Journal of International Business Policy (2021) 4, 28–57

www.jibp.net

Evidence-based policymaking and the wicked problem of SDG 5 Gender Equality

Lorraine Eden[1] and
M. Fernanda Wagstaff[2]

[1] *Department of Management, TAMU 4221, Mays Business School, Texas A&M University, College Station, TX 77843-4221, USA;*
[2] *Department of Marketing and Management, The University of Texas at El Paso, El Paso, TX 79968, USA*

Correspondence:
L Eden, Department of Management, TAMU 4221, Mays Business School, Texas A&M University, College Station, TX 77843-4221, USA
e-mail: leden@tamu.edu

Abstract
Evidence-based policymaking (EBP) contends that policy decisions are successful when informed by evidence. However, where policy problems are "wicked" (systemic, ambiguous, complex, and conflictual), politics trumps evidence and solutions are never first best or permanent. Applying an EBP approach to solving wicked problems (WPs) therefore appears to be a daunting, impossible task. Despite the difficulties, we contend that blending insights from the EBP and WP literatures can provide actionable and practical policy advice to governments and MNEs for dealing with the WPs of the UN Sustainable Development Goals (SDGs). We support our thesis with a case study applying EBP to the WP of SDG 5 Gender Equality. We compare the statistical evidence from gender inequality indexes to SDG 5's targets and indicators. We provide five insights from the EBP and WP literatures into why and how good evidence is necessary but not sufficient for progress on SDG 5. Building on these insights, we recommend that governments adopt an EBP approach employing public–private partnerships to address SDG 5. We also recommend that MNE executives use our new SDG Materiality Matrix, designed on EBP principles, to build SDG 5 into their global corporate social responsibility strategies.
Journal of International Business Policy (2021) **4,** 28–57.
https://doi.org/10.1057/s42214-020-00054-w

Keywords: sustainable development; gender equality; corporate social responsibility; evidence-based policymaking; wicked problems; gender inequality indexes; materiality; nonmarket strategy; policy-oriented studies

INTRODUCTION

On 25 September 2015, the United Nations (UN) General Assembly adopted the 2030 Agenda for Sustainable Development, a general pledge to "transform our world" and "leave no one behind" in terms of the economic, social, and environmental dimensions of sustainable development (UN, 2015b). The 2030 Agenda established 17 Sustainable Development Goals (SDGs) with a long list of targets and indicators that were to be collected, shared, and monitored by the UN Member States.

The 2030 Agenda is a form of goal-based global governance, where the 17 global goals define the sustainable development

Received: 27 May 2019
Revised: 19 May 2020
Accepted: 21 May 2020
Online publication date: 17 July 2020

aspirations of UN Member Countries and their major stakeholders (Fukuda-Parr, 2014). *Global goals* are instruments that "translate norms from the language of words to that of numbers, coupled with setting time bound targets" (Fukuda-Parr & McNeill, 2018: 6). Since the SDGs are voluntary, lack sanctions, and have few mechanisms to ensure their achievement, the 2030 Agenda is a form of "soft" international law (van Zanten & van Tulder, 2018: 212).

Progress on the 2030 Agenda has been slow, and criticisms have begun to appear. A January 2020 editorial in *Nature* (2020) argued that the SDGs were "not on track" and had a "bleak trend." *Nature* noted that only two targets were close to being achieved, and predicted that most SDGs would miss their 2030 target date. Lack of funding (an estimated shortfall of 2.5 trillion USD) and lack of government commitment were suggested as possible causes. There are also other reasons why the SDGs may be off track. For example, scholars have argued that many SDG targets are so conceptually complex that they cannot be translated into measurable indicators (Breuer, Janetschek, & Malerba, 2019). The SDGs themselves may be a problem, both their lack of prioritization (Breuer et al., 2019) and their number; there may be "too many goals, too little focus" for meaningful policymaking (Selin, 2015: 1).

The prospects for achieving the SDGs by 2030 have now become significantly worse since the COVID-19 pandemic spread across the globe in early 2020. In March, the estimated global cost of the pandemic for 2020 in terms of forgone world GDP was $1 trillion USD (Garten, 2020); by May, the estimate for 2020–2021 had risen to a 2-year loss of $8 trillion USD (UN, 2020: 1). The UN Secretary-General has called the coronavirus the "most challenging crisis since World War II", one that could cause a recession with "no parallel in the recent past" (Lederer, 2020). The effects are expected to be catastrophic for developing countries, and halt progress towards the SDGs (UNCTAD, 2020b).

We believe that slow progress on the 2030 Agenda was inevitable, even before the coronavirus pandemic, because the issues the UN are addressing are *wicked problems* (WPs; Rittel, 1972; Rittel & Webber, 1973; Alford & Head, 2017; Head, 2019). WPs are "systemic in nature, complexly interrelated, and materialize at the interface between public–private and profit–nonprofit interests"; as a result, they cannot be handled with "old management or leadership mindsets, or with old organizational structures" (van Tulder, 2018: 34). If WPs cannot be solved, policymakers may instead need to focus on managing or coping (Daviter, 2017; Head, 2019: 183). Assuming WPs scholars are correct, and slow progress on the SDGs was inevitable, what can policymakers do to ensure that the SDGs get back on track, or – given the global crisis now unfolding – that the SDGs are not derailed permanently?

We contend that insights from the evidence-based policymaking (EBP) literature can be helpful in spotlighting difficulties and suggesting policy directions for managing the WPs of the 2030 Agenda. EBP puts "the best available evidence from research at the heart of policy development and implementation" (Davies, 2004: 3). EBP scholars recognize that good evidence is necessary but not sufficient for good policymaking. Recognition that EBP is being applied to WPs can help dampen policymaker expectations, point out where difficulties and disputes are likely, and clarify achievable metrics for success.

We illustrate how insights from both the WP and EBP literatures can be useful for addressing the 2030 Agenda through a case study of SDG 5 Gender Equality, "achieve gender equality and empower all women and girls" (UN, 2015b: 14). Gender equality was originally a UN 2000–2015 Millennium Development Goal, which was carried over as SDG 5 in the 2030 Agenda (UN, 2015a, b). The global goal for SDG 5 is to achieve gender equality, and to empower women and girls by eliminating gender disparities, discrimination, and violence against women (UN, 2015b). A case study of SDG 5 is particularly appropriate because this year marks the 25th anniversary of the Beijing Declaration and Platform for Action that asserted that women's rights were human rights (UN Women, 1995). This year is also the 20th anniversary of the UN Security Council's Resolution 1325 on women, peace, and security, and the 10th anniversary of the establishment of UN Women (UN Women, 2020). Our study of SDG 5 is also especially salient, given that the harmful impacts of the COVID-19 pandemic are expected to fall disproportionately on women (Alon, Doepke, Olmstead-Rumsey, & Tertilt, 2020; UNCTAD, 2020a).

We begin by reviewing the WP literature and its applicability to the 2020 Agenda, focusing on SDG 5. We next analyze the existing evidence on gender equality. We employ country-based comparisons of the best-available gender inequality indexes, and

assess their ability to appropriately measure SDG 5's targets and indicators. We then turn to the EBP literature and show how EBP can provide useful insights for policymaking when faced with WPs. We argue that the generation and dissemination of high-quality, reliable evidence is necessary but not sufficient for making progress on the 2030 Agenda. Policymakers must also prepare for the many "slips between the cup and the lip" that bedevil EBP in addressing WPs. Lastly, we build on these insights to develop policy recommendations for governments and multinational enterprises (MNEs) for dealing with the WP of SDG 5. Figure 1 provides an outline of our paper.

THE WICKED PROBLEM OF SDG 5 GENDER EQUALITY

What are Wicked Problems?

The WP literature began as a critique of the systems approach to policymaking (Rittel, 1972; Rittel & Webber, 1973). In the systems approach, the planner implements a policy cycle with several steps: understand the problem; gather and analyze evidence; generate and assess solutions; and implement, test, and modify the solution. Rittel (1972) and Rittel and Webber (1973) criticized the systems approach, arguing that it could only handle "tame" problems. Since WPs were inherently unsolvable, their solutions could only be good or bad, not right or wrong; thus, a systems approach was unlikely to be successful.

Since the late 1970s, the WP literature has grown dramatically, in terms of both the number of scholarly disciplines and policy areas.[1] Ten key propositions underlie WP theory (see, e.g., McCall & Burge, 2016; Crowley & Head, 2017; Peters & Tarpey, 2019: 236), which are summarized below:

- *Problem*: (1) hard to define and no definite formulation; (2) no stopping rule; (3) unique; (4) multiple explanations; and (5) may be symptomatic of another problem.
- *Solution:* (1) not true or false but only good or bad; (2) no immediate or ultimate test; (3) no clear solution or even set of possible solutions; (4) attempts at solutions have effects that may not be reversible or forgettable; and (5) policymakers bear the responsibility for wrong solutions.

WP scholars believe that most policy problems have some degree of wickedness, so they cannot be solved using rational-scientific methods (Newman & Head, 2017).[2] Complexity is viewed as a key driver of wickedness, both the complexity of the problem (factual uncertainty) and of the actors/institutions involved (Alford & Head, 2017; Daviter, 2019). As the number and diversity of stakeholders and institutional contexts rise, so does the heterogeneity of preferences and interests, increasing the likelihood of conflict (Bannink & Trommel,

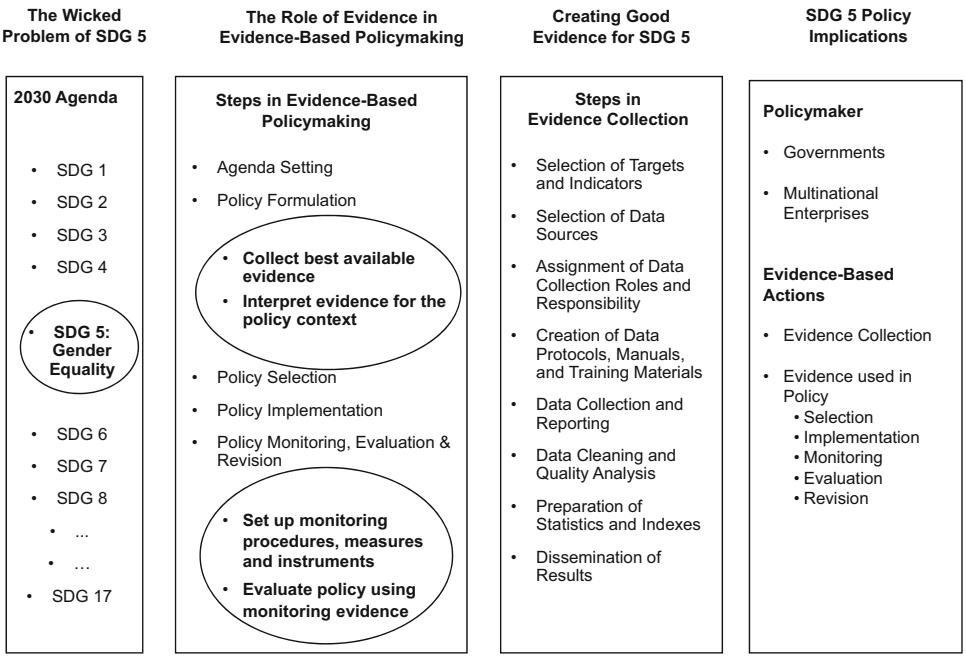

Figure 1 Applying evidence-based policymaking to the wicked problem of SDG 5 Gender Equality.

2019). Uncertainty and ambiguity exacerbate both complexity and conflict, increasing the degree of wickedness (van Tulder, 2018).

Key issues in the WP literature are how to address a WP and how to define success. Since WPs "are never solved. At best they are only *re*solved – over and over again" (Rittel & Webber, 1973: 160), policymakers cannot achieve first-best solutions. A frequent recommendation for addressing WPs, building on Rittel and Webber (1973), is that new and different ways of thinking are needed, and so also is collaboration or partnering among societal actors where they share joint responsibility for the solutions (Ney & Verweij, 2015; Crowley & Head, 2017; Daviter, 2017; Termeer & Dewulf, 2019).

An unexpected and paradoxical consequence of framing policy problems as wicked is that policymakers may choose to do nothing and just live with the problem (Bannink & Trommel, 2019). "Paralysis occurs when people experience or define the wickedness as so overwhelming that it discourages them and prevents them from doing anything about it" (Termeer, Dewulf, & Biesbroek, 2019: 176). Paralysis can be particularly wicked when grand societal challenges are framed negatively (van Tulder, 2018: 19). To avoid choice paralysis, policymakers are encouraged to explore "intelligent modes of imperfect governance" (Bannink & Trommel, 2019: 198), and to look for solutions that are "clumsy" or "just viable", which "everyone can more or less agree to live with" and are "responsive to different rationalities" (van Tulder, 2018: 39). Policymakers are also encouraged to focus on identifying, valuing, and learning from "small wins" (Termeer & Dewulf, 2019). Small wins are preferable to either doing too little (i.e., settling for paralysis or "cherry-picking" the least wicked parts of a problem) or expecting too much (i.e., the solving of an inherently unsolvable problem).

SDG 5 Gender Equality as a Wicked Problem

Scholars clearly view the 2030 Agenda as an example of a WP (Head, 2019; van Tulder, 2018). A core thesis of van Tulder (2018: 37) is that all 17 SDGs are WPs. The SDGs are "systemic in nature, complexly interrelated and materialize at the interface between public–private and profit–nonprofit interests. They are wicked both by nature and design" (van Tulder, 2018: 36) because they suffer from uncertainty, complexity, erratic dynamics, and ambiguity – all symptoms of WPs. The SDGs are societal problems, not tame or technical problems; for example, the guiding principle of the 2030 Agenda is that no one must be left behind – a huge societal challenge. Specific vulnerable groups are also regularly mentioned (e.g., women, children, minorities, migrants, refugees) and no vulnerable groups can be left behind.

To explore the WP of the 2030 Agenda in more depth, we provide a case study of SDG 5 Gender Equality "achieve gender equality and empower all women and girls" (UN, 2015b: 14). SDG 5 is decomposed into 9 targets and 14 indicators (UN, 2019); see Table 1. Ten of the other SDGs also include gender-specific indicators; as a result, 22% of the indicators for the 17 SDGs are gender specific (UN, 2019: 21–23), implying that gender equality is an important, cross-cutting goal in the 2030 Agenda.

The word "gender" refers to "the socially-constructed roles and responsibilities that societies consider appropriate for men and women" (Peace Corps, 2020). Gender equality means that "women and men enjoy the same rights and opportunities across all sectors of society, including economic participation and decision-making, [and that]…. the different behaviours, aspirations and needs of women and men are equally valued and favoured" (UNCTAD, 2016: 31). Gender equality is a "fundamental human right" and "keystone of a prosperous, modern economy that provides sustainable inclusive growth" (OECD, 2017: 3).

