The Agenda 2030 Compass: SDG expert assessment process

Work Package 2.1 Report for the Agenda 2030 Compass project

Authors: Timothy Suljada, Eric Kemp-Benedict, Emily Ghosh, Stefan Bößner

Acknowledgements: the authors would like to acknowledge the contributions of Anisha Nazareth and Susie Bresney

November 2021

This report presents the findings from the Agenda 2030 Compass project's Work Package 2.1 focusing on the development of the SDG expert assessment process. The project's overall findings are presented in the Agenda 2030 Compass synthesis report, which can be found along with all available project reports at: www.sei.org/agenda2030compass.

The research and development work has been carried out by a consortium consisting of the Stockholm Environment Institute (SEI), Jernkontoret (the Swedish Iron and Steel Producers' Association), the MIT Center for Collective Intelligence and Swedish software developer Swedwise.

Introduction

Initially designed to provide an expert assessment element to the Agenda 2030 Compass tool, Work Package 2.1 evolved to design the interactions assessment methodology for the project. The 2030 Agenda for Sustainable Development and its 17 Sustainable Development Goals can serve as a widely accepted and comprehensive set of criteria for assessing what is of value to a sustainable society or put another way, what represents societal value.

Expert assessments about how SDGs are expected to interact are at the centre at the centre of this methodology. These experts judgements were then supplemented by a wide array of historical data sets. Taken together, these two inputs (expert assessment and data sets) delivered the foundations for the Compass tool's analysis and strategies.

This report presents the methodology in the context of the literature on SDG interaction assessments and identifies its advantages and limitations as well as ideas for further work.

Methodology

Our methodology to analyse interactions between SDGs as a basis to inform decision-making incorporates both qualitative and quantitative methods., The goal of analysis is to construct a "map" of pairwise interaction scores, denoting whether an interaction between two SDGs is supportive or detrimental (positive or negative) and the strength of the interaction (on a scale of 1 to 3). The methods we adopted include expert judgment to provide assessments and scores of interactions, assignment of indicators from global data sets to quantify assessments, literature reviews and historical analysis of data sets as they relate to goals or targets. In drawing together these qualitative and quantitative methods, we adopt an approach from social science literature referred to as "triangulation". Following Denzin (1978), triangulation is "the combination of methodologies in the study of the same phenomenon" (Jick 1979). For our purposes, the triangulation approach allows us to cross-validate the results of our analysis from distinct methods. Beyond validation, the method can, according to Olsen (2004), also support interdisciplinary research like ours and serve to deepen and widen one's understanding of a research question.

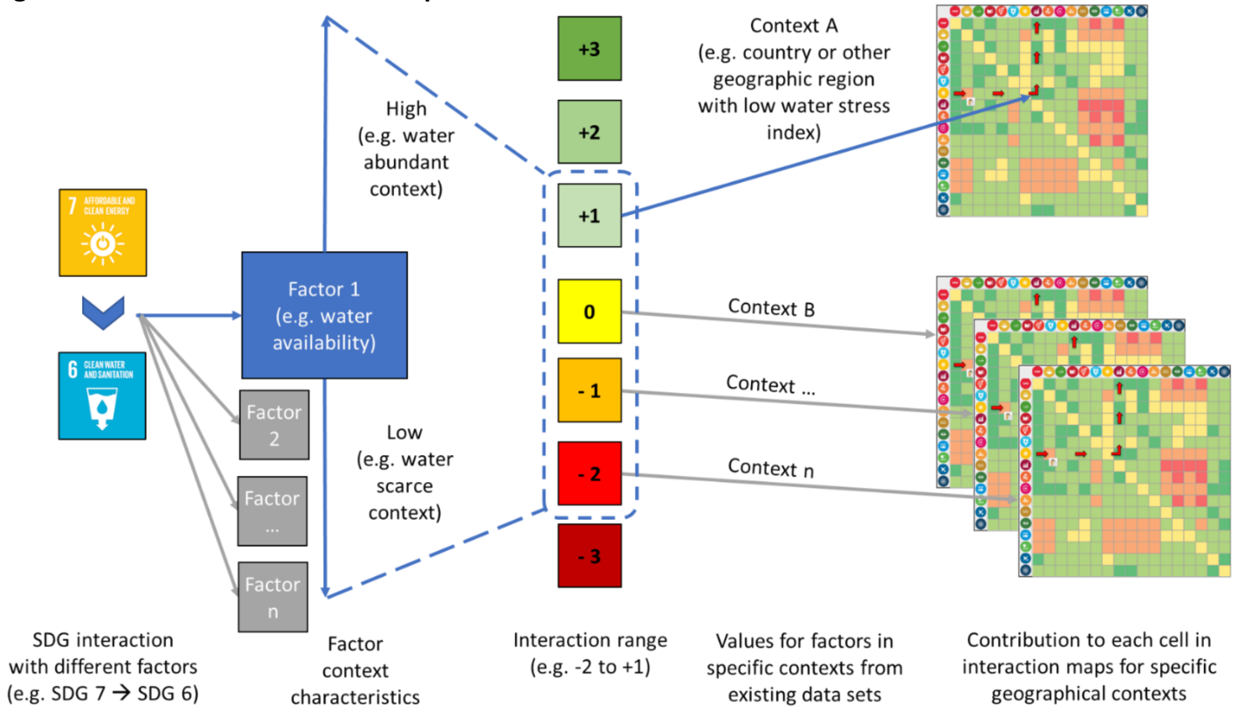
Context and factor maps

SDG interactions depend strongly on the context in which they take place (Nilsson et al. 2018; Weitz et al. 2018). For instance, analysing a SDG interaction pairing in Sweden might yield different assessments than an assessment of the same pairing in Kenya due to geographical, economic and societal differences amongst many other factors (see Annex 2 for a brief review of the SDG interactions literature). This suggests an expert elicitation or historical analysis targeted at a national context. However, drawing on country expertise for each of the roughly 200 countries in the world is challenging. We therefore built a methodology that uses expert input to identify explanatory factors upon which an SDG interaction would depend in different contexts (see Annex 5 for an explanation of the interpretation of factor and context scores). For example, one explanatory factor might be income level, another annual rainfall, yet another public expenditure on education, and so on. The goal of analysis is to construct a "factor map" showing SDG interactions corresponding to a particular value of the factor. Any given country, at any given time, will take on one value or another for each of the factors, so national context maps can be constructed by combining the factor maps corresponding to each value of the factor.

Box 1: Schematic overview of the assessment

The schematic in Figure 1 describes an example of how the method uses factors to develop an understanding of SDG interactions across contexts.

Figure 1 – The detailed assessment process



The example in Figure 1 identifies "water availability" as one of several possible factors in how making progress on SDG 7 affects making progress on SDG 6 with an interaction score ranging from +1 (enabling") in water abundant contexts to -2 ("counteracting") in water scarce contexts. The interaction range used is based on the work of (Nilsson, Griggs, and Visbeck 2016; Weitz et al. 2018) who suggest a 7 point scale when assessing SDG interactions ranging from +3 to -3. It is likely that several different factors can play a role in characterizing a specific interaction as indicated in the example above. Factors are then quantified for specific contexts using available global data sets. When all the interaction cells have been filled with information about factors and context specific scores (based on existing global data sets), a specific SDG interaction matrix could be generated for any given context.

SDG Expert Panel

The SDG Expert Panel solicited the input of 17 experts on the Sustainable Development Agenda, each of them being assigned one specific SDG. The experts were sourced from the Stockholm Environment Institute and from external partners where appropriate.[1] Using an online Expert Assessment Interface (EAI), specially developed for the process (see Annex 2), experts were invited to discuss SDG interactions, basing their judgement on the -3/+3 scale proposed by (Nilsson, Griggs, and Visbeck 2016; Nilsson et al. 2018). The guiding research question was *"If one makes progress on SDG X, how does this impact progress on SDG Y".*

The role of an SDG Expert Panellist was to take part in the development of a robust methodology for SDG interaction analysis and conduct assessments based on expert judgment. The Panel was an integrated part of the project's "collective intelligence" work with an important quality assurance role at the centre of the triangulation process including identification of factor and scores, feedback on data mining, review and reassessment based on feedback from testing of assessments through case studies and crowd sourcing exercises. This involved conducting assessments of the interaction between one SDG and each of the other 16 SDGs in the Agenda 2030 framework. The exercise was conducted in pairs with each panellist representing an SDG. The process primarily took place over the period from November 2019 until February 2020. The expert panel made the primary identification of "factors", which were subsequently summarized in a set of "factor themes" and a smaller set of "meta-factors". The factor themes aimed to capture as much of the initial identification of factors as possible, while the meta-factors aimed to provide a compact set of factors with high explanatory power when using the larger set is impracticable.

