on a different question: How far does the G7 accountability process use the potential of internal accountability? Three features stand out:

- The process does not draw a clear line between what G7 national administrations do to implement a G7 commitment and what they would have done anyway because it is part of their normal routine (*additionality*). Many commitments and indicators are not designed to distinguish between group action and measures taken independently (see subsection 4.3.1) and reporting starts without a clear picture of what national administrations have done before the commitment was agreed (see subsection 4.2.2). This can lead to over-reporting and may harm credibility.
- The process does not sufficiently control the risk of national administrations painting a benevolent picture of their own activity. Consensus, hierarchy (see subsection 3.2), and broad reliance on “self-reporting” as a data source (see subsection 4.2.4) all work against critical assessment. The AWG is committed to delivering a true and fair view of what the G7 has achieved and does its best to control the bias inherent in the process. But it often still relies on “self-reporting”.
- Finally, the process does not deliver the learning effects that it could because Leaders focus on forward-looking parts of the summit agenda rather than implementation reports (*disconnect*, see subsection 2.1)) while nobody likes to publish his or her failures. That bias affects every level of the G7 process (see subsection 2.3). This is unfortunate because it misses an important point of internal accountability, which is to give the

principal access to the agent's implementation experience so that both can make better decisions (see subsection 2.2.3). This feedback and learning effect can even compensate for the bias that comes with the agent reports about his/her own work because it makes the principal learn to ask the right questions.

5 Ways to improve accountability

The G7 can use its accountability process to learn and make better decisions, without fundamentally changing its framework. Here is how:

- When making a summit commitment, the G7 should ask ministers to meet again and take stock of the situation after a period of implementation, for example two years. Or it could set conditions that define when political guidance is needed and that trigger a ministers' meeting if they occur. Ministers would then look at the state of implementation and decide on the way forward, using the latest accountability report. Knowing how its reports will be used, the AWG would focus them on the implementation questions that need political guidance. Accountability reports would become more actionable. This step would close the feedback loop without affecting the summit agenda. It makes sense because political guidance on implementation does not necessarily need to come from Leaders: ministers can do that job. But it would only work if fewer commitments were followed up. At the present time, the G7 accountability process for development-related commitments covers 46 commitments.
- The G7 should design its commitments and methodologies to allow it to guide implementation. Every commitment should have an *explicit* assumption about how the efforts made by G7 national administrations will achieve the intended goal (*theory of change*), so that the commitment can be tested against implementation experience. An idea drawn from the theory of change could even support the drafting and negotiation of the commitment (see subsection 4.1.1) because it would show G7 members what efforts it would require from their administrations and what risks they have to consider: the G7 should set goals which are specific enough to tell at least whether it has moved towards or away from them. It should also be clear what G7 administrations had already done *before* the G7 made a commitment (*base line*), so that ministers can see whether administrations have made any *additional* effort to implement it. Indicators should distinguish between the effects of administrations' effort on the one hand and outside interference on the other, so that ministers can make an informed decision either to increase efforts or to change their approach (if a goal was to be changed, they would of course have to turn to Leaders). Accountability reports should use implementation experience to see whether the theory of change underlying the commitment is right or whether ministers need to correct it.
- The G7 may even consider not publishing the actionable parts of accountability reports. Reports can only improve decision-making if they are open about failures and setbacks. While current G7 accountability reports are far from uncritical, the prospect of publication may still stop certain sensitive issues being mentioned. This step would ensure that the legitimate demands of external accountability do not stop internal accountability from taking place. The two functions should be divided: while non-public reports would focus on managing the internal relationship between G7 Leaders and national administrations, public reports would serve the external relationship between the G7 and outside

stakeholders. The former would help ministers manage implementation and the latter would show the public how far the G7 had delivered on its promises.