In the international business (IB) literature, gender equality has typically been defined as equal treatment of women and men *in the workplace* (Eden & Gupta, 2017; O'Brien, Fitzsimmons, Crane, & Head, 2017; UNCTAD, 2014, 2018; UN Economic Commission for Europe, 2019b). Workplace gender inequality is viewed as having many causes, including gender discrimination and stereotyping, undervaluation of women's work, gender-based labor market segmentation, traditions and culture that treat men and women unequally, and work–life balance issues (UNCTAD, 2014: 4). O'Brien et al. (2017) hypothesize that workplace gender inequality has three WP characteristics: divergence, complexity, and uncertainty. There are divergent views about the problem, no agreed definitions, and large differences in values, underlying beliefs and interpretations of findings. Workplace gender inequality also suffers from complexity due its multiple causes, lack of a dominant solution, and complex linkages with other societal issues. Lastly, uncertainty affects problem definition, prevents optimal solutions, and causes unintended consequences. The authors argue that policymaking requires

Table 1 SDG 5 targets, indicators, and gender equality metrics

SDG 5 Target	SDG 5 Target Indicators	GII	GGGI	SIGI	EM 2030	SDG Index
5.1 End all forms of discrimination against all women and girls everywhere	5.1.1 Whether or not legal frameworks are in place to promote, enforce and monitor equality and non-discrimination on the basis of sex				✓	
5.2 Eliminate all forms of violence against all women and girls in the public and private spheres, including trafficking and sexual and other types of exploitation	5.2.1 Proportion of ever-partnered women and girls aged 15 years and older subjected to physical, sexual or psychological violence by a current or former intimate partner in the previous 12 months, by form of violence and age				✓	
	5.2.2 Proportion of women and girls aged 15 years and older subjected to sexual violence by persons other than an intimate partner in the previous 12 months, by age and place of occurrence					
5.3 Eliminate all harmful practices, such as child, early and forced marriage and female genital mutilation	5.3.1 Proportion of women aged 20–24 years who were married or in a union before age 15 and before age 18			✓	✓	
	5.3.2 Proportion of girls and women aged 15–49 years who have undergone female genital mutilation/cutting, by age			✓		
5.4 Recognize and value unpaid care and domestic work through the provision of public services, infrastructure and social protection policies and the promotion of shared responsibility within the household and the family as nationally appropriate	5.4.1 Proportion of time spent on unpaid domestic and care work, by sex, age and location					✓
5.5 Ensure women's full and effective participation and equal opportunities for leadership at all levels of decision-making in political, economic and public life	5.5.1 Proportion of seats held by women in (a) national parliaments and (b) local governments	✓	✓		✓	✓
	5.5.2 Proportion of women in managerial positions		✓			
5.6 Ensure universal access to sexual and reproductive health and reproductive rights as agreed in accordance with the Programme of Action of the International Conference on Population and Development and the Beijing Platform for Action and the outcome documents of their review conferences	5.6.1 Proportion of women aged 15–49 years who make their own informed decisions regarding sexual relations, contraceptive use and reproductive health care				✓	✓
	5.6.2 Number of countries with laws and regulations that guarantee full and equal access to women and men aged 15 years and older to sexual and reproductive health care, information and education				✓	
5.a Undertake reforms to give women equal rights to economic resources, as well as access to ownership and control over land and other forms of property, financial services, inheritance and natural resources, in accordance with national laws.	5.a.1 (a) Proportion of total agricultural population with ownership or secure rights over agricultural land, by sex; and (b) share of women among owners or rights-bearers of agricultural land, by type of tenure					
	5.a.2 Proportion of countries where the legal framework (including customary law) guarantees women's equal rights to land ownership and/or control			✓	✓	
5.b Enhance the use of enabling technology, in particular information and communications technology, to promote the empowerment of women	5.b.1 Proportion of individuals who own a mobile telephone, by sex					
5.c Adopt and strengthen sound policies and enforceable legislation for the promotion of gender equality and the empowerment of all women and girls at all levels	5.c.1 Proportion of countries with systems to track and make public allocations for gender equality and women's empowerment	✓				

Sources: for SDG 5 targets and indicators: https://unstats.un.org/sdgs/indicators/indicators-list/; for GII metrics see http://hdr.undp.org/en/content/gender-inequality-index-gii; for GGGI metrics see WEF (158, Table 1); for SIGI see Branisa et al. (2014 and Supplemental Material); for EM2030 see EM2030 and Supplemental Material at https://data.em2030.org/2019-sdg-gender-index/methodology/, for SDG Index see SDSN (2019, Tables 5 and 7). These metrics may cover partially or totally the target indicators.

methodological reflexivity, the ability to see multiple world views, and the need to pay attention to context. Rational approaches based on a single discipline cannot handle the WP of achieving workplace gender equality.

It is important to recognize, however, that SDG 5 is about more than workplace gender equality; it is also about the *empowerment of women and girls*. The goal of SDG 5 is equality between men and women in their rights and opportunities, their valuation and treatment, and their empowerment (i.e., the fostering of women's voice and agency). Women's empowerment can be defined as "a woman's sense of self-worth, her decision-making power, her access to opportunities and resources, her power and control over her own life inside and outside the home, and her ability to effect change" (Peace Corps, 2020).

Women's empowerment has been shown to be a fundamental and necessary input for economic and social development.[3] The most recent empirical evidence on the negative economic and social consequences of disempowering women (such as worse governance, more conflict, less stability, worse economic performance, and slower social progress) can be found in Hudson, Bowen and Nielsen (2020). The authors assert that women's disempowerment is pervasive and embedded in societies around the world. Disempowerment has four dimensions: status (whether male and female groups engage as equals or as subordinate and superordinate), decision-making (whether decisions are made by one or both groups), conflict resolution (how disagreements are resolved, whether one group can be coerced against its will), and resource distribution (whether control over resources is by one or both groups). Hudson et al. (2020) argue that patrilineal/fraternal networks (e.g., tribes and clans) support and encourage practices that disempower women, such as violence towards women, personal status laws benefitting men, laws that prevent women from owning property, preferences for sons over daughters, and polygyny. The authors create a women's disempowerment index to assess the presence or absence of these harmful practices in 176 countries over 2010–2015, finding that 56 countries score low on the index (mostly OECD countries), 40 countries score high (primarily in Sub-Saharan Africa, the Middle East, and West and South Asia), and the rest are in between (Hudson et al., 2020: 54). The pervasiveness of these harmful practices against women lead the authors to conclude that "the true clash of civilizations is not about religion or ethnicity but about the subordination of women" (Hudson et al., 2020: 377). Thus, whether measured as workplace gender equality or as women's empowerment, SDG 5 is a wicked problem.

MEASURING GENDER (IN)EQUALITY

The Gender Inequality Indexes

One way to assess gender inequality is to examine government policies and laws promoting gender equality; however, policies *on the books* are not the same as policies *in practice*. A second way is to examine gender-based statistics, and, in fact, statistical indexes are the most common empirical method for assessing differences between men and women. The two best-known indexes are the UN Development Programme's (UNDP's) Global Inequality Index (GII), available since 2010, and the World Economic Forum's (WEF's) Global Gender Gap Index (GGGI), available since 2006. Given that the GII and GGGI were not developed for SDG 5, a few international and non-governmental organizations (NGOs) have started to build new indexes for SDG 5, responding to the UN (2015a, b). These indexes (the OECD's SIGI, the SDSN SDG Index, and the Equal Measures 2030 2019 SDG Gender Index) attempt to measure gender equality across countries and across time by collecting a broad set of socio-economic indicators on workplace gender inequality and women's empowerment. Below, we briefly review the five indexes and compare them in terms of SDG 5's targets and indicators.

The UNDP global inequality index (GII)

The GII is designed to measure the human development costs of gender inequality; i.e., the higher the GII value, the greater the gender gap and the larger the loss in human development (UNDP, 2018, 2019). The GII measures inequality between men and women in terms of economic opportunity, reproductive health, and empowerment. The GII combines women-specific indicators with indicators for both men and women; some scholars view this as an "odd mixture" of "women status" (level) and "gender inequality" (gap) (Permanyer 2013a, b: 940). Table 1 compares the components of GII to the SDG 5 targets and indicators; Table 2 summarizes the GII targets, indicators and data sources.

The WEF global gender gap index (GGGI)

The GGGI measures the gap between men and women in four target areas: economic participation and opportunity, health and survival, political empowerment, and educational attainment. The GGGI is reverse-coded from the GII so a higher GGGI score implies a country is closer to equality (GGGI = 1; GII = 0). Table 1 compares the components of GGGI to the SDG 5 targets and indicators, and Table 2 summarizes the GGGI targets, indicators and data sources. Table 2 also compares the GII and GGGI in terms of their targets and indicators; differences between the two indexes are shown in bold text.

The index is designed to disassociate the gender gap from a country's level of economic development, i.e., the index uses the male–female gap, not the actual level, for each indicator. The GGGI also uses ratios, and caps each ratio at 1 (gender equality) for countries where women outperform or have reached parity with men on the indicator.[4] The index is also designed to capture outcomes rather than causal factors of gender inequality, such as culture or government policies. In sum, the GGGI measures gaps not levels, outcomes not inputs, and equality not empowerment (WEF, 2020: 45).

The OECD social institutions and gender index (SIGI)

The SIGI was created to track progress on gender equality for the 2015 Millennium Development Goals (Branisa, Klasen, Ziegler, Drechsler, & Jutting, 2014: 31–32). Its five targets (discrimination in the family, restricted physical integrity, restricted access to productive and financial resources, and restricted civil liberties) are meant to capture the deprivation of women caused by gender gaps in social institutions (OECD, 2018a, b, 2019). Data from the OECD's Gender, Institutions and Development Database are used to create the SIGI, which is currently available for 4 years (2009, 2012, 2014, 2018). The SIGI scores 120 countries and organizes them into quintiles; thus, the SIGI not only has a shorter time series but also includes fewer countries than either the GII or the GGGI (OECD, 2019). Table 1 compares the components of the SIGI to the SDG 5 targets and indicators. The index is reverse-coded like the GII where 0 represents perfect equality and 1 perfect inequality.

Equal Measures 2030 SDG gender index (EM 2030 gender index)

The EM 2030 index uses indicators developed by the Inter-Agency and Expert Group on SDG Indicators. The index was developed from several frameworks: the UN Women SDG Indicator Framework and Women Turning Promises into Action report, the UN Minimum Set of Gender Indicators agreed by the UN Statistical Commission in 2013, and the Ready to Measure study produced by DATA2x (EM 2030, 2019; Buvinic, Furst-Nichols, & Koolwal, 2014; Buvinic & Levine, 2015). There are two EM 2030 indexes; a broad index including 51 gender-related indicators from 14 SDGs and a narrower index for only SDG 5 indicators. Table 1 compares the components of the narrower index to the SDG 5 targets and indicators. Higher scores represent greater gender equality.

The SDSN SDG gender index (SDSN gender index)

The Sustainable Development Solutions Network (SDSN) and Bertelsmann Stiftung developed the SDG Index and Dashboards to cover all 17 SDGs (Schmidt-Traub, Kroll, Teksoz, Durand-Delacre, & Sachs, 2017; Sachs, Schmidt-Traub, Kroll, LaFortune, & Fuller, 2019). The SDG targets are grouped into "five P's: Prosperity, People, Planet, Peace, and Partnership." Country scores range from 0 to 100, with higher scores representing greater gender equality (Sachs et al., 2019: 1). The components of the SDSN gender index relative to the SDG 5 targets and indicators are reported in Table 1.

Statistical Comparison of the Indexes

To date, only the GII and GGGI have sufficient years of data to compare them over time; it will be some years before the other indexes are sufficiently developed to make robust historical comparisons. We focus below on the GII and GGGI, building on Eden and Gupta (2017).

The UNDP gender inequality index (GII)

Table 3 shows the GII scores for several years between 2000 and 2017. Note that lower GII values represent greater movement toward SDG 5. The world average GII score fell from 0.432 in 2000 to 0.350 in 2017; thus, the gender gap shrank by 18.9% between 2000 and 2017. Given that gender equality is defined as GII = 0, a 35% gap still exists between men and women as of 2017.

Table 2 GII and GGGI targets and indicators

GII targets and indicators	GGGI targets and indicators
Economic opportunity • Labor force participation for women age 15 and older (ILO) • Labor force participation for men age 15 and older (ILO)	Economic participation and opportunity • Labor force participation: female/male (ILO, ILOSTATAT) • **Wage equality between women and men for similar work (WEF-Executive Opinion Survey)** • **Estimated earned income: female/male (WEF-UNDP methodology)** • **Legislators, senior official and managers: female/male (ILOSTAT)** • **Professional and technical workers: female/male (ILOSTAT)**
Reproductive health • **Number of women who died from pregnancy-related causes for every 100,000 live births (UN Maternal Mortality Estimation Group)** • **Birth per 1000 women ages 15–19 (UNDESA)**	Health and survival • **Sex ratio at birth: female/male (UN Population Division, World Population Prospects)** • **Healthy life expectancy: female/male (WHO, Global Health Observatory)**
Empowerment • Percentage of parliamentary seats held by women (IPU) • Percentage of women age 25 and older with at least some secondary education (UNESCO Institute for Statistics; Barro & Lee, 2016) • Percentage of men age 25 and older with at least some secondary education	Political empowerment • Seats in parliament: female/male (Inter-Parliamentary Union, Women in Politics) • **Positions at ministerial level: female/male (IPU, Women in Politics)** • **Number of years with a female head of state (last 50 years): female/male (WEF calculations)**
	Educational Attainment (UNESCO Institute for Statistics Education Indicators) • **Literary rate: female/male** • **Net primary enrolment: female/male** • Net secondary enrolment: female/male • **Gross tertiary enrolment: female/male**

Data sources in parentheses. Bold text shows that indicators are included in that index but missing from the other index.
ILO International Labor Organization; WEF World Economic Forum; *UNESCO* United Nations Educational, Scientific and Cultural Organization; *UN* United Nations; *WHO* World Health Organization; *UNDESA* United Nations, Department of Economic and Social Affairs; *IPU* Inter-Parliamentary Union.

The UNDP also calculates and publishes the Human Development Index (HDI) in which countries are grouped into clusters based on their economic and social development levels. The HDI groupings can help us see how different country groups are progressing towards gender equality relative to one another and over time. We calculate and report the GII score in Table 3 for the four HDI groups in two different ways. First, for each row (HDI group), we calculate the difference in GII scores between the base year (2000) and the most recent year (2017), as a raw score and as a percent of the base year. Since lower GII scores reflect greater gender equality, negative numbers represent improved performance and closer movement toward SDG 5. Looking from left to right across each row (each HDI group), we find that gender equality improved (i.e., the GII score fell) for all four HDI groups. The improvement was largest in the High and Very High HDI countries. Second, looking vertically (by year), we calculate the gap in GII scores between the bottom (Low HDI) and top (Very High HDI) groups, as a raw score and ratio. The gap between the raw scores falls slightly between 2000 and 2017, implying a small narrowing of the gender gap. The ratio of the raw scores (Low HDI/Very High HDI) rises from 2000 to 2017, suggesting that Low HDI group lagged behind the other HDI groups in improvements in gender equality.

The WEF global gender gap index (GGGI)

We also collected GGGI data for 2006–2017 and linked the GGGI data with the UNDP's HDI country groups; our statistical results are reported in Table 4. The world average GGGI rose between 2006 and 2017 from 0.663 to 0.698, implying a modest 3.5% improvement in gender equality. As of 2017, the gender gap recorded by the GGGI is 30.2%, a bit smaller than the 35% gap recorded by

Table 3 The UNDP gender inequality index, 2000–2017

Country average by Human Development Index (HDI) Group	2000	2005	2010	2013	2014	2017	Point gap 2000–2017	Percent change 2000–2017
World average	0.4315	0.4199	0.3886	0.3754	0.3659	0.3500	−0.0815	−18.89
Very High HDI countries	0.2112	0.1991	0.1825	0.1532	0.1500	0.1520	−0.0592	−28.04
High HDI countries	0.4927	0.4119	0.3887	0.3560	0.3458	0.3457	−0.1470	−29.83
Medium HDI countries	0.5419	0.5419	0.4939	0.4852	0.4716	0.4761	−0.0658	−12.14
Low HDI countries	0.6746	0.6396	0.6035	0.5953	0.5929	0.6024	−0.0722	−10.70
Point gap between Low and Very High HDI countries	0.4633	0.4404	0.4210	0.4421	0.4429	0.4504		
Ratio of Low HDI to Very High HDI countries	3.1936	3.2116	3.3064	3.8866	3.9520	3.9632		
No. of countries with data points	83	137	133	152	155	160		

Source: Authors' calculations based on GII data from the UNDP website: http://hdr.undp.org/en/content/gender-inequality-index-gii.

the GII. Looking across the rows in Table 4, gender inequality fell in all four HDI groups, with the largest improvement in the Low HDI group. Looking down the columns, the gap between the raw scores of the low and very high HDI countries fell over the period, while the ratio of the raw scores (Low HDI/Very High HDI) rose slightly, indicating that the GGGI gap narrowed across the HDI groups over time. Thus, measured across HDI groups, the gender gap between Low and Very High HDI countries narrowed, suggesting that improvements in gender equality were larger in Low HDI countries.

Assessing the Evidence on Gender Equality

Comparing the GII and GGGI: Why are their country rankings so different?