The discussion between experts representing SDGs went beyond a general assessment of SDG interaction. Experts were instructed to identify and discuss the factors on which a positive or negative assessment would depend. For example, when discussing interactions between SDG 7 (modern energy for all) and SDG 6 (water and sanitation for all), experts identified water availability as an important factor. If solar PV was used to increase access to energy, water use might go up in order to clean solar panels. Therefore, SDG 7 could have a negative impact on SDG 6 if water availability is low, for instance in arid climates (where, in turn, solar irradiation is often abundant). Following each expert's 16 rounds of discussions, more than 700 factors were compiled, often containing detailed descriptions of how the experts saw the factor influencing SDG interactions.

Context mapper

The context mapper is a suite of software tools, written in R, for generating the factor and context maps.[2] The process involves several steps. First, it is necessary to identify historical cases of "progress" vs. "no progress" on the SDGs. This was done by first carrying out a principal component analysis (PCA) on the indicators in the UN SDGs database. Those principal components were subsequently combined

[1] See a anonymised list of all experts in the Annex

[2] The code, which is released under an open-source license, is available at https://github.com/sei-international/Compass-Context-Mapper.

into a single indicator for each goal that can be unambiguously interpreted as reflecting progress if it increases. The details are provided in Annex 6.

First, a context-free "global" map was constructed, in which the "context" is the whole world, using the procedure described in Annex 5. The results are shown in Figure 1. Note that data for SDG 16 (peace, justice, and strong institutions) was insufficient, which is why it was blank. Also note that the methodology allows for asymmetric matrices. Most of the values are low, which is anticipated with the global context, but some SDGs show tendencies. In particular: several SDGs reinforce SDG 11 (sustainable cities); SDG 14 (life below water) reinforces SDG 15 (life on land).

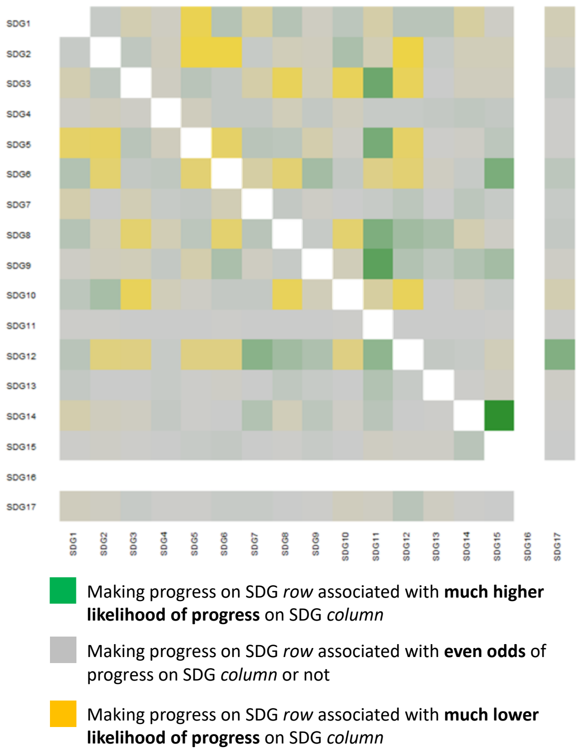


Figure 1: The global context map

Context maps are constructed from factor maps, which could be produced based on indicators that represent factors (i.e. factor indicators). Two different methods were used to determine the indicators that factor maps would be based upon: "factor themes" and "meta-factors", as described below.

Factor themes are a cluster of factors within a particular category such as agriculture, economic conditions land ownership/use, etc (the outputs of the categorisation are presented in the Results section). A Natural Language Processing (NLP) tool developed and administered by Oliver Miendl from the University of Augsburg assisted with the factor theme categorization process. To use the NLP-tool, the factors were first pre-processed into a form that could be understood by the tool's algorithm. The pre-processing involved removing numbers, punctuation, and words that do not provide value to the semantical meaning of a factor (like the words "example", "the", "could", etc.). The factors were then put through the NLP-tool which categorized the factors into 59 clusters. While the NLP-tool was useful for taking a first cut at categorizing the factors, many of the factors were incorrectly clustered together due to the lack of detail in the factor description. The project team then manually reviewed the clusters and assigned factor themes.

Multiple indicators were often available to represent factor themes. When that was the case, a PCA was carried out to identify linear combinations of the indicators that tended to be correlated. The PCA analysis showed which indicators were most representative of the factor theme and were retained as factor indicators for the theme.

Meta-factors were developed by Rob Laubacher as part of the analysis in Work Package 2.2 to identify a shorter list of indicators whose data sets showed a significant degree of orthogonality, that is independence from one another (see Work Package 2.2 report). This degree of independence was measured using p-values, which provides a regression co-efficient that ranged from X to Y with 0 meaning correlated and 1 meaning orthogonal. The fact that some degree of correlation remains in the data sets for these indicators means that combining meta-factors to describe interactions introduces some error to the results.

For both factor themes and meta factors, the representative indicators are used to construct a factor map with the following steps:

1. Categorise countries by quintiles of the factor indicator (e.g., the countries with the lowest 20% of per capita income, the next-lowest 20%, etc. – these make up the factor *levels*);
2. Calculate empirical SDG interaction scores across the countries within each factor level as explained in Annex 5;
3. Calculate a rough measure of the statistical significance of the interaction score (see Annex 5).

Once the factor maps were created, national context maps were constructed using the following steps:

1. Find the levels of all factors for the specified country;
2. Retrieve the factor map for each level;
3. Calculate a weighted sum of the factor maps (see Annex 5);
4. Subtract the context-free “global” map from the context map.

Results

SDG Expert Panel process

Following the assessment, the 17 SDG expert panellists were requested to take part in a survey to assess the method and identify lessons for the project’s future work. A total of 9 responses were received, focused on the clarity of guidance, the method and the workload involved in SDG interaction assessments using the SDG Expert Panel method. The full results of the survey are presented in Annex 3.

A short publication on the SDG Expert Panel process was published on the SEI website on 16 December 2019. The SEI Perspective, [Calibrating the Compass: How to factor in the context of SDG interaction](), outlined the process and its contribution to the project.

Outputs from the SDG Expert Panel

Of the 751 factors identified by the SDG Expert Panel, a total of 454 factors were deemed valid according to the level of detail provided for each factor, as well as the availability of relevant indicators to assess the status of the factors for the context maps. The factors were also grouped into 31 factor themes, reflecting the category of the factor identified (see Annex 4 for details).

Among the valid factors, the top five themes included water resources (49 factors), energy/natural resources (35 factors), government (30 factors), gender (30 factors) and agriculture (23 factors). A full list of the factor themes is provided in Table 1.

Table 1: Factor themes and the number of factors categorised within each theme

Factor Theme	Number of interactions with this factor theme
Agriculture	23
Climate/Disasters	13
Conservation and Ecosystems	16
Crime/Conflict/Violence	17
Economic conditions	7
Education	21
Employment	13
Energy/Natural Resources	35
Financial Resources	17
Fishing	13
Food	13
Gender Parity	30
Government	30
Health	9
Industry	4
Inequality	15
International Relations	8
Land Ownership/Use	12
Marine	9
Miscellaneous	1
Pollution	7
Social Policy/Context	22
Sustainability	10
Technology	18
Tourism	8
Trade	1
Transportation/Infrastructure	11
Urbanization	2
WASH	15
Waste management	5
Water resources	49
Grand Total	**454**

A total of 271 indicators were identified as proxies for the factors. The indicators were determined by first establishing an interpretation of each factor to be able to consistently evaluate each factor, and then finding a relevant, measurable indicator from a global data set that was easily accessible and had global coverage. Not all indicators were uniquely identified for factors; many indicators were used to describe more than one factor and some factors were described by more than one indicator.

Table 2 presents the most frequently identified indicators for the factors.