6 Summary and conclusions

One of the main purposes of internal accountability is to extract the experience gained from the implementation of a previous commitment in order to improve the way it is implemented going forward. G7 accountability for “development-related” summit commitments is essentially meant to deliver that feedback loop. Although the discontinuous nature of the G7 generally impedes effective feedback, the G7 can substantially improve the learning function of accountability by applying a small number of procedural fixes: In particular, G7 *portfolio ministers* can be the group where implementation experience is translated into forward guidance. Further improvements can be made if the G7 bases its commitments on *an explicit theory of change* and continuously tests this theory against implementation experience. This does not require building G7 accountability into a sophisticated evaluation process: designing commitments and methodology while retaining the basic tenets of contract theory in mind will do: Be clear about your purpose and the behaviour you want to see in G7 national administrations. Distinguish the difference they can make to your goal as opposed to outside interference. Prevent over-reporting. And be sure to get the information you require in order to know how far you have come. Finally, careful and transparent interpretation of the commitment text remains essential to the accountability process. A working-level follow-up must neither add to nor subtract from what Leaders have agreed. Attempts to “improve” Leaders’ decisions or make the G7 process more continuous than Leaders want it to be have their place in advocacy, not accountability. Progress on all these fronts may require reducing the number of commitments that are followed up – but this would be compensated by more meaningful reports.

Those same changes would also serve external accountability because stakeholders would then find it less hard to check G7 action against its words, have better reason to hope for effective implementation, and gain access to more meaningful reports. Be that as it may, the *main* burden of external accountability rests on how the G7 behaves (in terms of action and inaction) rather than how it designs its procedures. A large part of G7 action has been driven by the expectation that global interdependence be effectively managed and the G7 will certainly continue to be measured against this benchmark. G7 accountability can only solve a small part of the problem.

As G7 accountability will not change the informal and discontinuous nature of the G7, we need further research on how to reconcile spontaneous Leader-level policymaking with the continuity needed for effective implementation. Discontinuity has always been a feature of the G7 but changing events today seem to put an additional premium on flexibility. G7 accountability will only be robust to the extent to which it continues to accommodate this factor. My suggestion to enhance the role of G7 portfolio ministers is only one approach, with hopefully more to come. A larger question is how informal groups like the G7 can best deliver on their responsibilities under transformative agreements such as the 2030 Agenda. The limited scope of the G7 accountability framework means that it probably cannot evolve into an instrument for comprehensive reporting on all G7 actions relevant to the 2030 Agenda. A similar attempt at the G20 – the G20 Action Plan for the 2030 Agenda and its

Updates – is currently running into increasing difficulties, due in part to the similarly discontinuous nature of the G20.

Finally, accountability practice in the G7 is only one part of a large universe of accountability practices in informal international groups, including the G20, BRICS (Brazil, Russia, India, China and South Africa), and several more. Further comparative research on accountability in informal groups is of particular interest because it could fill at least two gaps: The description of accountability in informal international groups cannot easily apply analogies in the way, for instance, in which national administrative law helps one understand the World Bank's mechanism for grievance remedy. Informal groups are in a class by themselves (*sui generis*) also in the way they use accountability. Research must grapple with this specificity and describe the special forms accountability takes in informal groups. But research on accountability can also generate insight into the nature of informal international groups. With no fixed mandate, it is sometimes hard to tell how they actually work, what they are good at, and which questions they like to avoid. A closer look at their accountability practice may help to answer some of these questions.