A comparison between the results for the GII in Table 3 and the GGGI in Table 4 reveals a perplexing picture. Table 3 shows a large gap in GII scores between Low and Very High HDI country groups that widens over time. Table 4, on the other hand, shows that the gap in GGGI scores between Low and Very High HDI countries is much smaller and narrows over time. Both indexes are on a scale of 1 although they reverse code equality [gender equality = 0 (GII), 1 (GGGI)]. What explains the differences between the two indexes, when considered by HDI group?

We hypothesize that the puzzling findings noted above are due, first, to differences in the targets and indicators for the GII and GGGI. Table 2 puts in bold text the indicators that are included in one index but not the other. There are several differences; e.g., there are four economic indicators included in GGGI that are missing from GII (e.g., wage equality, female/male ratio of earned income).

A second reason is that the goals and how they are implemented differ between the two indexes. The GGGI is designed to remove country levels of economic development; i.e., the index reflects gaps in gender equality ignoring the development in each country. The GII, on the other hand, is designed to incorporate the loss in human development as a function of gender inequality (Permanyer, 2013a, b; Piper, 2019; WEF, 2018, 2020). The differential treatment of economic development is clear when we compare the GGGI and GII scores across the four HDI country groupings. Table 3 shows that the GII score for Low HDI countries is nearly four times the score for Very High HDI countries. Table 4, on the other hand, shows very small differences in GGGI scores across the four HDI groups. Even when one inverts the ratio because GGGI is reverse-coded from GII, there is little variance by HDI level in the GGGI compared with the GII. This result is deliberate: the GGGI is designed to disassociate the gender gap from country levels of development, while GII is designed to include the impact of gender inequality on potential human development. As a result, some scholars worry that the GII may be proxying not only for differences in gender equality but also for differences in living standards across countries; this concern does not apply to the GGGI (Sotsky, Shibuya, Kolovich, & Kebhai, 2016).

Two-country case study: Mozambique and Nicaragua

As a second exploration of the underlying differences between the GII and the GGGI, we provide a brief two-country case study. We selected two

Table 4 The WEF global gender gap index, 2006–2017

Country average by Human Development Index (HDI) Group	2006	2007	2008	2009	2010	2011	2012	2013	2016	2017	Point gap 2006–2017	Percent change 2006–2017
World average	0.6626	0.6677	0.6751	0.6784	0.6786	0.6799	0.6844	0.6864	0.6947	0.6978	0.0352	5.3110
Very High HDI countries	0.6903	0.6972	0.7016	0.7050	0.7121	0.7136	0.7154	0.7179	0.7223	0.7256	0.0353	5.1084
High HDI countries	0.6637	0.6670	0.6763	0.6774	0.6734	0.6728	0.6756	0.6790	0.6856	0.6944	0.0307	4.6226
Medium HDI countries	0.6543	0.6571	0.6621	0.6725	0.6709	0.6713	0.6758	0.6782	0.6936	0.6947	0.0404	6.1774
Low HDI countries	0.6005	0.6102	0.6228	0.6246	0.6226	0.6314	0.6458	0.6463	0.6645	0.6643	0.0638	10.6215
Point gap between Low and Very High HDI countries	−0.0898	−0.0871	−0.0789	−0.0804	−0.0895	−0.0822	−0.0696	−0.0716	−0.0578	−0.0613		
Ratio of Low HDI to Very High HDI countries	0.8699	0.8751	0.8876	0.8860	0.8743	0.8848	0.9027	0.9003	0.9200	0.9155		
No. of countries with data points	110	122	124	128	130	131	133	136	144	144		

In calculating the overall GGGI average we used all 144 countries in the GGGI dataset. However, because North Korea is missing an HDI score and group, we used 143 countries in calculating the HDI country groupings.

Source: Authors' calculations based on WEF data for the Global Gender Gap. Data for 2016 and 2017 are from WEF (2017). Data for 2006–2013 are from the Humanitarian Data Exchange at https://data.humdata.org/dataset/global-gender-gap-index-world-economic-forum.

developing countries: Mozambique and Nicaragua. Both countries are ranked, when compared with the world average, relatively low on the GII but relatively high on the GGGI. Table 5 provides data for the two countries in terms of all five gender indexes (where available). We highlight in bold text the score of the country (Mozambique or Nicaragua) that performs better on each indicator. We also provide average world scores for comparison.

As Table 5 shows, Nicaragua outperforms Mozambique on all five gender indexes and on most indicators. Both countries rank poorly on the GII (Mozambique = 138; Nicaragua = 106) but do well on the GGGI (Mozambique = 29; Nicaragua = 6). The first reason for why both countries do poorly on the GII but well on the GGGI is because their indicators also do so; i.e., selection of the indicators is a key reason for the anomaly. The second reason is also visible; i.e., the way the indicators are measured. By focusing on gaps not levels and outcomes not inputs, the GGGI takes economic development out of the equation, whereas the GII deliberately includes human development.[5] Thus, with GGGI – but not GII – women may be worse off (in absolute terms) even though the gender gap is small.

Workplace gender equality versus women's empowerment

Our analysis of the GII and GGGI shows clearly that the two indexes are more narrowly focused than SDG 5. The targets and indicators for SDG 5 are designed to capture both gender inequality and empowerment, particularly for marginalized groups and others "left behind." The SDG 5 targets also consider not only gender inequality outcomes but also their antecedents, such as government policies, laws, and customs. As Table 1 shows, with the exception of target 5.5 (women's leadership) and target 5.c (policies and laws), neither GII nor GGGI cover the other targets and indicators in SDG 5.

We conjecture that the narrow foci of the GII and GGGI reflects an implicit focus on *gender equality in the workplace* rather than *empowerment of women and girls*. For example, neither the GII nor GGGI incorporate measures that Hudson et al. (2020) see as critical components of women's disempowerment.[6] As a result, neither the GII nor the GGGI is as comprehensive as UN SDG 5, which treats gender equality from a holistic perspective, considering voice, agency, and the empowerment of women and girls.

Table 5 Comparison of Nicaragua and Mozambique gender equality scores, 2017–2018

	Mozambique	Nicaragua	World
Overall SDG and SDG 5 scores and ranks			
UNDP GII Score (higher → less gender equality)	0.552	**0.456**	0.441
UNDP GII Country Rank (out of 160 countries)	138	**106**	
WEF GGGI Score (higher → more gender equality)	0.741	**0.814**	0.680
WEF GGGI Country Rank (out of 144 countries)	29	**6**	
OECD SIGI Score (higher → less gender equality)	.24	**.19**	.29
OECD SIGI Country Rank (out of 120 countries)	55	**30**	
SDSN SDG 5 Score (higher → more gender equality)	60.0	**82.1**	60.17
SDSN SDG 5 Country Rank (out of 162 countries)	81	**14**	
EM 2030 SDG 5 Score (higher → more gender equality)	61.1	**73.4**	62.0
EM 2030 SDG 5 Country Rank (out of 129 countries)	114	**84**	
UNDP GII Indicators			
Female labor force participation rate (%, age 15+)	**82.5**	50.3	48.7
Male labor force participation rate (%, age 15+)	74.6	84.0	75.3
Maternal mortality ratio (deaths per 100,000 live births)	489	**150**	216
Adolescent birth rate (births per 1000 women aged 15–19)	135.2	**85.4**	44.0
Share of seats in parliament (% held by women)	39.6	**45.7**	23.5
Female population, some secondary education (%, age 25+)	16.1	**48.3**	62.5
Male population, some secondary education (%, age 25+)	27.3	46.6	70.9
WEF GGGI targets/indicators			
Economic participation and opportunity	**0.789**	0.702	0.65
Educational attainment	0.857	**1.00**	0.96
Health and survival	0.977	0.98	0.97
Political empowerment	0.34	**0.576**	0.20
SDSN SDG 5 Indicators			
Unmet demand for contraception	56.8	**7.2**	NA
Female/male mean years of schooling	54.3	**106.3**	NA
Female/male labor force participation rate	**110.6**	59.9	NA
Seats held by women in national parliaments	39.6	**45.7**	NA

Bold text highlights the score of the country (Mozambique or Nicaragua) that performs better on each indicator.
Sources: EM 2030 (2019), OECD (2019), Sachs et al. (2019), UNDP (2018), and WEF (2018).

Our assessment that the GII and GGGI do not capture women's disempowerment suggests that the newer gender inequality indexes – SIGI, EM 230, and SDSN SDG 5 – should be better proxies for the targets on SDG 5. However, reliable and complete data are very hard to find for many SDG 5 indicators, particularly those for marginalized groups, and it will be years before longer time series datasets become available.[7] Time series indexes take time to build. We conclude that, at least for the present, although the three new gender indexes based on SDG 5 look very different *on paper* from the old gender inequality indexes (GII and GGGI), the new indexes *in practice* are much closer to the GII and GGGI, particularly for developing countries for which data for the SDG 5 targets and indicators are very scarce. Still, as these new indexes mature, better evidence on women's empowerment should become available.[8]

Which is the best index? It depends

The five gender indexes that we have reviewed above vary enormously in focus, breadth and depth, and sophistication. Some indexes use indicators based on gaps between men and women; others look at women's levels. Some indexes vary with a country's level of economic development, while others attempt to remove the relationship between gender inequality and the standard of living. Some indexes are narrowly focused on women at work, while others include many sociocultural and legal indicators designed to capture broader issues of women's empowerment. Some indexes have long historical time-series datasets, while others are just beginning to finalize methodologies, collect data, and report statistics on a cross-section basis. Some indexes weight all their components equally, while others use unequal weights.

So how can policymakers decide which of the five gender inequality indexes is best for tracking their country's performance on SDG 5? Similarly, how do MNE executives select an index to track their company's gender equality performance, at home and in their foreign affiliates? We offer three insights into these questions. First, we argue that there is no single right answer. Policymakers and MNE executives need to match their selection of a gender inequality index with the specific gender equality goals and targets they are trying to achieve. The "best" gender inequality index for one government or one MNE may not be the same for another where their goals and targets differ. There may be no "best" index nor "off the shelf" solution; it may be that a single index is not sufficient and will need to be assembled from different indexes. Second, policymakers and MNE executives need to understand why and how the various indexes are constructed; that is, making a "deep dive" into the selection and measurement of indicators and construction of the indexes. A thorough analysis of the various indexes is a necessary first step to selecting or creating an inequality index that best fits the organization. Third, our overall assessment is that, given their strengths and weaknesses, the GII and GGGI are better able to usefully inform policymakers and MNE executives on how to move toward SDG 5 in terms of workplace gender equality. For more holistic measures of gender equality that take account of women's empowerment, policymakers must look to the new SDG 5 gender indexes. Policymakers should also study Hudson et al.'s (2020) new women's disempowerment measure.

EVIDENCE-BASED POLICYMAKING AND THE WICKED PROBLEM OF THE SDGS

We turn now to the second part of our analysis. Given the existing evidence on gender inequality, how can policymakers use this evidence to address the WP of SDG 5? We contend that EBP is the appropriate response.

What is Evidence-Based Policymaking?

The premise behind EBP is that policy decisions are more likely to result in better outcomes when informed by evidence (Scott, 2005). EBP uses "open, smart and trusted statistics [that are] relevant for the society" (Rademacher, 2019: 524), and puts "the best available evidence from research at the heart of policy development and implementation" (Davies, 2004: 3).[9] While EBP has been primarily used in OECD countries, EBP has also been applied in developing countries (Sutcliffe & Court, 2005, 2006; Bartlett, 2013; Hewlett Foundation, 2018).[10]

The original approach to EBP assumes a linear relationship between evidence and policy choice where the policymaker defines a problem, identifies what is known, collects the best available evidence, and makes a policy choice. The basic steps in EBP are outlined in Figure 1. The current approach to EBP embeds evidence into the policy cycle process in a circular fashion, as shown in Figure 2. The EBP policy cycle starts with agenda setting, moves through policy formulation, selection, and implementation, to monitoring, evaluation, and revision, in a circular fashion.[11]

Good evidence is a fundamental component of the EBP cycle. The key evidence steps in EBP are illustrated in the circled and bold text items in the second box in Figure 1; the sub-steps in the evidence collection phase are shown in the third box in Figure 1. The evidence steps are also highlighted in Figure 2 as steps 5 (obtaining evidence), 7 (interpreting evidence), 13 (monitoring), and 14 (evaluation). Thus, collecting, analyzing, and using good evidence are key elements in successful EBP.

Wicked Problems and Evidence-Based Policymaking

Can EBP be applied successfully to WPs? One might think that EBP should be applicable to WPs, given that both the WP and EBP literatures have their historical roots in the systems approach to planning. However, the WP literature arose as a *critique* of the systems approach, whereas the EBP literature *adopted* the systems approach and the policy cycle as part of EBP. As a result, neither literature has had much to do with the other for the past 40 years (Head, 2019).

In fact, the WP literature has been very critical of the EBP approach. WP scholars have long argued that the rational, step-based approach to policymaking illustrated in Figure 2 is incapable of addressing policy problems that are ill-defined, complex, and amorphous (Newman & Head, 2017; Daviter, 2019; Head, 2019; Termeer et al., 2019). Because poorly structured problems cannot be handled by "advanced and precise analytic methods", EBP "will predictably fail to grapple with the challenge of wicked problems" (Daviter, 2019: 67).

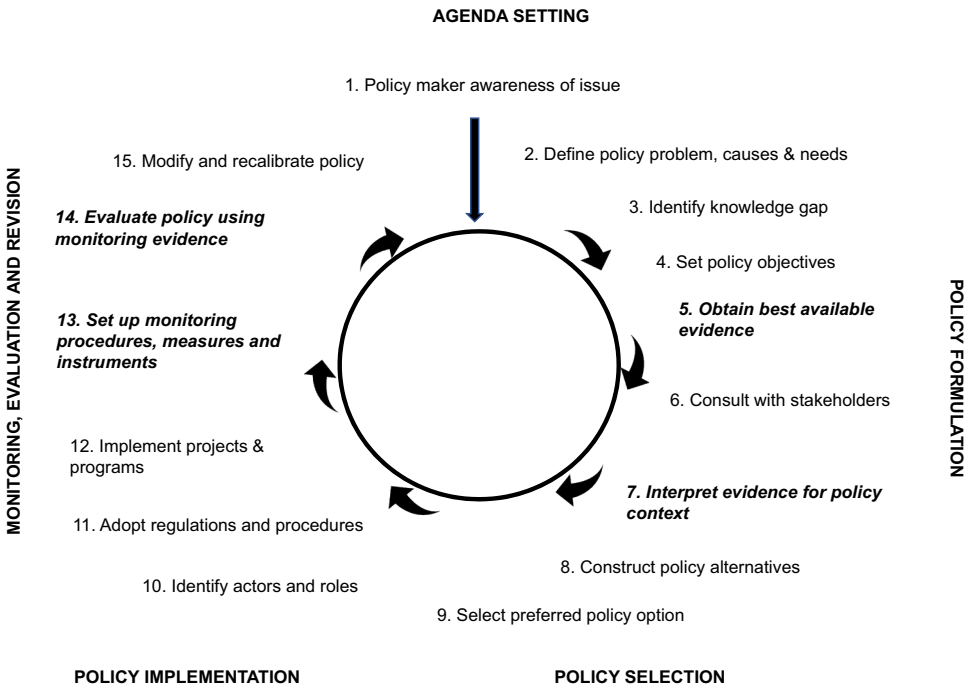

Figure 2 Evidence-based policymaking – a policy cycle approach.

EBP is also criticized for its implicit assumption that "access to better knowledge….will lead to greater consensus on how to improve policy outcomes". WP scholars argue instead that "political dynamics… cannot be neutralized by a dose of evidence, no matter how relevant or powerful" (Newman & Head, 2017: 419). WP scholars assert that policymaking is based, not on scientific evidence, but on "stakeholder perceptions, values and interests in explaining how issues are scoped, priorities are set and possible solutions considered" (Head, 2019: 188). Politics drives policy, not evidence, according to WP theorists.[12] As a result, the WP literature has been critical of and has ignored EBP for decades.

We contend that EBP *can* provide useful insights for addressing the WPs of the 2030 Agenda. Our thesis may appear overly optimistic, disingenuous, or even radical, given the criticisms of EBP by WP scholars noted above. We believe, however, that both literatures can inform each other and are particularly useful for analyzing the WP of the 2030 Agenda.