Table 2 Indicators most frequently identified proxies for factors (these 17 indicators were used to describe 10 factors or more)

Indicator	Dataset	Factor theme(s)	No. factors using indicator
Regulatory Quality: Percentile Rank	WB Governance Indicators	Government Water resources Financial Resources Climate/Disasters Fishing	21
Government Effectiveness: Percentile Rank	WB Governance Indicators	Government Financial Resources International Relations Gender Parity	19
Level of water stress: freshwater withdrawal as a proportion of available freshwater resources (%)	UN SDG DATA	Water resources Energy/Natural Resources	15
CPIA transparency, accountability, and corruption in the public sector rating (1=low to 6=high)	WDI	Government Financial Resources International Relations	14
Proportion of bodies of water with good ambient water quality (%)	UN SDG DATA	Water resources Waste management Conservation and Ecosystems Land Ownership/Use	13
Completion rate, by sex, location, wealth quintile and education level (%)	UN SDG DATA	Education Technology	12
Voice and Accountability: Percentile Rank	WB Governance Indicators	Government Water resources Gender Parity	12
GINI index	WDI	Inequality Economic conditions Climate/Disasters	12
Proportion of population using basic drinking water services, by location (%)	UN SDG DATA	WASH Water resources	12
Conflict Recurrence (# of Conflicts)	PRIO	Crime/Conflict/Violence Energy/Natural Resources	11
Average proportion of Marine Key Biodiversity Areas (KBAs) covered by protected areas (%)	UN SDG DATA	Conservation and Ecosystems Marine	11
Proportion of safely treated domestic wastewater flows (%)	UN SDG DATA	Water resources WASH Waste management	11
Land Use - Cropland Area Under Conservation Tillage	FAOSTAT	Agriculture Land Ownership/Use	10
Land Use Indicators - Agriculture area under organic agric. (Shares)	FAOSTAT	Agriculture Land Ownership/Use	10
Proportion of teachers who have received at least the minimum organized teacher training (e.g. pedagogical training) pre-service or in-service required for teaching at the relevant level in a given country, by sex and education level (%)	UN SDG DATA	Education Technology	10

Indicator	Dataset	Factor theme(s)	No. factors using indicator
Proportion of elected seats held by women in deliberative bodies of local government (%)	UN SDG DATA	Gender Parity Water resources	10
Rule Of Law, Value	WB GOV360	Government	10

The meta-factors are shown in Table 3.

Table 3: The ten meta-factors

Category	#	Meta-factor	Key indicator
Socio-economic	1	Per capita GNI (PPP)	Gross national income per capita PPP
Governance	2	Control of corruption	World Governance Indicators—Control of Corruption
	3	Political stability/ Absence of violence	World Governance Indicators—Stability
	4	Voice in governance	World Governance Indicators—Voice
	5	Social policy and services	Social protection expenditure as % of GDP
Natural resources	6	Agricultural land	Agricultural land as share of land area
	7	Forest resources	Forests as a % of land area
	8	Fresh water resources	Renewable fresh water resources per capita
	9	Marine and coastal resources	Fisheries production
	10	Fossil fuel and mineral resources	Natural resource rents as % of GDP

Context mapper

A pair of sample maps are shown in Figure 2 for Sweden and Tanzania. The maps are superficially similar, but by focusing on specific SDGs, the differences become clearer. In the figure, SDG 12 is highlighted: Sustainable Production and Consumption. The driver is on the row, whereas the response is on the column. So, for example, if progress on SDG 12 were enhanced through an industrial strategy, then historical data suggest that one consequence for Sweden is a significant likelihood of not making progress on lowering inequalities (the box in the figure). This is not necessarily a problem, as inequality is already low in Sweden. However, examining the scores in the column for SDG 12 shows that making progress on other SDGs has a dampening or neutral impact. This means that negative second-round feedbacks may be in effect, offsetting the initial positive impact of the strategy.

In Tanzania, progress on SDG 12 is, as in Sweden, associated with no progress on SDG 10. In this case, that might well be seen as a problem. However, based on historical data, the SDG 12-enhancing strategy increases the likelihood of making progress on SDG 7 (affordable and clean energy) and SDG 13 (climate action). What is more, the second-round impacts are mainly neutral rather than dampening. That is true in particular for SDGs 7 and 13, suggesting that action on SDG 12 could have beneficial effects that are not subsequently weakened through second-round interactions.

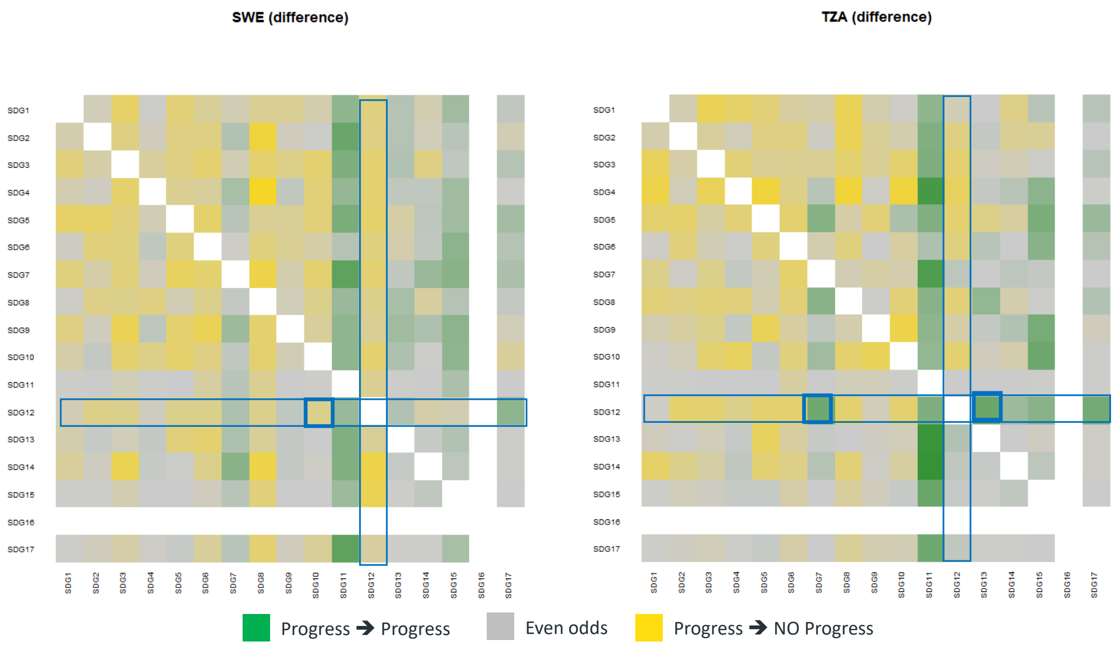


Figure 2: Context maps for Sweden (SWE) and Tanzania (TZA) built using the meta-factors, with annotations for SDG 12: Sustainable production and consumption

A rough test was carried out on the consistency between the historical analysis and the expert panel assessment. For this activity, the larger "factor themes" set of factor indicators was used, as this is closest to the set of factors suggested by the panel. The panel also provided estimates of ranges of scores for a given interaction. From the historical analysis, a range of scores was also calculated, taking the values for all levels of the factor indicator. Agreement was said to be "good" if the average absolute difference between the endpoints was less than one; it was "medium" if the average absolute difference was less than two; it was "poor" otherwise. The distribution was 26% good, 34% medium, and 40% poor. While not outstanding results, the distribution is significantly different than what would be expected if the agreement was purely random, namely 18% good, 40% medium, and 42% poor. More substantively, some disagreement is expected, because the reasoning expressed by experts did not necessarily match the implied interpretation of the SDGs as reflected in the official indicators.[3]

Discussion

Limitations of the methodology

- **Interaction assessments conducted by multiple experts across 272 different discussions introduces inevitable inconsistencies.** While each expert panellist was briefed individually during a full-day workshop by the project team, the analysis of results revealed subtly different understandings of the method and varying levels of detail and thoroughness in assessment and documentation. The inconsistencies made analysis difficult and reduced the volume of useable data. An interpretation of the factors needed to be made to ensure consistency for the indicator selection process. The interpretations were made by one person to ensure consistency, with a second person doing a secondary check. There are limitations of this interpretation process as it may not fully capture the intention behind a given factor by the expert panellist. Refer to examples set of factors and their interpretations in Annex 4.
- **The focus question for the SDG Expert Panel captures only part of the interaction assessment.** The focus question "How does making progress on SDG A affect making progress

[3] For example, education indicators tend to focus on years of schooling, education budgets, and similar high-level data. In contrast, the expert panel focused on less readily-measured factors.

on SDG B?” omits the converse question of “How does *not making progress* on SDG A affect making progress on SDG B?”.