References

- Brandi, C. (2019). Club governance and legitimacy: The perspective of old and rising powers on the G7 and the G20. *South African Journal of International Affairs*, 26(4), 685-702. doi: 10.1080/10220461.2019.1697354
- Buchanan, A., & Keohane, R. O. (2006). The legitimacy of global governance institutions. *Ethics & International Affairs*, 20(4), 405-437. doi: 10.1111/j.1747-7093.2006.00043.x
- Chayes, A., & Chayes, A. H. (1993). On compliance. *International Organization*, 47(2), 175-205. doi: 10.1017/S0020818300027910
- Cormier, B. (2018). Explaining G8 compliance: Reciprocity, ministerial institutionalization and accountability. In J. Kirton & M. Larionova (Eds.), *Accountability for effectiveness in global governance* (pp. 183-194). London: Routledge.
- Fischer, R. A. (2018). Towards sustainable global value chains: The role of the Group of Seven. In S. M. Jastram & A. M. Schneider (Eds.), *Sustainable fashion: Governance and new management approaches* (pp. 25-39). Cham, Switzerland: Springer. doi: 10.1007/978-3-319-74367-7
- G7 (Group of Seven). (2015a). *Think ahead. Act together. An morgen denken. Gemeinsam handeln. Leaders' Declaration G7 Summit (7-8 June 2015)*. Retrieved from https://www.g7germany.de/Content/EN/_Anlagen/G7/2015-06-08-g7-abschluss-eng_en__blob=publicationFile&v=3.pdf
- G7 (2015b). *Elmau progress report*. Retrieved from <https://www.bmz.de/g7/includes/Downloadarchiv/G7-Elmau-Progress-Report-2015-Biodiversity-A-vital-foundation-for-sustainable-development.pdf>
- G7 (2016). *Ise Shima progress report*. Retrieved from <https://www.mofa.go.jp/files/000158338.pdf>
- G7 (2017). *Taormina progress report*. Retrieved from https://www.esteri.it/mae/resource/doc/2017/10/taormina_progress_report_completo_sito.pdf
- G7 (2018). *Charlevoix progress report*. Retrieved from <https://www.mofa.go.jp/files/000498916.pdf>
- G7 (2019a). *Biarritz progress report*. Retrieved from https://www.diplomatie.gouv.fr/IMB/pdf/rapport-g7-bat-web_cle85fc7e.pfd
- G7 (2019b). *G7 Leaders' declaration: Biarritz, 26 August 2019*. Retrieved from <https://www.gouvernement.fr/en/g7-leaders-declaration>
- G8 (Group of Eight). (2009). *G8 preliminary accountability report*. Retrieved from www.g8.utoronto.ca/summit/2009laquila/2009-accountability.pdf
- G8 (2010). *G8 Muskoka accountability report*. Retrieved from www.g7.utoronto.ca/summit/2010muskoka/accountability/muskoka_accountability.pdf
- G8 (2011). *G8 Deauville accountability report*. Retrieved from www.g7.utoronto.ca/summit/2011deauville/deauville/2011-deauville-accountability-report.pdf
- G8 (2012). *G8 Camp David accountability report*. Retrieved from <https://www.usaid.gov/sites/default/files/documents/1868/CampDavidG8AccountabilityResport.pdf>
- G8 (2013). *Lough Erne accountability report*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/205437/Lough-Erne-Accountability-Report.pdf
- Hayford, P., & Kloke-Lesch, A. (2013). *Africa Partnership Forum evaluation report*. Retrieved from www.g8.utoronto.ca/scholar/APF-2013.09.18_Report_10.8.pfd
- Hilbrich, S., & Schwab, J. (2018). *Towards a more accountable G20?* (Discussion Paper 13/2018). Bonn: German Development Institute / Deutsches Institut für Entwicklungspolitik (DIE). Retrieved from https://www.die-gdi.de/uploads/media/DP_13.2018.pdf
- Keohane, R. O. (2002). *Global governance and democratic accountability*. Retrieved from https://www.researchgate.net/publication/228582436_Global_Governance_and_Democratic_Accountability