The timing may also be right for exploring connections between these two literatures. Recently, a few WP scholars have begun to take a second, more positive look at scientific approaches and EBP. For example, Head (2019: 183) contends that, after ignoring insights from the policy studies and public management literatures for 40 years, WP theory would benefit from reconnecting with these fields. He argues that WP theorists need to "draw more deeply on cutting-edge developments in contemporary policy sciences", and that policymakers need "to use best available evidence and to communicate the value of open processes for tackling complex and wicked problems" (Head, 2019: 192). Given "populist distain for expertise" and "widespread lack of trust in the institutions of public governance", Head (2019: 192) concludes that WP theorists need to use "best available evidence."

At the same time, EBP scholars have also moved away from the simple model in Figure 2, recognizing that "it is a long away from getting the facts straight to analyzing complex policy problems" (Daviter, 2019: 70). In pluralistic societies, political dynamics and embeddedness can matter as much or more than evidence in an EBP process, even when facing "tame" (technical or scientific) problems (Newman & Head, 2017). The current EBP literature is also sensitive to several issues that are inherent in WPs, such as complexity, uncertainty, ambiguity, and divergence (Head & Alford, 2015). Many ambiguities, for example, are likely involved in applying EBP to a WP: ambiguities in framing the issue (given interdependencies), determining the evidence (sorting opinions from facts, deciding

what is relevant), selecting targets (that are typically moving), selecting and implementing a policy or policies (which may have unintended consequences), and monitoring outcomes (relative to an unknown counterfactual) (van Tulder, 2018: 43–50).

Given these trends, we believe that the time is right for exploring how insights from the WP and EBP literatures could help devise appropriate public policies for managing the WPs of the SDGs.

Blending Insights from the EBP and WP Literatures

Below, we explore five insights drawn from the EBP and WP literatures, which we believe could help policymakers address the WP of the 2030 Agenda.

Insight #1: Good evidence is necessary but not sufficient

The proper application of EBP requires the creation and use of good evidence, which raises methodological questions among stakeholders over "what is evidence" and how to make it accessible to policymakers. High-quality evidence is defined as the "best available" evidence that is appropriate for the problem and has been "*systematically* searched, critically appraised, and rigorously analysed according to explicit and transparent criteria" (Davies, 2004: 7). Most EBP scholars agree that all forms of evidence collected through a systematic process should be included, e.g., theory building, data collection, and analysis, and practice-based wisdom (Sutcliffe & Court, 2006: 2). At the other end of the scale, EBP scholars also agree that using poor-quality research as evidence clearly reduces the efficacy of EBP.

However, once we move past these two zones of agreement (what is in/what is out), disputes emerge because the "devil is in the details." A first problem is that definitions of what constitutes appropriate evidence can vary across stakeholders, ranging from narrow (e.g., peer-reviewed academic journal articles) to broad (e.g., professional experience, stakeholder accounts). Stakeholders also disagree on what constitutes evidence in different issue areas. Nutley, Davies and Walter (2003: 31–32), for example, note that health care has an "established hierarchy of evidence for assessing what works"; whereas fields such as education and criminal justice are "riven with disputes" over what is appropriate evidence.

Moreover, evidence is generally assumed to be synonymous with empirical evidence, with the implicit assumption that "hard" evidence (official statistics, econometrics, indexes) is more rigorous than "soft" evidence (qualitative studies, expert evaluations). However, empirical findings are typically built on mathematical models that involve simplification and parsimony, which suggests that drawing policy recommendations from empirical results may not only be meaningless but potentially harmful. Saltelli and Giampietro (2017: 64) note that it is "one thing…to use a model to simulate a policy, another story is the leap whereby the same model is used to justify one."

Thus, the first step in EBP, the collection and provision of good evidence, is a necessary but onerous and contentious requirement for successful EBP. Moreover, while collecting, building, and analyzing "good evidence" are necessary steps, they are not sufficient. A successful EBP process needs to recognize and account for factors other than evidence that influence the policy process. Awareness of these potential problems is critically important for governments and MNE executives as they attempt to choose and implement the WP of the 2030 Agenda.

Insight #2: Good evidence may be misunderstood or misused by policymakers

Even though the adage "better statistics, better decision-making" is widely accepted, there can be "slips between the cup and the lip" that derail an EBP process. Policymakers may be "flying blind: trying to make policy decisions in the dark", even when good evidence exists, if the evidence cannot be found, or is misunderstood (Hewlett Foundation, 2018: 3).

For example, high-quality evidence may not be sufficient to motivate policymakers if they do not have the policy capacity or level of expertise needed to understand or use the evidence (Newman, Cherney, & Head, 2017). Successful EBP requires the government sector to have skilled analytical staff that are trained in data analysis and policy evaluation (Head, 2010). The level of training and professional competence of civil servants and government legislators varies across countries, along with their access to resources and institutional infrastructure. Policy capacity is typically much lower in developing than in developed countries (Howes, Betteridge, Sause, & Ugyel, 2017).

Misuse and misunderstanding of research findings by policymakers can also be caused by the researchers themselves, who seldom consider "policymaker friendliness" when disseminating their research findings. Researchers need to communicate well to "make their research relevant and readable" (Oliver & Cairney, 2019: 3). There are also multiple ways in which numbers can be misunderstood or misinterpreted, creating "mutant statistics" that lead to bad policymaking (Best, 2012). Examples include problems of generalizability, definition, inadequate measurement, bad samples, changing the meaning of statistics, garbling complex statistics, and compounding errors by creating chains of bad statistics. To prevent such technical biases, new institutions for good governance of evidence may be needed if EBP is to function effectively (Parkhurst, 2017).

It is also possible that the specialized knowledge of academics can be counterproductive, creating contentious debates over highly technical alternatives among experts on different sides of the problem. The uncertainty of social science and the different status of knowledge fields can cause policymakers to doubt or refuse to include particular studies as evidence, leading to disputes between policymaking groups. Where WPs are involved, there may be equally compelling evidence that supports competing perspectives, causing stakeholders to privilege the evidence that favors their own position (Daviter, 2019: 67). A related concern is the politicization of science whereby policymakers bury, misuse, manipulate, or cherry-pick evidence to support and promote their own policy preferences. EBP can be turned into policy-based evidence, generated by mechanisms such as knowledge monopolization, blame avoidance, black-boxing, or oversimplification of the evidence (Strassheim & Kettunen, 2014: 263).[13]

Insight #3: Good evidence is often trumped by politics

While EBP scholars stress the role played by good evidence, they also recognize that exogenous and endogenous factors other than evidence are regularly – and legitimately – involved in policymaking. Examples include the experience, expertise, and judgment of policymakers; the constraints of finite resources; the importance of values such as ideology and political beliefs and the role of habit and tradition; the power exerted by lobbyists, pressure groups, and consultants; and, more generally, the pragmatics and contingencies of political life (Davies, 2004; Head, 2010; De Marchi, Lucertini, & Tsoukiàs, 2016; Richards, 2017; Saltelli & Giampietro, 2017).

In theory, the selection and measurement of targets and indicators for building good evidence should be decisions made on technical or scientific grounds. In practice, the evidence collection stage can become a highly contested, political activity because the measurement tools are essential to defining the goals (Fukuda-Parr & McNeill, 2018).

Getting appropriate "buy-in" from politicians and bureaucrats to using EBP can also be difficult. "[P]ractitioners need incentives to use evidence and to do things that have been shown to be effective. This also means not doing things that have been shown to be ineffective or even harmful" (Davies, 2004: 20). Policymakers must have the capability, opportunity, and motivation to use evidence before they engage in behavioral change (Langer, Tripney, & Gough, 2016: 4). Outside events, such as pressing but unqualified threats, can also cause policymakers to "throw caution to the wind", ignore evidence, and make quick, precautionary decisions (Monaghan, Pawson, & Wicker, 2012). Arguing that evidence should prevail over pragmatism is difficult for politicians where politics and budget constraints are likely to be more important than evidence in driving policymaking (Richards, 2017).

These political realities at every stage in the EBP process accord well with the concerns of WP theorists. One of the core insights of the WP literature is that the political arena is the true battlefield in policymaking because public policy problems are inherently political problems (Newman & Head, 2017; Head, 2019). As the wickedness of a problem increases, conflicts among stakeholders with differing values become more important and the usefulness of scientific evidence diminishes.

Insight #4: Good evidence needs networks and partnerships

Networks and partnerships among the various stakeholders in EBP – government, business, nongovernmental organizations (NGOs), policy think tanks, academics, and the general public – are necessary for successful EBP, particularly in democratic systems and when dealing with WPs (Ney & Verweij, 2015; Crowley & Head, 2017; Daviter, 2017; Termeer & Dewulf, 2019). Partnerships are "voluntary and collaborative relationships between various parties, both State and non-State, in which all participants agree to work together to achieve a

common purpose or undertake a specific task and to share risks and responsibilities, resources and benefits" (UN, 2005, par. 8, p. 4).

Partnerships between government policymakers and their stakeholders can offer benefits, such as access to resources, capacity and coalition building, more policy relevant research, and linkages with stakeholders (Richards, 2017). EBP is more likely to be successful when the "actors in the evidence-to-policy ecosystem…..have regular opportunities to connect with each other, develop relationships of mutual trust and respect, and exchange ideas and learning" (Hewlett Foundation, 2018: 8).

Academics and policymakers have historically been viewed as "two communities" with few contacts and bridges, either formal or informal, between them (Caplan, 1979). While evidence suggests the situation is better now than in the 1970s (Newman, Cherney, & Head, 2016), it is still the case that factors which could encourage networking, such as knowledge brokers, shared agendas, and common meeting places (e.g., joint conferences), are often missing or weak (Lugo-Gil, Jean-Baptiste, & Livia Frasso Jaramillo, 2019). For example, academic scholars are seldom motivated to work on policy issues nor encouraged to disseminate their research findings in ways that reach policymakers or to interact in policy circles. The "publish-or-perish" syndrome and the rewards systems in universities do not privilege policy-based research or reward linkages between government and academia, viewing them as consulting or secondary contributions to knowledge (Eden, Lund Dean, & Vaaler, 2018, Chapter 28).[14] Academics need to learn the "dos and don'ts" of how to influence policy (Oliver & Cairney, 2019). For expert advice to be helpful, academic experts may also need to be more formally embedded in the policymaking process and able to think "outside the box" of their own discipline and specialty (Daviter, 2019).

Insight #5: Good evidence here may not work there

A concern particularly salient for IB scholars arises from the difficulty of applying EBP in different country settings. Context matters, and "what works here may not work there." A simple but telling example is the need for and difficulty of adapting policies that work in OECD countries to the very different and differing institutional contexts of developing economies. While EBP has been used mostly in OECD countries,[15] most EBP experts agree that the potential economic gains from successfully implementing EBP in developing countries are likely to be large, possibly much larger than those found in case studies of EBP in OECD countries (Sutcliffe & Court, 2005, 2006; Bartlett, 2013; Hewlett Foundation, 2018).

Only a few developing countries have embraced and attempted to implement EBP; for example, Malawi (Government of Malawi, 2016) and Tanzania (Lubua & Maharaj, 2012). The reasons for why EBP initiatives have failed in developing countries include the application of orthodox approaches inappropriate to countries at different stages of development, weaker economic conditions and more difficult political environments (e.g., political volatility, corruption), and institutional voids (Data for African Development Working Group, 2014; Hewlett Foundation, 2018). Even where international organizations have developed EBP "toolboxes or kits" for developing countries, a common concern is that the toolkits may be ideologically-based, promoting orthodox policies inappropriate for countries at different stages of development (Sutcliffe & Court, 2005, 2006; European Commission, 2017b). The heterogeneity of developing economies also makes it difficult to translate the lessons from EBP case studies carried out in one developing country to another because the environmental contexts are so different.

Even for a single developing country, the hurdles can be interactive and reinforce one another. A particularly difficult "cocktail" is the mix of poor evidence, political realities, and weak academic–policymaker networks (Hantrais, Lenihan, & MacGregor, 2015). While there is some research on how to successfully apply EBP in developing country contexts (see, e.g., the Overseas Development Institute's RAPID Framework; Sutcliffe & Court, 2005, 2006; Court & Young, 2006), clearly the hurdles are higher and the success stories fewer in number.[16]

Problems at the single country level are compounded in a multi-country framework where attempting to implement EBP generates huge coordination issues. Examples are Bartlett's (2013) study of the hurdles faced by applying EBP to labor-skill policies in the EU enlargement countries[17] and Lofstedt & Schlag's (2017) study of the debate over banning the chemical Bisphenol A in the European Union.[18]

The problems of applying EBP to multiple countries are even more acute in developing countries. A useful (and timely, given the current COVID-19 pandemic) analysis of the challenges is the study by

Andrus, Jauregui, De Oliveria and Ruiz Matus (2011) of the Pan American Health Organization's ProVac Initiative in the Americas, which was designed to ensure that developing countries had equitable access to new vaccines. The authors found that more lives were saved more quickly when national governments had sufficient policy capacity, took responsibility for helping to pay for and distribute the vaccines, and were supported by strong partnerships with international organizations.

In sum, "what works there may not work here" reminds us that an EBP approach to WPs will be particularly difficult in multi-country cases with multiple stakeholders and wide differences in institutional contexts and levels of development. Top–down, "one size fits all" policies – even where the definition of success is tackling, managing or coping – are unlikely to be accepted or successful.

POLICY RECOMMENDATIONS FOR THE WICKED PROBLEM OF SDG 5

As we have shown above, determining what is and how to use high-quality evidence in an EBP framework is far from easy and especially difficult where WPs are involved. Despite these difficulties, we argue that policymakers can benefit from applying the EBP and WP literatures to the WPs of the SDGs. Drawing on these insights, we make some policy recommendations for governments and MNEs for addressing the WP of SDG 5 Gender Equality.

SDG 5 Policy Recommendations for Governments

Recommendation #1: Formally adopt EBP and prioritize completion of evidence collection

The UN recognized early in the multi-stakeholder negotiations led to the 2030 Agenda that "good quality, verifiable evidence on progress toward achieving the 2030 Agenda" would be necessary, together with a formal monitoring and evaluation mechanism to ensure accountability and benchmark country progress (MacFeely, 2019b: 3). Implementation of the 2030 Agenda was to be based on "sound evidence and science, taking advantage of contemporary approaches from the sustainability sciences including systems thinking and analysis and quantitative modelling" (Allen, Metternicht, & Wiedmann, 2018: 1454). Thus, the United Nations in 2015 implicitly, if not explicitly, recognized the importance of evidence and made a commitment to an EBP process for the 2030 Agenda.

UN Women has been actively leading the process for SDG 5, working with PARIS21, a worldwide network of statisticians and other stakeholders that are committed to EBP in the public sector. Their goal is to develop better gender statistics through assessing country-level data and statistical capacity gaps, with the ultimate goal of using these assessments to develop better strategies for national statistics (UN, 2013, 2016; UN Women and PARIS21, 2019). UN Women (2018, 2019a) has set up a Women Count Data Hub with an SDGs Dashboard (https://data.unwomen.org/countries) where raw data are posted and analyzed, by country, for as many of the SDG indicators as possible.[19] The OECD is also heavily involved in the SDGs project, providing "snapshots" of OECD member countries and their performance on the SDGs (see, e.g., OECD, 2018). Once SDG 5's targets and indicators are finalized, governments and intergovernmental agencies have committed to collecting and reporting statistics annually (UN Statistics Division, UNSD, 2020; UN, 2015b; Allen et al., 2018). Sociodemographic statistics are to be collected on everyone, including marginalized populations (e.g., the homeless, migrants, minorities, and the underground economy) that are typically difficult or impossible to track.[20] Data collection so far has been difficult, and less than half the SDG indicators are high quality and many are completely missing.[21]

The activities above involving SDG 5 are steps in the "evidence collection" stage of an EBP process. As Figure 1 shows, once a goal has been defined, the policymaker selects targets and indicators, identifies data sources, and assigns data collection roles and responsibilities. Where multiple entities are involved, a central entity typically coordinates and oversees the process by creating ex ante protocols, manuals, and training materials to guide collection, and by ex post cleaning, merging, and analyzing the submitted data. The last step is the generation and dissemination of the results.