- **The context mapper does not show regression, rather high odds of not making progress.** The Beysian method used in the context mapper measures the odds of making progress on SDG B if there is progress on SDG A. However, low odds of making progress can only be interpreted as very likely to be no progress rather than regression.

Reflections and insights

- **The SDG Expert Panel was a novel but cumbersome SDG interactions method.** It was challenging to recruit and maintain subject matter experts to represent each SDG throughout the interactions assessment process. Over the period from expert panel initiation at the October 2019 workshop to the conclusion of assessments in February 2020, 14 of 17 SDGs were represented by one expert panellist and the remaining three SDGs switched experts mid-way through the assessments. This is a good and consistent outcome, however, it required management oversight and support throughout, including providing an extension of the initial deadline for assessments by two months for all panellists and three months for three of the panellists (a total of 5 months instead of the anticipated 2 months for assessments).
- **SDG expert panellists assessed the method positively overall but had differing views on the burden of work involved.** The survey received nine responses (of a potential 17). These responses indicated that factor method had potential to bring benefits to SDG interactions, provided it was understood and used consistently. Respondents indicated that approaches and methods were explained well (77% of respondents gave a score of 4 or 5 out of 5) and supporting documents and guidance were rated average to above average (3 or 4 out of 5 with 66% scoring materials 4 out of 5). Despite these strong responses, there are areas for improvement in follow up and training to ensure consistency as we discovered when processing inputs. Seventy-seven per cent of respondents deemed the workload of an SDG Expert Panelist “manageable” and respondents indicated that each interaction discussion took under 90 minutes (66% of respondents took less than 60 minutes).
- **Assessments by SDG expert panellists varied in quality.** Of the 751 factors identified across all SDG interactions, 454 were assessed by the project team to be distinct and clear enough to be able to assign an indicator. There were differing levels of motivation amongst experts, some were more reactive and quicker to complete their interactions than others. This could be explained by existing workload from other projects and the relatively small allocation of time that the SDG Expert Panel represented (up to 80 hours input over the assessment period).
- **The SDGs, at the level of specific targets, are not internally consistent.** For example, SDG 13 on climate goal includes two distinct aspects, mitigation and adaptation, that makes assessments at the goal level difficult. Assessments at the target level rather than goal level may be a solution to the sometimes diverse coverage of goals. While this changes the interactions matrix from 17 x 17 to 169 x 169, feedback from SDG expert panellists in the survey highlights that many interaction discussions were in fact at target level. Also that many target-target interactions could be ignored for specific SDG pairings.
- **Indicators and data sets limit the reliability of the factor approach.** It is a useful method to assess interactions across contexts, but indicators and global data sets were limited.
- **SDGs emerged from a process with divergent views and the SDG indicators reflect this.** Process for coming up with the indicators was more political than the process for the goals and targets (Fukuda-Parr 2019).
- **A strength of the methodology is that it meaningfully combines expert inputs with historical data.** This is relevant in the context of the SDGs, where the process of creating the goals and targets was separate from that of generating the indicators, and the results diverged at some

points (Fukuda-Parr and McNeill 2019; Fukuda-Parr 2019). In a preliminary comparison of the results from each method, we found that some of discrepancies might be explained by substantive differences between the language in the goals and targets and the implicit meaning behind the indicators.

Conclusions

The context mapper allows a rapid assessment for further analysis. The context maps focus the attention of the user on the strength of direct and to some extent indirect interactions. It is not a definitive assessment of interactions, but one that is based on a balance between an expert driven approach (with its inherent biases) and a data-driven approach (along with its limitations due to indicator selection and data reliability).

The context mapper developed in this work package has been used in the Agenda 2030 Compass project's Strategy Analyser, which is a workshop-based process and toolbox to analyse the sustainability implications of a planned intervention within that context. The context mapper is part of the toolbox to generate an SDG interaction matrix for societal contexts, usually at country level. as a basis for workshop case study discussions.

Annex 1: Brief review of the literature on SDG interactions

Analysing and understanding the interactions between the SDGs and their different targets is a quite new scholarly field. Nevertheless, there is already some significant activity and several approaches and methods have been applied by researchers around the world, some rather qualitative, some more quantitative.

In the more qualitative camp, some scholars suggest using a nexus approach – inspired by the water-energy-food nexus in development literature (Allan, Keulertz, and Woertz 2015) - to investigate the interconnections between different SDGs (Boas, Biermann, and Kanie 2016). More concretely, (Nilsson, Griggs, and Visbeck 2016) suggest to describe SDG interactions using a scale from -3 (cancelling) to +3 (indivisible). Researchers Weitz et al. (2018) use this methodology when discussing SDG interactions with experts in their specific fields. By assigning aforementioned values to each SDG interaction, the researchers together with the experts were able to construct a matrix of all the 272 (17x16) interactions, using Sweden as an example (Weitz et al. 2018). The International Council for Science (ICS) used this approach by conducting an in depth analysis of interactions between SDGs 2, 3, 7 and 14 (International Council for Science 2017). A similar approach can be found in (Coopman et al. 2016) who explore SDG interconnection by using a classification and scoring system to assess linkages between SDGs in the European Union. Another of those rather qualitative approaches can be found in (Nerini et al. 2019) who investigate how progress on SDG 13 (Climate Action) interacts with the other 16 goals using academic publications to justify their analysis. One lesson learnt from those investigations is that SDG interactions are highly context specific and depend on which geographical space or temporal horizon is chosen for the analysis amongst other factors. (Nilsson et al. 2018).

Another line of inquiry has been the application of network analysis to SDG interactions. For instance, (Le Blanc 2015) uses qualitative data such as the wording used in SDG targets and goals descriptions to inform his networks and show how SDGs are interconnected but doesn't assess these interconnections. (Dawes 2020) also used qualitative information found in the International Council of Science report to construct networks but assess those networks quantitatively using dynamic models. Others such as (Lusseau and Mancini 2019) use a more data-centric approach to construct their networks. Feeding world bank data about SDG relevant indicators (such as access to electricity etc.) into a linear mixed effect model (MEM), they derive two set of networks which they then control for level of income per country. They conclude that the importance of an SDG (its centrality in the network) depends on the income level of a country and that more conflicts between SDGs might arise in high-level income countries while progress on one SDG usually would have a positive impact on other SDGs in lower income countries. Another data intensive approach is used by (Pradhan et al. 2017) who use statistical correlation analysis, based on available indicators data sets on SDG progress by the UN statistical division, looking for synergies and trade-offs between data pairs and, consequently, between SDGs. They find that positive correlations among SDGS would outweigh the negative trade-offs but also that the relationship is varies across countries. Nevertheless, some SDGs like SDG 3 (good health and well-being) would be connected to others in a more synergetic way, while SDG 12, responsible consumption and production would have the largest potential of trade-offs (negative impacts) on other SDGs. Another example of using network analysis to analyse SDG interactions can be found in a report published by Institute for Global Environmental Strategies (IGES) (Zhou and Moinuddin 2017).

Other scholars combine two methodologies such as cross-impact analysis and network analysis (Zelinka and Amadei 2019b) or a logistic technology innovation model with a system dynamics model (Zelinka and Amadei 2019a) to investigate SDG interactions, showing that SDGs vary in influence and impact on the network and that context (for instance, which countries are investigated) matters.

The approaches and methods described above, have also been applied to research, investigating specific cases such as countries (or groups of, singular or few SDGs or regions. Allen, Metternich and Wiedmann (2019) apply a multi-criteria analysis with a network analysis to 22 countries of the Middle Eastern region, finding that delivering on SDGs 2, 6, 7 and 12 would be of particular importance for the region (Allen, Metternicht, and Wiedmann 2019). Other investigations of SDG interactions based on case studies include (Collste, Pedercini, and Cornell 2017) assessing interactions between SDGs 3, 7 and 4 using an integrated assessment model (iSDG), Bastos-Lima et al. who describe qualitatively interactions of the SDG systems with the REDD+ framework (Bastos Lima et al. 2017) or Requejo-Castro et al. (2016) who pursue Bayesian network approach to investigate interlinkages between SDG 6 and the other SDGs (Requejo-Castro, Giné-Garriga, and Pérez-Foguet 2020). SDG 6 is also at the heart of a case study by the UN Economic and Social Commission for Asia and the Pacific (ESCAP) investigating SDG interaction from and to SDG 6 using mapping tools and a qualitative driven “systems approach” (ESCAP 2017).