- Kirton, J., & Larionova, M. (Eds.). (2018). *Accountability for effectiveness in global governance*. London: Routledge.
- Leininger, J. (2009). *Think big! Future prospects of the international summit architecture – the G20, G8, G5, and the Heiligendamm Dialogue Process* (Discussion Paper 6/2009). Bonn: German Development Institute / Deutsches Institut für Entwicklungspolitik (DIE). Retrieved from https://www.die-gdi.de/uploads/media/DP_6.2009.pdf
- O'Neill, J., & Terzi, A. (13 June 2018). *The G7 is dead, long live the G7* [Blog post]. Retrieved from <https://www.bruegel.org/2018/06/the-g7-is-dead-long-live-the-g7/>
- Schedler, A. (1999). Conceptualizing accountability. In A. Schedler, L. J. Diamond & M. F. Plattner (Eds.), *The self-restraining state: Power and accountability in new democracies* (pp. 13-28). London: Boulder.
- Schmidt, K. M. (2017). Contributions of Oliver Hart and Bengt Holmström to contract theory. *The Scandinavian Journal of Economics*, 119(3), 489-511. doi: 10.1111/sjoe.12245

Publications of the German Development Institute /
Deutsches Institut für Entwicklungspolitik (DIE)

Discussion Papers

- 18/2020 Römling, Cornelia, & Sarah Holzapfel. *Monitoring in German bilateral development cooperation: A case study of agricultural, rural development and food security projects* (49 pp.). ISBN 978-3-96021-129-7. DOI:10.23661/dp18.2020.
- 17/2020 Laudage, Sabine. *Corporate tax revenue and foreign direct investment: potential trade-offs and how to address them* (46 pp.). ISBN 978-3-96021-128-0. DOI:10.23661/dp17.2020.
- 16/2020 Cavatorta, Francesco, & Fabio Merone. *Never-ending reformism from above and dissatisfaction from below: The paradox of Moroccan post-Spring politics* (32 pp.). ISBN 978-3-96021-127-3. DOI:10.23661/dp16.2020.
- 15/2020 Eppler, Mirko, Stella Gaetani, Francy Köllner, Jana Kuhnt, Charles Martin-Shields, Nyat Mebrahtu, ...Carlotta Preiß. *Information and communication technology in the lives of forcibly displaced persons in Kenya* (31 pp.). ISBN 978-3-96021-126-6 DOI:10.23661/dp15.2020.
- 14/2020 Hackenesch, Christine, Julia Leininger, & Karina Mross. *What the EU should do for democracy support in Africa: Ten proposals for a new strategic initiative in times of polarisation* (45 pp.). ISBN 978-3-96021-125-9. DOI:10.23661/dp14.2020.
- 13/2020 Never, Babette, Jose Ramon Albert, Hanna Fuhrmann, Sebastian Gsell, Miguel Jaramillo, Sascha Kuhn, & Bernardin Senadza. *Carbon consumption patterns of emerging middle classes* (42 pp.). ISBN 978-3-96021-124-2. DOI:10.23661/dp13.2020.
- 12/2020 Wehrmann, Dorothea. *Transnational cooperation in times of rapid global changes: The Arctic Council as a success case?* (29 pp.). ISBN 978-3-96021-123-5. DOI:10.23661/dp12.2020.
- 11/2020 Erforth, Benedikt. *The future of European development banking: What role and place for the European Investment Bank?* (36 pp.). ISBN 978-3-96021-122-8. DOI:10.23661/dp11.2020.
- 10/2020 Kaplan, Lennart. *Systemic challenges and opportunities of Franco-German development cooperation* (64 pp.). ISBN 978-3-96021-121-1. DOI:10.23661/dp10.2020.
- 9/2020 Stoffel, Tim. *Socially responsible public procurement (SRPP) in multi-level regulatory frameworks: Assessment report on policy space for SRPP regulation and implementation in Germany and Kenya* (55 pp.). ISBN 978-3-96021-120-4. DOI:10.23661/dp9.2020.
- 8/2020 Müngersdorff, Maximilian, & Tim Stoffel. *Strategies to strengthen socially responsible public procurement practices in German municipalities: A mapping exercise* (78 pp.). ISBN 978-3-96021-119-8. DOI:10.23661/dp8.2020.

[Price: EUR 6.00; publications may be ordered from the DIE or through bookshops.]

For a complete list of DIE publications:
www.die-gdi.de