Evidence collection occurs early in the EBP process. Why, in May 2020, 5 years after the SDGs were launched, are UN agencies and Member States still in the evidence-gathering stage for SDG 5[22]? The slow progress appears to have multiple causes. UN Women (2018: 54) lists three problems: the uneven coverage of gender indicators across goals and targets, the absence of internationally agreed standards for data collection, and the uneven availability of gender statistics across countries and over time. The existing gender inequality

indexes, the GII and GGGI, were not closely tied to the broader targets and indicators of SDG 5 so new indexes had to be created; data collection for the new metrics has been hampered by lack of established methodologies for collecting and measuring the indicators.

A second reason why SDG 5 is still in the evidence collection phase is that national governments have failed to mainstream gender by not prioritizing gender statistics and/or by having weak and under-resourced statistical agencies (Thomas, Cordova Novion, de Haan, de León, Forest, & Iyer, 2018; UN, 2016; UN Women, 2018). Resource constraints on statistical agencies are likely to be even more important given the coronavirus pandemic and the global recession now underway.

A third causal factor is that selection and measurement of targets and indicators becomes a political activity when WPs are involved (Fukuda-Parr & McNeill, 2018). Politics can affect evidence collection in terms of, for example, determining priorities within complex targets, handling country composition changes over time, deciding which entities (national or international) are responsible for providing data, and allocating the financial costs of measurement (MacFeely, 2019b). Governments may also be unwilling to collect and provide gender-related data to an international agency, especially governments with poor records on gender equality, which may prefer to hide or tamper with their statistics, fearing the reputational risks (UN Women, 2018).

Our assessment is that SDG 5 needs more momentum. The evidence collection phase needs completion and the process needs to move forward on other steps in the EBP process. We therefore recommend that the UN and UN Women first commit explicitly and publicly to an EBP approach to SDG 5 and, second, prioritize completion of the evidence collection phase. The symbolic effect of a formal restatement of commitment to EBP we believe would be useful for all stakeholders in the process. The commitment would also raise awareness among stakeholders of the need to set up formal milestones and to prioritize completion of the evidence collection phase in the EBP process.

Recommendation #2: Expand the role of partnering in evidence collection

Our second policy recommendation is that the UN and UN Women expand their partnering arrangements so they can move faster on the evidence collection phase for SDG 5. We recommend expansion both in terms of partners and data sources.

The UN has worked for many years to ensure independence and impartiality of official statistics through the UN Fundamental Principles of Official Statistics and UN Statistical Quality Assurance Frameworks (UN-SQAF, 2018). Only official or accredited statistical agencies that supply official statistics are currently acceptable sources of data for the SDG indicators (MacFeely & Nastav, 2019).

Imagine a 2×2 matrix with Agency (official/accredited vs. non-official/non-accredited) on one side and Data Source (official/accredited vs. non-official/non-accredited) on the other side. The UN's current rules and procedures limit evidence collection for the SDGs to only one of the four cells in the 2×2 box: official/accredited agencies and official/accredited data sources. Given that the lack of data for many SDG indicators has slowed down completion of the data collection stage, we support proposals to – carefully – open up and include agencies and data sources in the other cells of the 2×2 matrix.

We recommend that UN statistical agencies should push forward rapidly on expanding their partnering to include non-official partners, as argued in MacFeely (2019b) and MacFeely and Nastav (2019). UN Women could, for example, build on the unrealized potential in DATA2x, the collaborative technical and advocacy platform at https://data2x.org/, which is estimating gaps in gender data, finding potential sources, and collecting data (Buvinic & Levine, 2015; Buvinic et al., 2014).

We also support recent proposals that unofficial data and statistics, both national and international, be certified and used as sources (MacFeely, 2019a; MacFeely & Nastav, 2019). Expansion of acceptable data metrics is particularly important for developing countries where policy capacity and resources are limited. "Big data" could prove to be a more cost-effective, efficient, and fine-grained data source than official sources and of better quality than survey data (MacFeely, 2019a). For example, DATA2x provides several fascinating case studies of gender inequalities, which were done using big data in developing countries (DATA2x, 2017, 2019).

In order to expand partnering in terms of agencies and data sources, the UN and its statistical agencies will need to be more open to using unofficial routes for SDG indicators (MacFeely, 2019a; MacFeely & Nastav, 74). We therefore also

support a rapid certification process whereby unofficial agencies and data sources can be accepted by the UN specifically for the SDG indicators.

Recommendation #3: Build public–private partnerships for the long-term stages of EBP

Generating high-quality good evidence for SDG 5 is only the first step in the EBP process. Governments must also mobilize and allocate resources to the achievement of gender equality. This means prioritizing gender-responsive investments, policies, and programs. Implementation, monitoring, and accountability are also needed. These next steps in the EBP process are likely to be very difficult for the UN and its Member States, for at least three reasons. First, short-run dislocation costs caused by the COVID-19 pandemic and global recession will make it difficult for governments to prioritize and fund gender equality initiatives. A second, longer-term reason is that the SDGs are a goal-based institution built on soft international law with little to no enforceability. "As a non-binding political commitment, the 2030 Agenda lacks enforceability. There are no defined consequences if countries fail to make serious efforts to meet the goals and targets" (UN Women, 2018: 257). The third reason is simply that SDG 5 is a WP where politics may trump evidence.

How can national governments move forward on the policy implementation, monitoring, and evaluation stages of the EBP cycle to address the WP of SDG 5? We argue that a key requirement will be that governments build successful partnerships with other stakeholders, in particular with multinational enterprises. Public–private partnerships (PPPs) with MNEs will be critical, for example, in developing and implementing firm-level policies for gender equality and empowerment of women in the workplace. Potential benefits from partnering with other stakeholders include knowledge-sharing and capacity building, mobilization of resources, and achievement of joint goals through collaboration (Bull & McNeill, 2019).

Our commitment to PPPs as a solution to the "evidence using" stages of an EBP approach to the 2030 Agenda builds on the three evidence "lenses" in Head (2008): scientific knowledge (the research-based knowledge of specialists), political knowledge (the know-how, analysis and judgment of political actors), and practical implementation knowledge (the wisdom and practices of government bureaucrats). Head (2008) argues that viewing evidence more broadly as three lenses can help policymakers address the complexity and conflict inherent in WPs. Our modification contends that PPPs can also bring together "three lenses" on evidence: scientific (academics, scientists), political (government policymakers and bureaucrats), and practical implementation (MNEs) knowledge, to deal with the WP of SDG 5. We explore this idea below.

Policy Recommendations for Multinational Enterprises

Recommendation #1: Commit to a global corporate social responsibility strategy

How MNE executives view the role of business in society has changed significantly over the years (Bull & Miklian, 2019; Eden, 2020). Definitions of social issues and corporate social responsibility (CSR) have broadened significantly as the social responsibility of business has evolved from its historical goal of "do no harm" to the more activist role of "doing good."[23] More recently, CSR scholars have found that MNEs, in particular the largest ones are moving from "doing good" to "going above and beyond" mandated levels of government social policies (Eden, 2020; Schlegelmilch & Szöcs, 2020). Simply meeting government CSR regulations is no longer viewed as a differentiating factor; MNEs must exceed mandated levels of social and environmental activities to build a reputation and positively affect their financial performance (Miller, Eden, & Li, 2020).

The 2030 Agenda provides an opportunity for MNE executives to rethink their CSR strategies. We recommend that MNE executives shift from viewing CSR as a stand-alone activity located in their marketing departments to recognizing that CSR is an activity that can and should be linked strategically and dynamically to the MNE's overall global strategy (Eden, 2020; Schlegelmilch & Szöcs, 2020). This new role should include building a commitment to global CSR into the MNE's goals, scope, rules of engagement, capabilities, and management systems. Making a commitment to at least 1 of the 17 SDGs should be a core component of the global corporate social strategy for every MNE.

Recommendation #2: Build public–private partnerships for the 2030 Agenda

The 2030 Agenda is a new form of global governance by goal setting (van Zanten & van Tulder, 2018; Bull & McNeill, 2019). The agenda offers an opportunity for MNEs to use PPPs as a vehicle for solidifying a new role for business in society. MNEs,

especially large MNEs with a global footprint, can be proactive agents of change that can serve society and address global problems (Kolk, Kourula, & Pisani, 2017; van Tulder, 2018). We therefore recommend that MNEs work pro-actively with UN agencies and national governments through PPPs to further the 2030 Agenda.

MNEs have been partnering with international organizations at least as far back as 1946 when the International Chamber of Commerce was given consultative status at the United Nations (Seitz, 2019). PPPs began to play a major role in MNE–state relations starting in the late 1990s (Bull & McNeill, 2019). For example, over the past 20 years, the primary forum for UN–business networking has been the UN Global Compact (UNGC), initiated in 1999 by UN Secretary-General Kofi Annan.[24] Other industry groups have also formed to support the SDGs; for example, the World Business Council for Sustainable Development (WDCSD, 2017, 2018) whose members are primarily large MNEs.

The 2030 Agenda provides many opportunities for MNEs to work with national governments and international NGOs. Van Zanten and van Tulder (2018: 226), in their study of MNE engagement with the SDGs, found that MNEs were using PPPs where the SDGs were complex and externally actionable, notably, SDG 1 (no poverty), SDG 2 (zero hunger), and SDG 4 (quality education). While PPPs and strategic alliances are costly, they enable resource collaboration and knowledge-sharing benefits that often cannot be achieved by firms on their own.

We recommend that MNEs use PPPs especially for the evidence collection stage for the SDGs, in particular for SDG 5. The UN has asked businesses to partner with governments and NGOs to build more robust statistical indexes for the SDGs (Business for 2030, 2020). There will be some targets and indicators where the private sector has better and more direct access to data than governments (e.g., SDG 5 data on wages and salaries, access to childcare and maternity leave policies, and share of women in management and leadership roles). These are areas where MNEs can play a powerful role in improving gender statistics in the evidence collection stage, particularly where statistical agencies are weak.

Recommendation #3: Adopt an EBP approach to the wicked problem of the 2030 Agenda

Our third recommendation is that MNEs adopt an EBP approach to selecting and implementing their own CSR strategy for engagement with the 2030 Agenda. Given that the SDGs are voluntary goals without formal binding commitments or penalties, each MNE has the flexibility to select from among the "menu" or "smorgasbord" of the 17 SDGs, prioritizing/ignoring and spending/not spending on the SDGs' multiple targets and indicators, as the firm's executives so choose (van Zanten & van Tulder, 2018).

Our proposed framework builds on the EBP policy cycle in Figure 2 with the key difference that the EBP approach is applied to MNE decision-making. We call our framework the SDG Materiality Matrix because a key component of the matrix is the analysis of both ex ante and ex post materiality of SDG targets to the MNE. The concept of materiality reflects the impact of a decision on a firm or actor; *materiality analysis* is designed to determine "what really matters to company sustainability performance, commitment and strategies" (Calabrese, Costa, Ghiron, & Menichini, 2017: 440; Bellantuono, Pontrandolfo, & Scozzi, 2018). Materiality analyses are typically carried out on an ex post basis, assessing the performance of the MNE's CSR activities. Here, we expand the concept to include both ex ante and ex post estimates of materiality of SDG targets to the MNE. We argue that MNE executives can use our SDG Materiality Matrix to develop their internal evidence-based policies, first, for creating and disseminating good evidence for the SDGs, and second, for policy design and implementation of SDG policies inside their organizations. We illustrate our SDG Materiality Matrix in Figure 3. Given the huge number of SDG targets and indicators, we recommend that SDG target selection by the MNE should be based on four factors: Quality of evidence for the target, Salience of the target to the MNE, Actionability of the target by the MNE, and Ethicality of the target for the MNE.

An EBP approach starts with the requirement of good evidence. Thus, *quality* – the "best available evidence" – for each SDG in terms of its targets and its indicators is the first factor in our SDG Materiality Matrix. MNE executives should start by mapping and assessing the 17 SDGs and their targets and indicators, for example, by consulting the most recent edition of the *E-Handbook on the SDGs* (UNSD, 2020). Summaries of available datasets and questions are also provided in GRI and the UN Global Compact (2017) and UNSD (2020).[25]

Our second factor is the *salience* of the evidence for the MNE. We argue that salience has two components: fit and materiality. In terms of *fit*,

we argue that MNE executives should prioritize SDG targets based on their relevance to and fit with the MNE's core business purpose (Szöcs & Schlegelmilch, 2020). The MNE's corporate goals for performance (market, financial, and social) and organizational legitimacy must also be considered (Donoher, 2017; Terpstra-Tong, 2017). The MNE should also consider existing in-house programs (e.g., CSR activities) and how the SDG target would fit with the MNE's capabilities (Szöcs & Schlegelmilch, 2020) and CSR programs (Schönherr, Findler, & Martinuzzi, 2017).

The second component of salience is *materiality* of the SDG target to the MNE. Here, we recommend that MNE executives estimate the expected benefits and costs (both private and social) from acting on the SDG target, the likely impact on firm performance, and in what ways the SDG target is relevant to the MNE's internal and external stakeholders. How, for example, would adopting the SDG target likely affect the MNE's exports and imports, employment, foreign direct investment (FDI), and global value chains?[26] Based on an assessment of fit and materiality, the MNE can determine which SDG targets have the greatest salience.

Our third factor is *actionability* by the MNE, which has three components: target scope, action type, and actors (we expect the three components to be interdependent). The first component is *target scope*; that is, whether the target is aimed at the regional/global, country, industry, or firm level. The second component is the *type of action* required by the SDG target. Addressing the target could require the MNE to engage in actionability through market-based actions (e.g., new products), operational actions (e.g., processes and value chains), or regulatory actions (e.g., standard-setting) (Szöcs & Schlegelmilch, 2020). The third component of actionability – *actors* – addresses which entity or entities are responsible for the actions and what roles they play. Both target scope and action type should affect the choice of actors and roles. For example, if the target is actionable at the firm level, the MNE can, but does not have to, move on its own to address the target. Targets aimed at the industry are likely to require partnering with firms in the same industry. Targets at the national or international levels are more likely to require PPPs and alliances with NGOs. Similarly, standard-setting actions may need industry alliances or PPPs, whereas new product launches can be done inside the MNE.

Our last factor in the SDG Materiality Matrix is *ethicality* of the target, which is the level of effort that the MNE chooses for the SDG target. Here, we consider two components: the standard or norm embedded in the SDG target and the MNE's choice of compliance level. First, SDG targets can be written as either proscriptive ("do not harm") or prescriptive ("do good") *norms or standards,* which

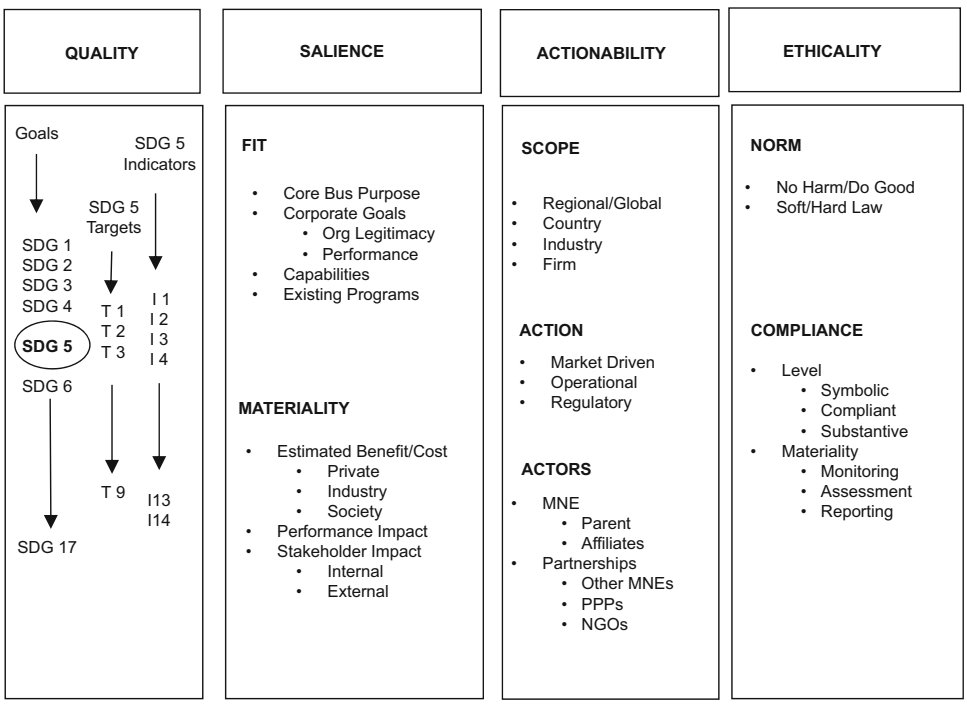

Figure 3 The SDG Materiality Matrix applied to SDG 5 Gender Equality.

can also be either voluntary or mandatory. Where the standard is mandatory, additional issues involve enforceability and sanctioning. Because the SDGs are "soft" law, only the first component (prescriptive or proscriptive behavior) is at issue here. The second component of ethicality is the MNE's chosen level of *compliance* with the standard: below (symbolic), at (compliant), or above (substantive) the norm (Miller et al., 2020). We also consider the *materiality* of the MNE's engagement with the SDG target in terms of monitoring, impact assessment, and reporting to governments and other stakeholders (Calabrese et al., 2017; Bellantuono et al., 2018). Note that, while the salience factor assesses materiality on an ex ante basis, the ethicality factor does the same on an ex post basis.