References

Allan, Tony, Martin Keulertz, and Eckart Woertz. 2015. “The Water–Food–Energy Nexus: An Introduction to Nexus Concepts and Some Conceptual and Operational Problems.” *International Journal of Water Resources Development* 31 (3): 301–11. https://doi.org/10.1080/07900627.2015.1029118.

Allen, Cameron, Graciela Metternicht, and Thomas Wiedmann. 2019. “Prioritising SDG Targets: Assessing Baselines, Gaps and Interlinkages.” *Sustainability Science* 14 (2): 421–38. https://doi.org/10.1007/s11625-018-0596-8.

Bastos Lima, Mairon G., Gabrielle Kissinger, Ingrid J. Visseren-Hamakers, Josefina Braña-Varela, and Aarti Gupta. 2017. “The Sustainable Development Goals and REDD+: Assessing Institutional Interactions and the Pursuit of Synergies.” *International Environmental Agreements: Politics, Law and Economics* 17 (4): 589–606. https://doi.org/10.1007/s10784-017-9366-9.

Boas, Ingrid, Frank Biermann, and Norichika Kanie. 2016. “Cross-Sectoral Strategies in Global Sustainability Governance: Towards a Nexus Approach.” *International Environmental Agreements: Politics, Law and Economics* 16 (3): 449–64. https://doi.org/10.1007/s10784-016-9321-1.

Collste, David, Matteo Pedercini, and Sarah E. Cornell. 2017. “Policy Coherence to Achieve the SDGs: Using Integrated Simulation Models to Assess Effective Policies.” *Sustainability Science* 12 (6): 921–31. https://doi.org/10.1007/s11625-017-0457-x.

Coopman, Anna, Derek Osborn, Farooq Ullah, Emily Auckland, and Graham Long. 2016. “Seeing the Whole. Implementing the SDGs in an Integrated and Coherent Way.” Stakeholder Forum.

Dawes, Jonathan H.P. 2020. “Are the Sustainable Development Goals Self-Consistent and Mutually Achievable?” *Sustainable Development* 28 (1): 101–17. https://doi.org/10.1002/sd.1975.

Denzin, Norman K. 1978. *The Research Act: A Theoretical Introduction to Sociological Methods*. 2nd Edition. McGraw-Hill.

ESCAP. 2017. “Integrated Approaches for Sustainable Development Goals Planning: The Case of Goal 6 on Water and Sanitation.” Economic and Social Commission for Asia and the Pacific (ESCAP).

Fukuda-Parr, Sakiko. 2019. “Keeping out Extreme Inequality from the SDG Agenda – the Politics of Indicators.” *Global Policy* 10 (S1): 61–69. https://doi.org/10.1111/1758-5899.12602.

Fukuda-Parr, Sakiko, and Desmond McNeill. 2019. “Knowledge and Politics in Setting and Measuring the SDGs: Introduction to Special Issue.” *Global Policy* 10 (S1): 5–15. https://doi.org/10.1111/1758-5899.12604.

International Council for Science. 2017. “A Guide to SDG Interactions: From Science to Implementation.” *A Guide to SDG Interactions: From Science to Implementation*. International Council for Science, Paris. https://doi.org/International Council for Science (2017). A Guide to SDG Interactions: from Science to Implementation. International Council for Science, Paris. DOI:10.24948/2017.01 <https://doi.org/10.24948/2017.01>.

Jick, Todd D. 1979. “Mixing Qualitative and Quantitative Methods: Triangulation in Action.” *Administrative Science Quarterly* 24 (4): 602–11. https://doi.org/10.2307/2392366.

Le Blanc, David. 2015. “Towards Integration at Last? The Sustainable Development Goals as a Network of Targets.” *Sustainable Development* 23 (3): 176–87. https://doi.org/10.1002/sd.1582.

Lusseau, David, and Francesca Mancini. 2019. “Income-Based Variation in Sustainable Development Goal Interaction Networks.” *Nature Sustainability* 2 (3): 242–47. https://doi.org/10.1038/s41893-019-0231-4.

Nerini, Francesco Fuso, Benjamin Sovacool, Nick Hughes, Laura Cozzi, Ellie Cosgrave, Mark Howells, Massimo Tavoni, Julia Tomei, Hisham Zerriffi, and Ben Milligan. 2019. “Connecting Climate Action with Other Sustainable Development Goals.” *Nature Sustainability* 2 (8): 674–80. https://doi.org/10.1038/s41893-019-0334-y.

Nilsson, Måns, Elinor Chisholm, David Griggs, Philippa Howden-Chapman, David McCollum, Peter Messerli, Barbara Neumann, Anne-Sophie Stevance, Martin Visbeck, and Mark Stafford-Smith. 2018. “Mapping Interactions between the Sustainable Development Goals: Lessons Learned and Ways Forward.” *Sustainability Science* 13 (6): 1489–1503. https://doi.org/10.1007/s11625-018-0604-z.

Nilsson, Måns, Dave Griggs, and Martin Visbeck. 2016. “Policy: Map the Interactions between Sustainable Development Goals.” *Nature News* 534 (7607): 320. https://doi.org/10.1038/534320a.

Olsen, Wendy. 2004. “Triangulation in Social Research: Qualitative and Quantitative Methods Can Really Be Mixed,” 30.

Pradhan, Prajal, Luís Costa, Diego Rybski, Wolfgang Lucht, and Jürgen P. Kropp. 2017. “A Systematic Study of Sustainable Development Goal (SDG) Interactions.” *Earth’s Future* 5 (11): 1169–79. https://doi.org/10.1002/2017EF000632.

Requejo-Castro, David, Ricard Giné-Garriga, and Agustí Pérez-Foguet. 2020. “Data-Driven Bayesian Network Modelling to Explore the Relationships between SDG 6 and the 2030 Agenda.” *Science of The Total Environment* 710 (March): 136014. https://doi.org/10.1016/j.scitotenv.2019.136014.

Weitz, Nina, Henrik Carlsen, Måns Nilsson, and Kristian Skånberg. 2018. “Towards Systemic and Contextual Priority Setting for Implementing the 2030 Agenda.” *Sustainability Science* 13 (2): 531–48. https://doi.org/10.1007/s11625-017-0470-0.

Zelinka, David, and Bernard Amadei. 2019a. “A Systems Approach for Modeling Interactions Among the Sustainable Development Goals Part 1: Cross-Impact Network Analysis.” *International Journal of System Dynamics Applications (IJSDA)* 8 (1): 23–40.

———. 2019b. “Systems Approach for Modeling Interactions Among the Sustainable Development Goals Part 1: Cross-Impact Network Analysis.” *International Journal of System Dynamics Applications (IJSDA)* 8 (1): 23–40. https://doi.org/10.4018/IJSDA.2019010102.

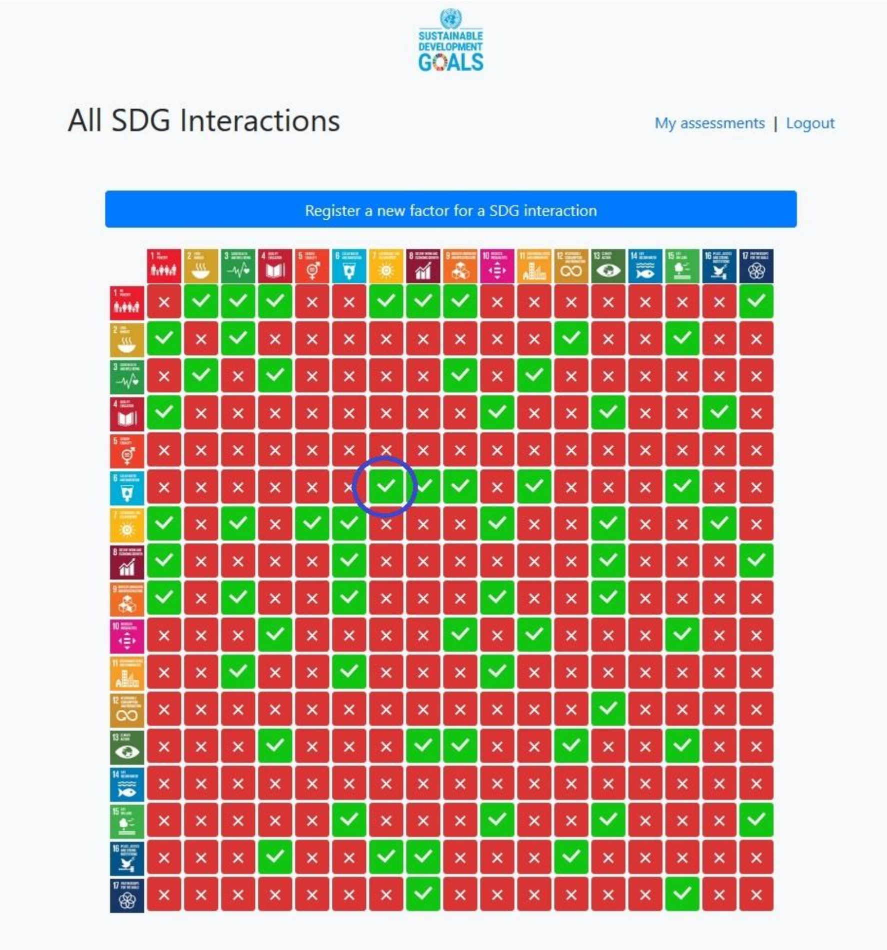
Zhou, Xin, and Mustafa Moinuddin. 2017. “Sustainable Development Goals Interlinkages and Network Analysis: A Practical Tool for SDG Integration and Policy Coherence.” Institute for Global Environmental Strategies. https://www.iges.or.jp/en/pub/sustainable-development-goals-interlinkages/en.

Annex 2: the SDG Expert Assessment Interface (SDG-EAI):

The following guidance was provided to SDG Expert Panellists to support the SDG interactions assessments.

Log into the [interface, accessible following this link.] In order to guide you when using the interface, we provided a [short instructional video, available on the TEAMS site here].[4] Once you have logged onto the SDG-EAI, you will see this landing page.

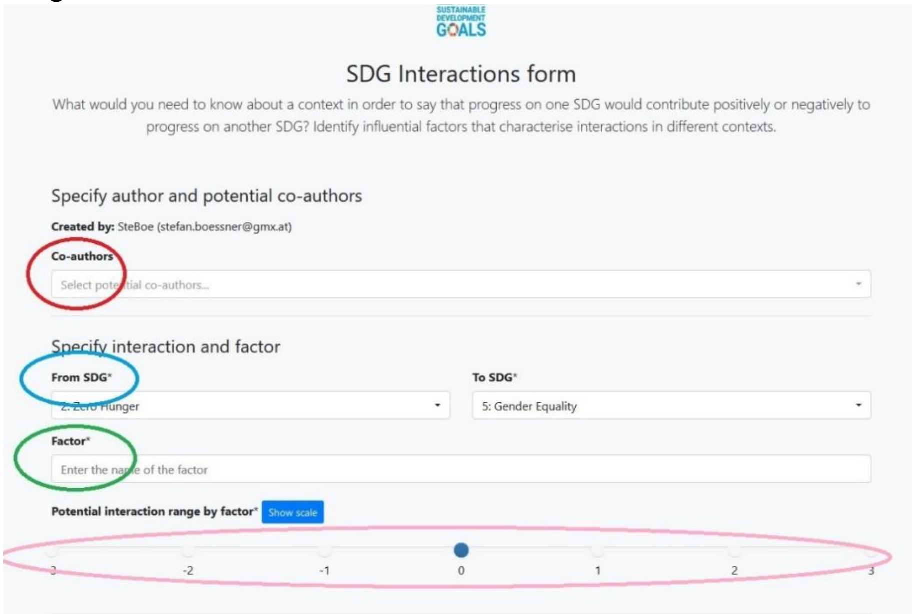
Image 1 – The landing page of the SDG-EAI



Once you are on the landing page, click on the intersection of the x-axis and y-axis of your interaction. In our example case of SDG6 and SDG7, the blue circle marks the interaction. Alternatively, you can also click on the blue header "*register a new factor for a SDG interaction*". This should take you to the next, sub-layer of the interface.

[4] If you have any issues accessing Microsoft Stream where the video is hosted, an alternative link to download the video is available here.

Image 2 – The SDG-EAI[5]



You can add your expert pairing partner as *co-author,* indicated by the circle in **red**. The **blue** circle allows you to choose the *direction* of the interaction while the **green** circle highlights the *factor* each interaction assessment is based on.

The Direction

SDG interaction can be done in two different directions. It can either be discussed from SDG A to SDG B or vice versa. Depending on the direction you choose, your discussions will be different. For instance, discussing the impact of progress made on *SDG 7 – Affordable and Clean Energy* on *SDG 6 – Clean Water and Sanitation* is different to discussing the impact of progress made on SDG 6 on SDG 7. Please note that you can discuss both the interaction concerning the overall goals or the interactions of individual targets. A good overview of the different SDGs as well as their targets can be found here and further reading is provided in the FAQ.
However, it is the **green** circle, the "factor" which is one of the most important aspects to discuss.

The Factors

The concept of factors is one of the unique elements of the SDG interaction assessment method used in this project. When considering the main question for the SDG interaction assessment, "How does making progress on SDG A affect making progress on SDG B?" The intuitive answer is: "it depends on the context". Of the range of methods used to assess SDG interactions, most either prescribe the context in which an interaction takes place (e.g. in Sweden, how does making progress on SDG A on SDG B) or make a general assessment of interactions independent of specific contexts. The method used in the Agenda 2030 Compass project is different because we aim to take into account the range of possible contexts in which an interaction takes place between two SDGs and consider what distinguishes that interaction in different contexts.

In order to contextualise SDG interactions, we ask our experts to identify "factors" which might influence the interaction between the SDGs. For instance, the increase of renewable energies such as

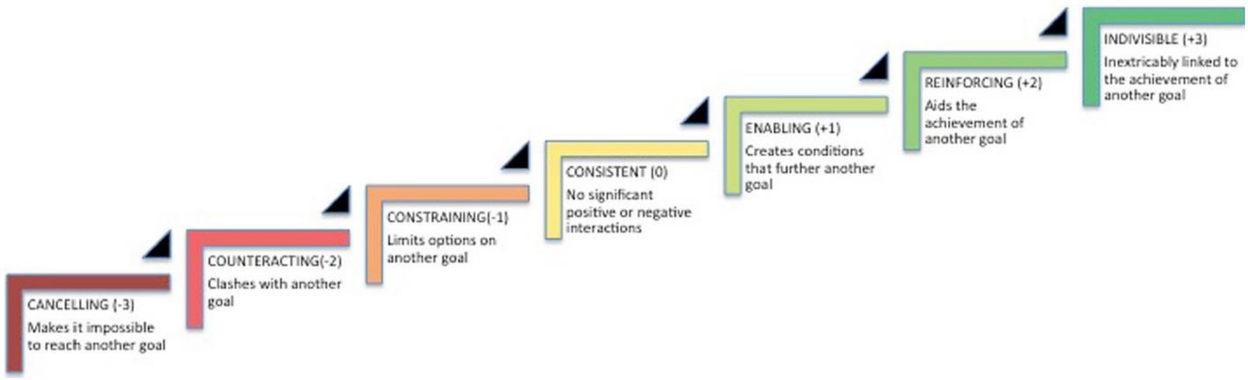
[5] This image is for a newly created interaction in the interface; the design will differ slightly from what you will see when you click on an already existing interaction

solar PV (a potential strategy to deliver on SDG7, target 7.2.) might increase the water usage (for cleaning), particularly in dry, hot desert climates. Whether this increased water usage brought about potentially by progress on SDG7 has a negative impact on SDG6 (Clean Water and Sanitation) depends on many factors. One factor identified during our first workshop was "*water availability*". If water is used efficiently and waste-water treatment facilities support the potentially increasing need for water to clean solar PV installations, then the interaction from SDG7 and SDG6 might be assessed as slightly positive. It is important to note, that there are no right or wrong answers when it comes to factors. Each factor you deem important should be noted down and its inclusion supported by a few lines explaining why you chose this factor and how it might impact SDG interaction (see next section on where to note that down). Another thing to keep in mind is that we don't want you to assess the factors themselves, but the interactions based on these factors. In our example, you don't assess the factor *"water availability"* but the interaction SDG7 to SDG6 based on the factor *"water availability"*. Nevertheless, it is highly likely that the factors themselves will have to be attributed a certain value. For instance, is *water use efficiency* high or low in any specific context? If you can, please note down any quantification of a given factor, using for example scales such as "high, moderate, or low".