The four factors in our SDG Materiality Matrix for MNE engagement with the SDGs have their direct parallels in the EBP policy cycle recommended for government policymakers. As Figure 2 shows, given the best available evidence, policymakers should consult with stakeholders, interpret the evidence for the context, select their preferred option(s), determine the actors and their roles, implement the policies, and monitor and assess performance.

Recommendation #4: Use the SDG materiality matrix to mainstream gender equality

We hypothesize that MNEs will prioritize at least some SDG 5 targets and indicators as part of their global CSR strategy. Below, we apply the SDG Materiality Matrix to explore how this can be done.

In terms of selecting evidence based on its quality, we argue that businesses are best placed to "make a difference" if they select targets that focus on gender equality and empowering women *in the workplace*. In terms of indicators, we recommend that MNEs start with the GII and GGGI rather than the newer SDG gender indexes to assess and design policies for addressing the gender inequality gaps in their organizations. As we have argued above, the GII and GGGI are better measures (at least at present) of workplace gender inequalities than the newer indexes. MNEs should use metrics from the GII and GGGI to collect and monitor their own internal statistics on gender equality and empowerment.

In terms of salience, van Zanten and van Tulder (2018: 220) report that MNE engagement with the SDGs has been "particularly high" for SDG 5. Their result accords with other literature on the importance and salience of gender equality to business both in terms of fit and materiality; see, for example, Accenture (2019), Ike, Donovan, Topple and Masli (2019), PwC (2016), and WBCSD (2018).

In terms of actionability, we recommend that MNEs should focus, first, on the collection and dissemination of internal evidence on SDG 5 and, second, on designing and implementing internal policies to foster gender equality. MNEs that adopt SDG 5 as one of their key social strategies need to mainstream gender in their organizations and set up EBP management and reporting systems to ensure implementation (Thomas et al., 2018).

We also recommend that the MNE's workplace, for the purposes of SDG 5, be defined as including all domestic and foreign affiliates. This does not imply that the exact same gender equality policies must apply throughout the MNE group on a worldwide basis. Given the large differences across countries in both gender inequality antecedents and outcomes, a proactive and substantive strategy will be a difficult, expensive, and contentious undertaking (Terpstra-Tong, 2017). EBP predicts that top–down strategies based on global integration are likely to fail when "what works here does not work there." We therefore recommend a bottom–up, locally responsive approach. MNEs should build partnerships with key stakeholders at the country level (e.g., employees, governments, suppliers and buyers, civil society) and use the partnerships to develop appropriate country-based gender equality policies. MNEs should also address "missing links", such as second- and third-tier suppliers that often "fly below the radar" in (non)-compliance with the lead firm's and first-tier suppliers' CSR initiatives (Serdijn, Kolk, & Fransen, 2020). The commitment to mainstreaming gender equality throughout the MNE group should be, initially, to "go above and beyond" country-level requirements for gender equality and, in the longer term, to "lift all boats" to the *highest* common denominator across the MNE's local and foreign affiliates.

Lastly, in terms of ethicality, assuming MNEs do prioritize SDG 5, what norm should they adopt – do no harm or do good – and what level and materiality? Van Zanten and van Tulder (2018) hypothesized that good citizens (including MNEs) would be more likely to choose compliance with expected norms. Assuming that a "do good" standard is more costly than a "no harm" standard, the authors found that MNEs preferred SDG targets with "no harm" standards.[27] On the other hand, if MNE executives view CSR as a socially responsible

contract with their stakeholders, the executives are more likely to choose pro-active levels of engagement with the SDG targets (Eden, 2020; Schlegelmilch & Szöcs, 2020). There is also empirical evidence that MNEs that "go above and beyond" mandated CSR levels earn a positive reputation, which positively affects firm performance (Miller et al., 2020). We therefore expect MNEs to benefit from positive reputation gains that come from exceeding government mandates in terms of their global CSR strategies. We therefore recommend that MNEs commit publicly to exceeding national standards as they mainstream gender equality throughout their organizations.

Finally, adopting SDG 5 as a corporate social strategy also involves dissemination to others and reporting standards. The most common of these are CSR reports (Eden, 2020; Schönherr et al., 2017); other examples for how to build SDG 5 activities into corporate reports can be found in GRI and UNGC (2018). In addition, MNE CEOs are now being asked to publicly sign and post the seven Women's Empowerment Principles, designed to mainstream gender in business organizations (UN Women, 2019).[28] When CEOs of large MNEs commit to making gender equality a top strategic priority for their organizations, the positive signaling effect encourages others to follow suit.[29] The visibility of large MNEs encourages others to emulate their practices, creating a bandwagon effect (Van Zanten & van Tulder, 2018: 225). We therefore recommend that, as part of mainstreaming gender equality in their organizations, MNE CEOs sign on to the seven Women's Empowerment Principles.

In sum, our recommendations, drawn from the SDG Materiality Matrix, would mainstream gender equality throughout the MNE group and send a strong signal to stakeholders and other MNEs.

CONCLUSION

Gender equality is 1 of 17 "wicked problems" in the 2030 Agenda. The issue fulfills all the criteria: systemic, complexly interrelated, with material involvement by multiple actors at multiple levels across multiple countries (van Tulder, 2018). There is wide recognition that better evidence is needed to foster women's voice and agency, and to meet the UN challenge to leave no one behind (Thomas et al., 2018; UN Women, 2018). Our paper explores how insights from EBP can be useful for tackling the WP of SDG 5 and the 2030 Agenda.

Our paper makes several contributions to the IB literature. First, to the best of our knowledge, there have been no articles in *Journal of International Business Policy* (*JIBP*) or *Journal of International Business Studies* (*JIBS*) on WPs or EBP. A keyword search of *JIBS* and *JIBP* failed to find a single article on either topic.[30] We believe both frameworks offer much value added for thinking about IB policies and problems. IB theory has a long history of adopting insights from other disciplines. We hope that our work will encourage other IB scholars to use these theoretical frameworks for their own research topics.

We contribute to the EBP literature by expanding it in three different contexts: first, from the single-country (domestic) level to the multi-country level of the 2030 Agenda; second, by applying EBP to the problem of gender inequality; and third, by developing a EBP for multinational enterprises in our new SDG Materiality Matrix. In addition, we contribute to the WP literature by applying it to SDG gender equality, building on O'Brien et al. (2017). We also bring together two literatures – WPs and EBP – that historically have had little to do with one another, and show that they can usefully inform each other. Our work here builds on Daviter (2019) and Head (2019).

We contribute to the literature on the SDGs, and SDG 5 in particular, by showing how high-quality evidence and EBP can assist in the attainment of the 2030 Agenda. Our paper also contributes to the literature on gender equality by exploring differences between five gender inequality indexes, building on Eden and Gupta (2017), and assessing their relative appropriateness as evidence in an EBP approach to SDG 5. Our paper also addresses the research agenda items outlined by Witte and Dilyard (2017), including how government policies on the SDGs affect MNE strategies and predict which firms will engage and how they will engage with the SDGs.

A key innovation in our paper is the introduction of a new EBP framework for MNEs, the SDG Materiality Matrix, based on four factors (Quality, Salience, Actionability, and Ethicality), which MNEs can use to select and implement policies for their preferred SDG targets. We also contribute to the CSR literature by expanding it to encompass the SDGs, applying CSR insights into our SDG Materiality Matrix, and developing the concepts of ex ante and ex post materiality.

Going forward, we argue that IB scholars need to better understand how, why, and where MNEs are

involved in the SDG process. Our SDG Materiality Matrix could be used as the framework for case studies of MNE interactions with other SDGs. Lastly, our paper could be expanded to discuss the role of evidence and EBP for the SDGs from the perspective of other organizations, such as business schools and professional associations. To what extent are they involved in the 2030 Agenda and what policy processes are they adopting to ensure that no one is left behind and everyone's voice is heard?

ACKNOWLEDGEMENTS

This paper was first presented at a *JIBP* Workshop at the Academy of International Business meetings in Copenhagen in June 2019. We thank the *JIBP* Special Issue Editors, especially Rob van Tulder, and the anonymous reviewers, for their thoughtful, detailed, and helpful comments on the paper. We thank Reilly Smith for her assistance with the datasets. Any remaining errors are the authors.

NOTES

[1]For good literature reviews, see the August 2012 special issue of *Journal of Agricultural and Environmental Ethics* introduced by Whyte and Thompson (2012), and the December 2019 special issue of *Policy and Society* introduced by Termeer et al. (2019).

[2]In an interesting test of Newman and Head (2017), Peters and Tarpey (2019) surveyed researchers on six policy problems, asking respondents to assess each problem's wickedness according to the ten propositions. The respondents viewed all six as having some degree of wickedness, which depended on complexity and conflict.

[3]See, for example, Hudson, Ballif-Spanvill, Caprioli and Emmett (2012), Klugman, Hanmer, Twigg, Hasan, McCleary Sills and Santamaria Bonilla (2014), Hudson (2015), McKinsey Global Institute (2015), and Hudson et al. (2020).

[4]To create the index, all data are converted to indicator ratios that are truncated at 1; scores between 0 and 1 for each target are calculated using weighted averages of the individual indicators; and the final score is calculated as an unweighted average of the four targets (World Economic Forum, WEF, 2020: 17: 45–48).

[5]Piper (2019) argues that Nicaragua's fast rise up the GGGI country rankings was due to the GGGI ignoring national levels of economic development; "if the situation is bad for both genders, but similarly bad, the Index will report (accurately) a narrow gender gap" (Piper, 2019: 1395). Constantine (2017) reached a similar conclusion. Piper (2019) also tested whether Nicaraguan women reported greater life satisfaction when the gender gap narrowed; once GDP growth rates were accounted for, no differences between Nicaragua and other countries remained.

[6]Measures of violence against women, such as forced marriages, gender-selective abortions, human trafficking, and sexual violence in armed conflicts, are missing. Both indexes are silent on discriminatory laws and stereotypes pervasive in many countries that deter gender equality. In addition, women's rights to economic resources, as well as access to ownership and control over land and other forms of property, financial services, inheritance, and natural resources, are typically absent from the indexes.

[7]A look at the indicators in the SDSN's gender index provides some evidence on the difficulties of capturing "no one left behind." Of the 14 target indicators for SDG 5 (i.e., 5.1.1 to 5.c.1 in Table 1), the actual measures used appear to be a combination of GII and GGGI indicators.

[8]Barnat, MacFeely and Peltola (2019a, b) conducted a principal component analysis of the GII, GGGI, and SIGI; the indexes clustered on four components: education and women's social conditions, women's economic and labor market participation, women's political participation, and health.

[9]Histories and analyses of EBP can be found in, for example, Brownson, Gurney and Land (1999), Nutley et al. (2003), Davies (2004), Sutcliffe and Court (2006), Head (2010), Jayaraman and Rocholl (2017), Saltelli and Giampietro (2017), Haskins (2018), and Hewlett Foundation (2018).

[10]EU (2017a) provides a useful case study of efforts by the European Institute for Gender Equality to launch, implement, and evaluate the gender mainstreaming platform.

[11]On EBP, see, for example, Davies, Nutley and Smith (2000), Sutcliffe and Court (2005), Davies (2012), Strehlenert, Richter-Sundberg, Nystrom and Hasson (2015), and European Commission (2017b).

[12]Newman and Head (2017) provide mini-case studies of climate change, genetically modified foods, and hydraulic fracturing. In all three cases,

there was abundant scientific evidence on which almost all scientists agreed. Despite the evidence, public policy was driven by political dynamics and embedded biases. The authors conclude that good evidence alone cannot address wicked problems; rather, the cure lies in "untangling the political dynamics and values-based discourse" (Newman & Head, 2017: 424).

[13]The well-known phrase "lies, damned lies, and statistics" also points to a broader critique of evidence: the crisis of reproducibility, integrity, and legitimacy that currently plagues academic research. Ethical pitfalls bedevil academia, especially in research, as evidenced by growing numbers of shoddy research practices, article retractions, and predatory journals (Eden et al., 2018; Nielsen, Eden, & Verbeke, 2020). The crisis in science creates a crisis of trust; how can policymakers trust the research findings when scientists themselves are behaving badly?

[14]The problem of weak academic-policymaker linkages may be changing. A recent example is Responsible Research in Business and Management, a global network of business and management faculty that encourages research on societal issues with practical relevance for policymakers and businesses (Responsible Research in Business and Management, Community, 2017). Some academic associations are also setting up policy journals, such as the Academy of International Business's *Journal of International Business Policy*, specifically to encourage scholars to engage in scholarly policy-based research.

[15]For examples, see Davies et al. (2000), Nutley et al. (2003), Mulgan (2005), Shepherd (2007), Janssen and Forbes (2014), and Commission on Evidence-Based Policymaking (2017).

[16]A useful example of how internationally endorsed EBP can fail in developing countries is Behague, Tawiah, Rosato, Some and Morrison's (2009) study of the conflict between the neonatal and maternal health care targets in the Millennium Development Goals (UN, 2015a). Developing country governments were asked to collect statistics and adopt EBP for neonatal health care, but, hampered by lack of policy capacity and resources, ended up diverting resources from maternal to neonatal health care. The authors concluded that international norms ignored local policy needs, weakening the benefits of EBP.

[17]The assistance of four international organizations (the European Commission, World Bank, OECD, and UNDP) proved to be a mixed blessing since "conflicting advice received from multiple donors and external advisers only provides an incentive for playing the system and producing inconsistent policy formulas" (Bartlett, 2013:464).

[18]Using EBP at the regional level is explored in Lofstedt and Schlag's (2017) account of how European governments handled the heated debate over whether to allow the use of the chemical Bisphenol A in Europe; the authors concluded that policy decisions were ideological, not evidence- or risk-based.

[19]The data collection and analysis are part of the UN Women Count Project established in 2016 to "improve the production and use of gender data and help countries monitor the SDGs from a gender perspective" (Seck & Maskey, 2019: 3).

[20]UN Women (2019a: 5) estimates that socio-demographic data are missing for up to 350 million people.

[21]The SDG indicators are classified into three tiers: tier 1 (highest quality = established methodology with widely available data), tier 2 (mid-quality = established methodology but data not easily available) and tier 3 (lowest quality = no internationally accepted methodology). As of April 2019, only 44% of the SDG indicators were tier 1 and 15% were tier 3 (MacFeely & Nastav, 2019: 311).

[22]The Statistics Division of the UN Department of Economic and Social Affairs maintains a list of proposed changes to the SDG indicators at https://unstats.un.org/sdgs/iaeg-sdgs/2020-comprev/UNSC-proposal/ (last accessed 15 May 2020).

[23]See, for example, Eden (2020), Kolk (2016), Schlegelmilch and Szöcs (2020), Schönherr et al. (2017), van Tulder (2018), and van Zanten and van Tulder (2018).

[24]The UNGC now has more than 14,000 member firms; its members committed originally to 10 principles related to human rights, labor, the environment, and anti-corruption. In 2015, the UNGC also signed onto the 2030 Agenda and developed its own set of 10 Action Platforms to encourage networking between businesses, governments, and NGOs. (https://www.unglobalcompact.org/sdgs/action-platforms).