Should you know of any relevant data sources available to assign a value to a certain factor, please feel free to note this data source as well. Any information which makes the interaction more concrete and illustrative is highly appreciated. However, we suggest you focus on having fruitful discussions and not waste your time too much on finding the appropriate data sources. Indeed, one of the tasks for the analytical team will be to find data sets of the factors you suggest.

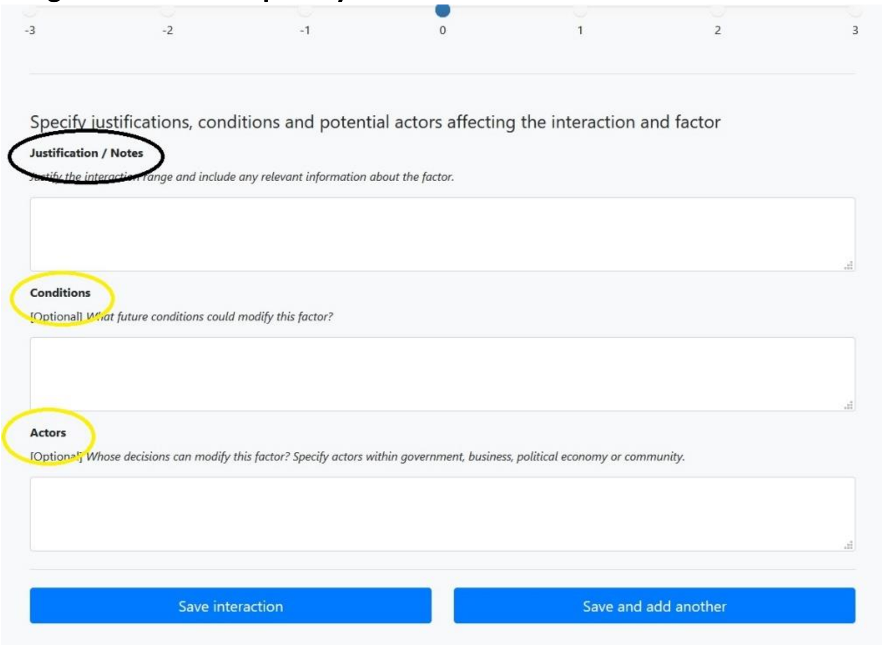
The assessment based on the factors is at the heart of this exercise. We are using the assessment framework developed by Nilsson et al.[6] and Weitz et al.[7] to quantify the strength of each of the SDG interactions on a scale from minus three (-3) to plus three (+3). The relevant function in the interface is circled in pink in image No.2 (above). While the lowest value denotes a negative interaction (cancellation), meaning that progress on a specific goal/target makes it impossible to reach another goal/target, the highest value denotes an indivisible relationship, meaning that progress on one goal/target is inextricably linked to the progress of another goal/target. The graph below describes the scale more in detail.

Image 3 – The SDG interaction scale



[6] (Nilsson, Griggs, and Visbeck 2016)
[7] (Weitz et al. 2018)

Image 4 – where do I put my notes?



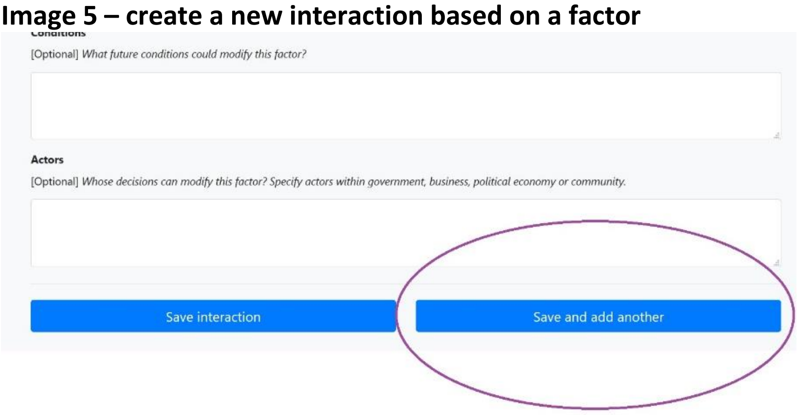
Once the discussion is finished, roles are swapped and an interaction is discussed the other way round. If you discussed SDG X to SDG Y, now you discuss SDG Y to SDG X. We encourage you to have a quite free and semi-structured discussion; But please remember to agree on a factor and to discuss the interaction based on the factors using the -3 to +3 scale mentioned above.

Besides assessing the scale of interactions, the interface also offers two other fields where you could leave notes (circled in yellow). If you come across certain conditions which might impact the interactions in the future or important actors which themselves might influence SDG interactions, please note them down in the appropriate fields.

Once the discussions successful terminated and the main points of the discussion entered in the appropriate boxes, please make sure you save the interaction.

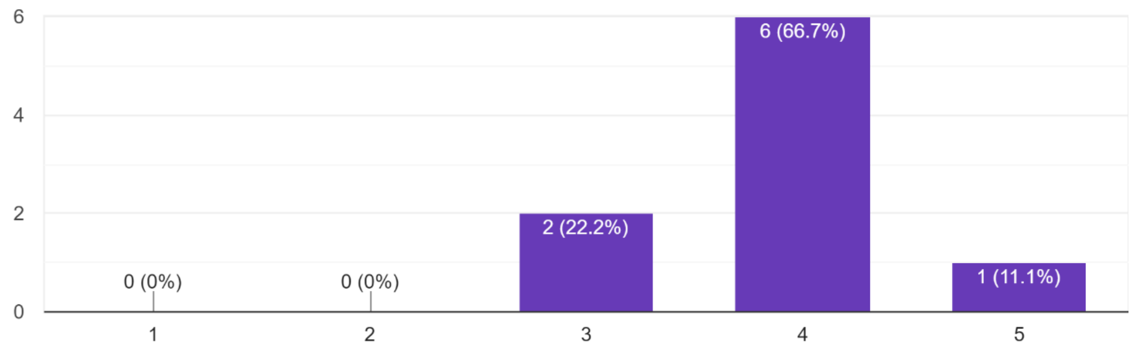
IMPORTANT: Please remember to save additional factors for a specific SDG interaction as new entry by clicking "save and add another" as shown in the following image. Repeat the previous steps.

Image 5 – create a new interaction based on a factor

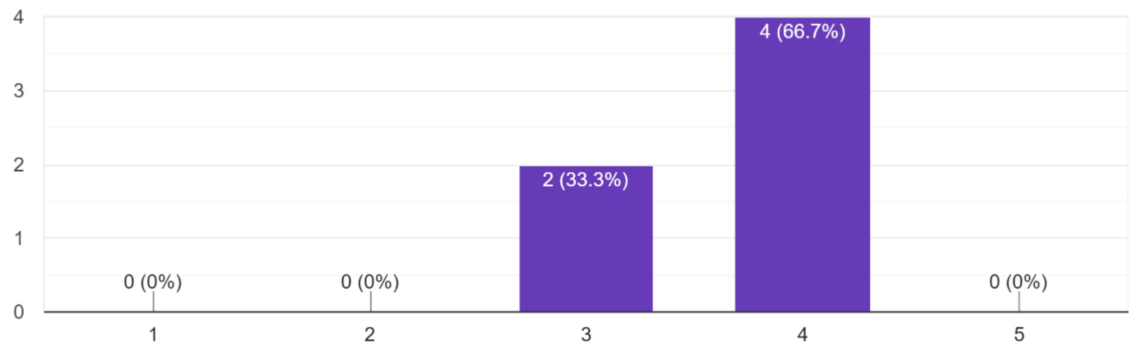


Annex 3: Survey of participants in the SDG Expert Panel

1. On a scale from 1 to 5, how well were the approaches and methods explained to you prior of pursuing the pairing exercise? The value 1 is badly ...ned, and the value 5 indicates perfect explanation.
9 responses



2. Did you use the supporting documents (guidance document, instruction video)? If so, how would you rate them? The value 1 stands for 'not u...while 5 indicates very useful guidance documents.
6 responses



3. Using the Expert Assessment Interface (EAI, https://agenda2030-expertassessment.herokuapp.com/login), is there anything you would improve from a usability perspective?