[25]For example, pages 206–207 of the GRI and UNGC (2017) Report assess which SDG 5 targets are most closely and least closely tied to business, and then provide a detailed assessment of the best available indicators for each target. Detailed commentaries are provided on each SDG 5 indicator with suggestions for gender statistics that the MNE

could collect itself (see GRI, and UN Global Compact, UNGC, 2017: 59–68).

[26]The OECD has created a new set of FDI Qualities Indicators designed to measure the impact of FDI on sustainable development. OECD (2019, Ch.5) takes a "women at work" approach, arguing that there are four channels through which FDI can affect SDG 5: employment, wages, top management positions, and entrepreneurship (see also UNCTAD, 2014). The OECD's new FDI qualities index may prove useful for tracking SDG performance for MNEs that adopt SDG 5 as their corporate social strategy.

[27]There is also some evidence that MNEs may choose to make a symbolic commitment to the SDGs, or, worse, use the process to impede government regulation by lobbying (Seitz, 2019). The "scope, intent and impact of business' involvement in the SDGs is often "vague and hard to measure", according to Abshagen, Cavazzini, Graen and Oberland (2018: 7). The lack of a common and systematic approach to interactions between MNEs and UN agencies can also create risks and conflicts of interest.

[28]The principles were developed under the WE EMPOWER program, a program funded by the European Union and implemented by UN Women and the International Labor Organization (https://www.weps.org/join).

[29]For example, at the January 2020 Davos meetings of the World Economic Forum, the UNGC together with SAP and Accenture announced, "SDG Ambition", challenging its members to raise their level of commitment to the 2030 Agenda (UN Global Compact, UNGC, 2020).

[30]Every scholar stands on the shoulders of earlier scholars. We owe a particular debt to Robert van Tulder for his seminal work applying insights from the WP literature to the 2030 Agenda (van Tulder, 2018).

REFERENCES

Abshagen, M.-L., Cavazzini, A., Graen, L., & Obenland, W. 2018. *Highjacking the SDGs? The private sector and the sustainable development goals.* Berlin: Global Policy Forum. https://www.globalpolicy.org/images/pdfs/GPFEurope/Hijacking_the_SDGs.pdf. Accessed 2 July 2020.

Accenture. 2019. *The decade to deliver: A call to business action. The United Nations Global Compact—Accenture Strategy CEO Study on Sustainability 2019.* United Nations Global Compact.

Alford, J., & Head, B. W. 2017. Wicked and less wicked problems: A typology and a contingency framework. *Policy and Society*, 36(3): 397–413.

Allen, C., Metternicht, G., & Wiedmann, T. 2018. Initial progress in implementing the Sustainable Development Goals (SDGs): A review of evidence from countries. *Sustainability Science*, 13: 1453–1467.

Alon, T., Doepke, M., Olmstead-Rumsey, J., & Tertilt, M. 2020. *The impact of COVID-19 on gender equality*. NBER Working Paper No. 26947, April. https://www.nber.org/papers/w26947. Accessed 2 July 2020.

Andrus, J. K., Jauregui, B., De Oliveira, L. H., & Ruiz Matus, C. 2011. Challenges to building capacity for evidence-based new vaccine policy in developing countries. *Health Affairs*, 30(6): 1104–1112.

Bannink, D., & Trommel, W. 2019. Intelligent modes of imperfect governance. *Policy and Society*, 38(2): 198–217.

Barnat, N., MacFeely, S., & Peltola, A. 2019a. Comparing global gender inequality indices: How well do they measure the economic dimension? *Journal of Sustainability Research*, 1: e190016. https://doi.org/10.20900/jsr20190016. Accessed 2 July 2020.

Barnat, N., MacFeely, S., & Peltola, A. 2019b. Comparing Global Gender Inequality Indices: What can they tell us about development? In *Presented at the UN Economic Commission for Europe Conference of European Statisticians Work Session on Gender Statistics*, Neuchatel, Switzerland, 15–17 May. Working Paper No. 16.

Barro, R. J. & Lee, J.-W. 2016.*Education matters. Global schooling gains from the 19th to the 21st century*, Oxford: Oxford University Press.

Bartlett, W. 2013. Obstacles to evidence-based policy-making in the EU Enlargement Countries: The case of skills policies. *Social Policy and Administration*, 47(4): 451–467.

Behague, D., Tawiah, C., Rosato, M., Some, T., & Morrison, J. 2009. Evidence-based policy-making: The implications of globally-applicable research for context-specific problem-solving in developing countries. *Social Science and Medicine*, 69: 1539–1546.

Bellantuono, N., Pontrandolfo, P., & Scozzi, B. 2018. Guiding materiality analysis for sustainability reporting: The case of the agri-food sector. *International Journal of Technology, Policy and Management*, 18(4): 336–359.

Best, J. 2012. *Damned lies and statistics: Untangling numbers from the media, politicians and activists.* Los Angeles: University of California Press.

Branisa, B., Klasen, S., Ziegler, M., Drechsler, D., & Jutting, J. 2014. The institutional basis of gender inequality: The social institutions and gender index (SIGI). *Feminist Economics*, 20(2): 29–64.

Breuer, A., Janetschek, H., & Malerba, D. 2019. Translating sustainable development goal (SDG) interdependencies into policy advice. *Sustainability*, 11(April): 1–20.

Brownson, R. C, Gurney, J. G., & Land, G. H. 1999. Evidence-based decision making in public health. *Journal of Public Health Management Practice*, 5(5): 86–97.

Bull, B., & McNeill, D. 2019. From market multilateralism to governance by goal setting: SDGs and the changing order of partnerships in a new global order. *Business and Politics*, 21(4): 464–486.

Bull, B., & Miklian, J. 2019. Towards global business engagement with development goals? Multilateral institutions and the SDGs in a changing global capitalism. *Business and Politics*, 21(4): 445–463.

Business for 2030. 2020. *UN metrics and indicators for the SDGs. Forging a path for business in the UN 2030 Development Agenda*. http://www.businessfor2030.org/metrics-indicators. Accessed 2 July 2020.

Buvinic, M., & Levine, R. 2015. What is wrong with data on women and girls? Data2x: Partnering for a gender data revolution. November. 8 pp.
Buvinic, M., Furst-Nichols, R., & Koolwal, G. 2014. Mapping gender data gaps, DATA2x: Partnering for a gender data revolution, March. 8 pp. https://data2x.org/resourcecenter/mapping-gender-data-gaps/. Accessed 2 July 2020.
Calabrese, A., Costa, R., Ghiron, N. L., & Menichini, T. 2017. Materiality analysis in sustainability reporting: A method for making it work in practice. *European Journal of Sustainable Development*, 6(3): 439–447.
Caplan, N. 1979. The two-communities theory and knowledge utilization. *American Behavioral Scientist*, 22(3): 459–470.
Commission on Evidence-Based Policymaking. 2017. *The promise of evidence-based policymaking. Report of the Commission on Evidence-Based Policymaking*, September 7. Washington, DC: U.S. Government.
Constantine, G. 2017. Does Nicaragua match up to its global gender gap report ranking on women's rights? *Eye On Latin America*, November 6. https://eyeonlatinamerica.com/2014/11/06/nicaragua-global-gender-gap-report-2014/. Accessed 2 July 2020.
Court, J., & Young, J. 2006. Bridging research and policy in international development: An analytical and practical framework. *Development in Practice*, 16(1): 85–90.
Crowley, K., & Head, B. 2017. The enduring challenge of "wicked problems": Revisiting Rittel and Webber. *Policy Science*, 50: 537–547.
Data for African Development Working Group. 2014. *Delivering on the data revolution in Sub-Saharan Africa.* Center for Global Development and African Population and Health Research Center. https://www.cgdev.org/publication/delivering-data-revolution-sub-saharan-africa-0. Accessed 2 July 2020.
DATA2x. 2017. *Big data and the well-being of women and girls: Applications on the social science frontier*, April. https://data2x.org/resource-center/big-data-and-the-wellbeing-of-women-and-girls/. Accessed 2 July 2020.
DATA2x. 2019. *Big data, big impact? Towards gender-sensitive data systems*, November. https://data2x.org/resource-center/big-data-report/. Accessed 2 July 2020.
Davies, P. T. 2004. *Is evidence-based government possible?* Jerry Lee Lecture. https://www.eldis.org/document/A18705. Accessed 2 July 2020.
Davies, P. T. 2012. The state of evidence-based policy evaluation and its role in policy formation. *National Institute Economic Review*, 219(January): R41–R52.
Davies, H. T. O., Nutley, S. M., & Smith, P. C. 2000. *What works: Evidence-based policy and practice in public services.* Bristol: The Policy Press.
Daviter, F. 2017. Coping, taming or solving: Alternative approaches to the governance of wicked problems. *Policy Studies*, 38(6): 571–588.
Daviter, F. 2019. Policy analysis in the face of complexity: What kind of knowledge to tackle wicked problems? *Public Policy and Administration*, 34(1): 62–82.
De Marchi, G., Lucertini, G., & Tsoukiàs, A. 2016. From evidence-based policy making to policy analytics. *Annals of Operation Research*, 236: 15–38.
Donoher, W. 2017. The multinational and the legitimation of sustainable development. *Transnational Corporations*, 24(3): 49–60.
Eden, L. 2020. The social responsibility of multinationals: From an after-thought to center stage. In B. Schlegelmilch and I. Szőcs (Eds.) *Rethinking business responsibility in a global context—Challenges to corporate social responsibility, sustainability and ethics*. Springer Series: CSR, sustainability, ethics and governance. Cham: Springer.
Eden, L., & Gupta, S. F. 2017. Culture and context matter: Gender in international business and management. *Cross Cultural and Strategic Management*, 24: 194–210.
Eden, L., Lund Dean, K., & Vaaler, P. 2018. *The ethical professor: A practical guide to research, teaching, and professional life.* New York: Routledge, Taylor and Francis.
Equal Measures 2030 (EM 2030). 2019. *Harnessing the power of data for gender equality: Introducing the 2019 EM2030 gender index.* Surrey: Equal Measures 2030.
European Commission. 2017a. *2017 Report on Equality between Women and Men in the EU.* Luxembourg: Publications Office of the European Union.
European Commission. 2017b. *Quality of public administration: A toolbox for practitioners*. Luxembourg: Publications Office of the European Union.
Fukuda-Parr, S. 2014. Global goals as a policy tool: Intended and unintended consequences. *Journal of Human Development and Capabilities*, 15(2–3): 118–131.
Fukuda-Parr, S., & McNeill, D. 2018. Knowledge and power in setting and measuring SDGs. *Global Policy Journal*, 10 (Supplement 1): 5–15.
Garten, M. 2020. Coronavirus update: COVID-19 likely to cost economy $1 trillion during 2020, says UN Trade Agency. *UN News*, March 9. https://news.un.org/en/story/2020/03/1059011. Accessed 2 July 2020.
Government of Malawi. 2016. *Guidelines for evidence use in decision-making in the health sector in Malawi*. Ministry of Health.
GRI, and UN Global Compact. 2017. *Business reporting on the SDGs: An analysis of the goals and targets.* https://www.globalreporting.org/information/SDGs/Pages/Reporting-on-the-SDGs.aspx. Accessed 2 July 2020.
GRI, and UN Global Compact (UNGC). 2018. *Business reporting on the SDGs: Integrating the SDGs into corporate reporting, a practical guide.* https://www.unglobalcompact.org/library/5628. Accessed 2 July 2020.
Hantrais, L., Lenihan, A. T., & MacGregor, S. 2015. Evidence-based policy: Exploring international and interdisciplinary insights. *Contemporary Social Science: Journal of the Academy of Social Science* 10(2): 101–113.
Haskins, R. 2018. Evidence-based policy: The movement, the goals, the issues, the promise. *Annals, AAPSS*, 678: 8–37.
Head, B. W. 2008. Three lenses of evidence-based policy. *The Australian Journal of Public Administration*, 67(1): 1–11.
Head, B. W. 2010. Reconsidering evidence-based policy: Key issues and challenges. *Policy and Society*, 29(2): 77–94.
Head, B. W. 2019. Forty years of wicked problems literature: Forging closer links to policy studies. *Policy and Society*, 38(2): 180–197.
Head, B. W., & Alford, J. 2015. Wicked problems: Implications for public policy and management. *Administration and Society*, 47(6): 711–739.
Hewlett Foundation. 2018. *Evidence-informed policymaking strategy*. Global Development and Population Program, March.
Howes, S., Betteridge, A., Sause, L., & Ugyel, L. 2017. *Evidence-based policy making in the Tropics: Are developing countries different?* Development Policy Centre Discussion Paper No. 59, July 26. ANU Research Paper Series. Canberra, Australia: Crawford School of Public Policy, Australian National University. https://ssrn.com/abstract=3009196. Accessed 2 July 2020.
Hudson, V. 2015. Leveling the field: A global inventory of gender equality for women. *World Politics Review*, September 1. https://www.worldpoliticsreview.com/articles/16572/leveling-the-field-a-global-inventory-of-gender-equality-for-women. Accessed 2 July 2020.
Hudson, V., Ballif-Spanvill, B., Caprioli, M., & Emmett, C. 2012. *Sex and world peace.* New York: Columbia University Press.
Hudson, V., Bowen, D. L., & Nielsen, L. 2020. *The First Political Order: How sex shapes governance and national security worldwide*. New York: Columbia University Press.
Ike, M. Donovan, J. D., Topple, C., & Masli, E. K. 2019. The process of selecting and prioritizing corporate sustainability

issues: Insights for achieving the sustainable development goals. *Journal of Cleaner Production*, 236: 1–13.
Janssen, T., & Forbes, S. 2014. The use of official statistics in evidence-based policymaking in New Zealand. In K. Makar, B. de Sousa & R. Gould (Eds.) *Sustainability in statistics education. Proceedings of the ninth international conference on teaching statistics (ICOTS9)*, Flagstaff, Arizona, USA, July 2014. Voorburg: International Statistical Institute. iase-web.org.
Jayaraman, R., & Rocholl, J. 2017. Research-based policy advice to the G20. G20 Insights. G20 Germany 2017 Think 20 Dialogue, May 29. https://www.g20-insights.org/policy_briefs/research-based-policy-advice-g20/. Accessed 2 July 2020.
Klugman, J., Hanmer, L., Twigg, S., Hasan, T., McCleary Sills, J., & Santamaria Bonilla, A. J. 2014. *Voice and agency: Empowering women and girls for shared prosperity*. Washington, DC: World Bank.
Kolk, A. 2016. The social responsibility of international business: From ethics and the environment to CSR and sustainable development. *Journal of World Business*, 51: 23–34.
Kolk, A., Kourula, A., & Pisani, N. 2017. Multinational enterprises and the Sustainable Development Goals: What do we know and how to proceed? *Transnational Corporations*, 24(3): 9–32.
Langer, L., Tripney, J., & Gough, D. 2016. *The science of using science: Researching the use of research evidence in decision-making*. London: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London.
Lederer, E. M. 2020. *UN chief says COVID-19 is worst crisis since World War II*. Associated Press. https://abcnews.go.com/US/wireStory/chief-covid-19-worst-crisis-world-war-ii-69905340. Accessed 2 July 2020.
Lofstedt, R., & Schlag, A. K. 2017. Looking back and going forward: What should the new European Commission do in order to promote evidence-based policy-making? *Journal of Risk Research*, 20(11): 1359–1378.
Lubua, E. W., & Maharaj, M. S. 2012. ICT policy and e-transparency in Tanzania. In P. Cunningham & M. Cunningham (Eds.) *IST-Africa 2012 conference proceedings*. IICM International Information Management Corporation.
Lugo-Gil, J., Jean-Baptiste, D., & Livia Frasso Jaramillo, L. 2019. Use of evidence to drive decision-making in government. *Mathematica Policy Research*, February 20, 80pp. https://www.mathematica.org/-/media/publications/pdfs/education/2019/useofevidenceindecisionmaking.pdf.
MacFeely, S. 2019a. The big (data) bang: Opportunities and challenges for compiling SDG indicators. *Global Policy*, 10(1): 121–133.
MacFeely, S. 2019b. *The political economy of measuring the sustainable development goals.* UNCTAD Research Paper No. 32. UNCTAD/SER.RP/2019/4. Geneva: United Nations.
MacFeely, S., & Nastav, B. 2019. "You say you want a [data] revolution": A proposal to use unofficial statistics for the SDG Global Indicator Framework. *Statistical Journal of the IAOS,* 35: 309–327.
McCall, R., & Burge, J. 2016. Untangling wicked problems. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 30: 200–210.
McKinsey Global Institute. 2015. *The power of parity: How advancing women's equality can add $12 trillion to global growth,* September. McKinsey Global Institute. https://www.mckinsey.com/global-themes/employment-and-growth/how-advancing-womens-equality-can-add-12-trillion-to-global-growth. Accessed 2 July 2020.
Miller, S. R., Eden, L., & Li, D. 2020. CSR reputation and firm performance: A dynamic approach. *Journal of Business Ethics*, 16(3): 619–636. https://doi.org/10.1007/s10551-018-4057-1. Accessed 2 July 2020.
Monaghan, M., Pawson, R., & Wicker, K. 2012. The precautionary principle and evidence-based policy. *Evidence and Policy*, 8(2): 171–191.
Mulgan, G. 2005. Government, knowledge and the business of policy making: The potential and limits of evidence-based policy. *Evidence and Policy*, 1(2): 215–226.
Nature. 2020. Get the Sustainable Development Goals back on track. *Nature*, 577(January 2): 7–8. https://doi.org/10.1038/d41586-019-03907-4. Accessed 2 July 2020.
Newman, J., & Head, B. W. 2017. Wicked tendencies in policy problems: Rethinking the distinction between social and technical problems. *Policy and Society*, 36(3): 414–429.
Newman, J., Cherney, A., & Head, B. W. 2016. Do policy makers use academic research? Reexamining the "Two Communities" theory of research utilization. *Public Administration Review*, 76(1): 24–32.
Newman, J., Cherney, A., & Head, B. W. 2017. Policy capacity and evidence-based policy in the public service. *Public Management Review*, 19(2): 157–174.
Ney, S., & Verweij, M. 2015. Messy institutions for messy problems: How to generate clumsy solutions? *Environment and Planning C*, 33: 1679–1696.
Nielsen, B. B., Eden, L., & Verbeke, A. 2020. Research methods in international business: Challenges and advances. In L. Eden, B. B. Nielsen & A. Verbeke (Eds.) *Research methods in international business.* JIBS Special Collections: 3–41. London: Palgrave Macmillan.
Nutley, S., Davies, H., & Walter, I. (2003). Evidence-based policy and practice: Cross-sector lessons from the United Kingdom. *Social Policy Journal of New Zealand*, 20: 29–48.
O'Brien, K., Fitzsimmons, T. W., Crane, M., & Head, B. 2017. Workplace gender inequality as a wicked problem: Implications for research and practice. In *Academy of Management annual meeting proceedings*, Briarcliff Manor, NY.
OECD. 2017. *The pursuit of gender equality: An uphill battle*. Paris: OECD.
OECD. 2018a. *Development Co-operation Report 2018: Joining forces to leave no one behind.* Paris: OECD.
OECD. 2018b. Gender equality: Gender, institutions and development database 2009 (GID-DB). *OECD Social and Welfare Statistics* (database). https://doi.org/10.1787/f3947736-en. Accessed 2 July 2020.
OECD. 2019. *SIGI 2019 Global Report: Transforming challenges into opportunities.* Paris: OECD.
Oliver, K., & Cairney, P. 2019. The dos and don'ts of influencing policy: A systematic review of advice to academics. *Palgrave Communications*, 5:21. https://doi.org/10.1057/s41599-019-0232-y. Accessed 2 July 2020.
Parkhurst, J. 2017. *The politics of evidence: From evidence-based policy to the good governance of evidence.* Routledge studies in governance and public policy. Abingdon: Routledge.
Peace Corps. 2020. *Global issues: Gender equality and women's empowerment*. https://www.peacecorps.gov/educators/resources/global-issues-gender-equality-and-womens-empowerment/. Accessed 2 July 2020.
Permanyer, I. 2013a. A critical assessment of the UNDP's gender inequality index. *Feminist Economics*, 19(2): 1–32.
Permanyer, I. 2013b. Are UNDP indices appropriate to capture gender inequalities in Europe? *Social Indicators Research*, 110: 927–950.
Peters, B. G., & Tarpey, M. 2019. Are wicked problems really so wicked? Perceptions of policy problems. *Policy and Society*, 38(2): 218–236.
Piper, A. 2019. An investigation into the reported closing of the Nicaraguan gender gap. *Social Indicators Research*, 144: 1391–1413.
PwC. 2016. *Make it your business: Engaging with the sustainable development goals*. www.pwc.com/sdg. Accessed 2 July 2020.
Rademacher, W. J. 2019. Governing-by-the-numbers/statistical governance: Reflections on the future of official statistics in a digital and globalised society. *Statistical Journal of the IAOS*, 35(4): 519–537.
Responsible Research in Business and Management, Community for (RRBM). 2017. *A vision of responsible research in business*

and management: Striving for useful and credible knowledge. Position Paper, November 22. https://rrbm.network/position-paper/. Accessed 2 July 2020.
Richards, G. W. 2017. How research-policy partnerships can benefit government: A win–win for evidence-based policy-making. *Canadian Public Policy/Analyse de Politiques*, June: 165–170.
Rittel, H. 1972. On the planning crisis: Systems analysis of the 'First and Second Generations'. BedriktsØkonomen NR.8: 390.
Rittel, H. W. J., & Webber, M. M. 1973. Dilemmas in a general theory of planning. *Policy Sciences*, 4: 155–169.
Sachs, J., Schmidt-Traub, G., Kroll, C., Lafortune, G., & Fuller, G. 2019. *Sustainable Development Report 2019*. New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN).
Saltelli, A., & Giampietro, M. 2017. What is wrong with evidence based policy, and how can it be improved? *Futures*, 91: 62–71.
Schlegelmilch, B., & Szöcs. I. (Eds.) 2020. *Rethinking business responsibility in a global context: Challenges to corporate social responsibility, sustainability and ethics.* Uitgever: Springer.
Schmidt-Traub, G., Kroll, C., Teksoz, K., Durand-Delacre, D., & Sachs, J. D. 2017. National baselines for the Sustainable Development Goals assessed in the SDG Index and Dashboards. *Nature Geosciences*, 10: 547–555. https://doi.org/10.1038/ngeo2985. Accessed 2 July 2020.
Schönherr, N., Findler, F., & Martinuzzi, A. 2017. Exploring the interface of CSR and the sustainable development goals. *Transnational Corporations*, 24(3): 33–42.
Scott, C. 2005. Measuring up to the measurement problem: The role of statistics in evidence-based policymaking. In *Proceedings of the 2005 CBMS Network meeting*: 35–92.
Seck, P., & Maskey, S. 2019. Use data to take the guesswork out of policymaking. UN Women, September 24. https://data.unwomen.org/features/use-data-take-guesswork-out-policymaking. Accessed 2 July 2020.
Seitz, K. 2019. *A Fatal Attraction? Business engagement with the 2030 Agenda*, February. Berlin: Global Policy Forum. https://www.globalpolicy.org/images/pdfs/A_Fatal_Attraction_online.pdf. Accessed 2 July 2020.
Selin, H. 2015. The risk of UN's Sustainable Development Goals: Too many goals, too little focus. *The Conversation*, September 25. http://theconversation.com/the-risk-of-uns-sustainable-development-goals-too-many-goals-too-little-focus-48083. Accessed 2 July 2020.
Serdijn, M., Kolk, A., & Fransen, L. (2020). Uncovering missing links in global value chain research—And implications for corporate social responsibility and international business. *Critical Perspectives on International Business*. https://doi.org/10.1108/cpoib-01-2020-0002. Accessed 2 July 2020.
Shepherd, J. 2007. The production and management of evidence for public service reform. *Evidence and Policy*, 3(2): 231–251.
Sotsky, J. G., Shibuya, S., Kolovich, L., & Kebhaj, S. 2016. *Trends in gender equality and women's advancement.* Working Paper No. WP/16/21, February. Washington, DC: International Monetary Fund.
Strassheim, H., & Kettunen, P. 2014. When does evidence-based policy turn into policy-based evidence? Configurations, contexts and mechanisms. *Evidence and Policy*, 10(2): 259–277.
Strehlenert, H., Richter-Sundberg, L., Nystrom, M. E., & Hasson, H. 2015. Evidence-informed policy formulation and implementation: A comparative case study of two national policies for improving health and social care in Sweden. *Implementation Science*. https://doi.org/10.1186/s13012-015-0359-1. Accessed 2 July 2020.
Sutcliffe, S., & Court, J. 2005. *Evidence-Based Policymaking: What is it? How does it work? What relevance for developing countries*? London: Overseas Development Institute.
Sutcliffe, S., & Court, J. 2006. *A toolkit for progressive policymakers in developing countries.* Research and policy in development programme. London: Overseas Development Institute (ODI).
Szöcs, I., & Schlegelmilch, B. 2020. Embedding CSR in corporate strategies. In B. Schlegelmilch & I. Szőcs (Eds.) *Rethinking business responsibility in a global context: Challenges to corporate social responsibility, sustainability and ethics.* Cham: Springer.
Termeer, C. J. A. M., & Dewulf, A. 2019. A small wins framework to overcome the evaluation paradox of governing wicked problems. *Policy and Society*, 38(2): 298–314.
Termeer, C. J. A. M., Dewulf, A., & Biesbroek, R. 2019. A critical assessment of the wicked problem concept: Relevance and usefulness for policy science and practice. *Policy and Society*, 38(2): 167–179.
Terpstra-Tong, J. L. Y. 2017. MNE subsidiaries' adoption of gender equality and women empowerment goal: A conceptual framework. *Transnational Corporations*, 24(3): 89–102.
Thomas, M., Cordova Novion, C., de Haan, A., de León, G., Forest, M., & Iyer, S. S. 2018. Gender Mainstreaming: A strategic approach. *Global Solutions Journal* (T20 Policy Brief), 1(2): 155–173.
UN Economic Commission for Europe. 2019. Towards a conceptual framework for measuring gender-in-trade in official statistics. In *Note by the United Nations Conference on Trade and Development. Conference of European Statisticians, Work Session on Gender Statistics*. Working Paper No. 25. April 8.
UN Global Compact, UNGC. 2020. *SDG ambition: Scaling business impact for the decade of action*, January 23. https://www.unglobalcompact.org/take-action/sdg-ambition. Accessed 2 July 2020.
UN Statistics Division, UNSD. 2020. *E-handbook on the sustainable development goals*. New York: United Nations. https://unstats.un.org/wiki/display/SDGeHandbook/Home. Accessed 2 July 2020.
UN Women. 1995. *Beijing Declaration and Platform for Action/ Beijing +5 Political Declaration and Outcome.* New York: United Nations.
UN Women. 2018. *Turning promises into action: Gender equality in the 2030 Agenda for Sustainable Development.* New York: United Nations.
UN Women. 2019a. *Progress on the Sustainable Development Goals 2019.* New York: United Nations.
UN Women. 2019b. *WE EMPOWER: Responsible business conduct in G7 countries.* New York: United Nations.
UN Women and PARIS21. 2019. *Assessing data and capacity gaps in gender statistics.* https://paris21.org/node/3286. Accessed 2 July 2020.
UN Women. 2020. *Gender equality: Women's rights in review 25 years after Beijing*. New York: United Nations.
UNCTAD. 2014. *Investment by TNCs and gender: Preliminary assessment and way forward. Investment for Development Policy Research Series.* Geneva: UNCTAD.
UNCTAD. 2016. *Development and globalization: Facts and figures*. Geneva: UNCTAD.
UNCTAD. 2018. *Better data and statistics for gender-responsive trade policy.* UNCTAD Policy Brief, No. 70, October. Geneva: UNCTAD.
UNCTAD. 2020a. COVID-19 requires gender-equal responses to save economies, April 1. https://unctad.org/en/pages/newsdetails.aspx?OriginalVersionID=2319. Accessed 2 July 2020.
UNCTAD. 2020b. UN calls for $2.5 trillion coronavirus crisis package for developing countries, March 30. https://unctad.org/en/pages/newsdetails.aspx?OriginalVersionID=2315. Accessed 2 July 2020.
UNDP. 2018. *Human Development Indices and Indicators: 2018 Statistical update, technical note.* http://hdr.undp.org/sites/default/files/hdr2018_technical_notes.pdf. Accessed 2 July 2020.

UNDP. 2019. *Human Development Report 2019: Beyond income, beyond averages, beyond today: Inequalities in human development in the 21st century.* Geneva: United Nations.
United Nations. 2005. *Enhanced cooperation between the United Nations and all relevant partners, in particular the private sector. Report of the Secretary-General*. New York: United Nations.
United Nations. 2013. *Economic and Social Council. Substantive session of 2013.* New York: United Nations.
United Nations. 2015a. *The Millennium Development Goals Report 2015*. New York: United Nations. https://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf. Accessed 2 July 2020.
United Nations. 2015b. *Transforming our world: The 2030 Agenda for Sustainable Development.* Resolution adopted by the General Assembly on 25 September 2015. A/RES/70/1. UN General Assembly, Seventieth Session. Agenda items 15 and 116. New York: United Nations.
United Nations. 2016. *Integrating a gender perspective into statistics.* Studies in Methods, Series F No. 111. New York: Department of Economic and Social Affairs, Statistics Division, United Nations.
United Nations. 2019 *Progress on the sustainable development goals: The gender snapshot 2019.* New York: Department of Economic and Social Affairs, United Nations.
United Nations. 2020. *World economic situation and prospects as of mid-2020*. New York: United Nations.
UN-SQAF. 2018. *UN Statistics quality assurance framework: Including a generic statistical quality assurance framework for a UN Agency.* New York: United Nations.
Van Tulder, R. 2018. *Business and the sustainable development goals: A framework for effective corporate involvement*. Rotterdam: Rotterdam School of Management, Erasmus University.
Van Zanten, J. A., & van Tulder, R. 2018. Multinational enterprises and the sustainable development goals: An institutional approach to corporate engagement. *Journal of International Business Policy*, 1: 208–233.
WBCSD. 2018. *Reporting matters. Six years on: The state of play. WBCSD 2019 Report.* Geneva: WBCSD.
Whyte, K. P., & Thompson, P. B. 2012. Ideas for how to take wicked problems seriously. *Journal of Agricultural and Environmental Ethics*, 25(August): 441–445.
Witte, C., & Dilyard, J. 2017. Guest editors' introduction to the special issue: The contribution of multinational enterprises to the Sustainable Development Goals. *Transnational Corporations*, 24(3): 1–8.
World Business Council for Sustainable Development, WBCSD. 2017. *CEO guide to the SDGs*, March 27. https://www.wbcsd.org/Overview/Resources/General/CEO-Guide-to-the-SDGs. Accessed 2 July 2020.
World Business Council for Sustainable Development, WBCSD. 2018. *Business and the SDGs*, July 9. https://www.wbcsd.org/qvl3. Accessed 2 July 2020.
World Economic Forum, WEF. 2018. *The Global Gender Gap Report, 2018. Insight Report*. Geneva: World Economic Forum.
World Economic Forum, WEF. 2020. *The Global Gender Gap Report, 2020. Insight Report*. Geneva: World Economic Forum.

ABOUT THE AUTHORS

Lorraine Eden is Professor Emerita of Management and Research Professor of Law (joint appointment) at Texas A&M University. Her current research interests include transfer pricing and MNE strategies in the digital economy. She was 2008–2010 Editor-in-Chief of the *Journal of International Business Studies* and 2017–2018 President of the Academy of International Business. She is incoming Dean of the AIB Fellows for 2020–2023.

M. Fernanda Wagstaff is an Associate Professor in the Department of Marketing and Management at the University of Texas at El Paso where she holds the Robert E. and Jacqeline Skov Professorship in Business Ethics. She received her PhD from Texas A&M University. Her current research interests include gender and diversity issues at work and international human resource management. She is the Associate Editor of *Management Research*: *The Journal of the Iberoamerican Academy of Management*.

Accepted by Rob van Tulder, Area Editor, 21 May 2020. This article has been with the authors for three revisions.