8 responses

Once an entry is "green" I would prefer to be able to click it to add a new interaction rather than to see the list of interactions.

No, it is easy to use.

There were some assessments that repeated factors from other assessments. It would have helped if these could have been easier to distinguish so that more factors could have been looked for.

It was pretty good in general - well designed, easy to follow.

I had to go back and forth to google the targets to have them in mind to know what to write (maybe there was a function to look at them in the interface - but I never found that - but it would have been good to have access to)

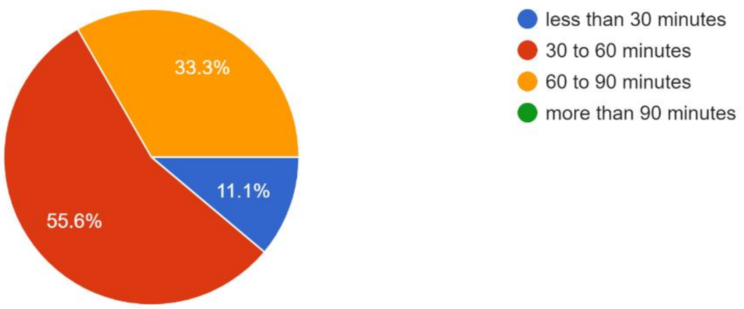
The way to go back to an interaction already assessed and change it is not very practical.

Rather unclear how the last part on actors I supposed to be used.

I feel the explanation of what we were required to do should have followed the template for interactions more thoroughly to ensure that we share a common interpretation of the tasks from the beginning

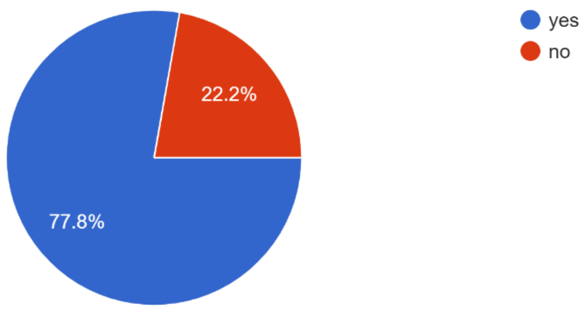
4. On average, how much time did each pairing discussion on interactions take?

9 responses



5. Was the workload in the panellist role manageable?

9 responses



6. If you found the workload unmanageable, could you please elaborate?

4 responses

n/a

It took longer than 30 minutes to start with, and if one got stuck, or if one did not thik that one factor explained it well enough, but otherwise one got faster and faster - 20 minutes in the end for the not so complex ones perhaps - (I tried to mark both -30 and 30-60, because the average was Close to 30 ... 30+/-15 then one got the hang of it - but I did a lot of scorings

While at the end we had more time to finalize the pairings, finding time to do them all before Christmas (as initially communicated), which is already a very busy period, was tough.

Meeting heavy work, very disruptive for work schedule.

6. What was most difficult part of the role of the panellist?

8 responses

Putting things in terms of "factors". We had interesting discussions but it was sometimes hard to reduce what we were discussing down to a "factor".

Challenging to not go into too much detail especially because it is unclear how all this work is going to be used.

Making appointments with some of the other panelists. Some were very busy and some did not respond to emails or suggestions

Sometimes it felt like I didn't have all the knowledge needed to be confident of a certain interaction. And it was hard to find time to meet everyone.

Factor-finding-thinking - and especially thinking about the "curvature" of the factor to the effect .. if the factor is high the score should be-arguing back and forth

To deal with the complexity of the SDGs and their targets

Finding time to prepare for the discussions and the discussions themselves.

The uncertainty of the method. What we were supposed to contribute to in terms of output/project and the difference in understanding of the concepts and the tasks within the interactions

7. How would you compare the pairing discussions you had with each panellist with (if you attended) those during the workshop in October 2019?

8 responses

It was pretty similar. It was nice to be in the room together, though.

n/a

Much better when we had more time as we did online. However, the exercise at the workshop was useful to get us into the swing of things.

Most worked well - in some there were more connections that others. But generally they were all quite similar in the way we worked, regardless of whether face-to-face or virtual.

It was easier to be two (I did most of mine by myself - there goes my anonymity)

We got more efficient as progressing. Also we could see certain types of factors coming back with clusters of SDGs.

Similar

it became easier with time (i also started with the most difficult interactions with SDG 16 and 17 at the workshop which was very hard to grasp). The WS interactions were, however, more connected to the presentation and the teams introduction to the concepts

8. Which part of the whole assessment process do you think could be improved? And how?

8 responses

In many pairings, we ended up going through all of the targets. It seems to me that it might have been possible to do a target-by-target mapping. I realize that would be a lot of interactions, but unlike the goals, many of them would be "no significant interaction".

Would be useful to have been provided an example of what a good and complete factor looks like. Also think panelists should have been asked to provide example indicators.

I would have appreciated a workshop or face-to face meeting or webinar with all the panelists midway in the process. That way we could have compared our processes and shared our concerns. The project leaders did a good job of getting back to us with practical advice but a midway content discussion might have helped.

Perhaps it could have helped to recruit two people for each SDG to make the process faster. If one had enough time the workload wasn't a problem, but it did get a bit tiring to go through all pairings, so maybe a second person could have ensured momentum by taking some of the load. Just a thought.

Nexus-seminars (like we did in the October-session) there people discussed with other having "similar" SDGs how they argued etc - could have been useful a quarter intp the process and halfway - and changing the nexus-Groups to avoid sub-groups thinking differenty

Going forward, having a better understanding of what we will be required from us and when and how it will be used would help, although I understand this is the first time it is being done and it is not necessarily possible.

Overall understanding of how this material will be used

More clarity of method, what the output will be used for and exact interpretation of the tasks when going into online interactions

9. Apart from workshops and the EIA, do you have a suggestion on what other tools and methods could be used to discuss SDG interactions?

6 responses

You might try some "holistic" exercises where the same people fill in all interactions.

No.

Since the SDG's varied as to which emphasized technical targets and indicators and which focused on broader more systemic change, it might have helped to have some way to deal with the differences. Perhaps an interface option that was framed around these two perspectives would have been useful.

I think webinar discussions - hard to facilitate, perhaps - could be useful, particularly to get people from around the world. Could do some sort of workshop where you get people to come into a room and work together virtually but with people in same room...sounds weird, but it could be like gaming or something.

Some five different good examples with different types of "curvature"

no

10. Was the factor approach a useful way of discussing SDG interactions?

9 responses

I think so. I understand that it was a way to get results more general than for a specific country. It was still sometimes a struggle to identify them.

Yes.

Factors were interpreted differently by the different panel participants. Some factors were abstract others more concrete. That made some of the assessments difficult and perhaps not as easily comparable.

I liked the idea of factors. Sometimes they were easier to pin down than others. I tended to use the phrase "It depends on..." as my way of understanding what we were looking for when looking for factors.

It is tricky - and we might run into traps - but if it works out well it is a brilliant way to get a context-"machine" for any context

I think it depends on how it is/will be used considering how subjective it is

It was tricky, especially with indirect relations.

Is this the right question? The factor approach is the method perse, right? This is unavoidable given the setup, no?

Partly useful but the "factor" was differntly interpreted in different interactions

11. Are there any reflections about the SDGs, interactions or other things that came up during this project that you wish to share?

9 responses

The SDGs, at the level of specific targets, are not internally consistent. It is also clear that the climate goal really has to be split into mitigation and adaptation. Either that or we should work at target level. I realize that could potentially be a lot more work, but in fact many of my discussions were at target level, and many target-target interactions could probably be ignored. It might help to work at target level when assessing strategies as well. But I would not recommend revising the approach at this stage.

No.

It would be interesting to have a new round of assessments in a year or two. The interruptions already caused by the coronavirus outbreak and their influence on world economy, personal consumption habits, etc. may significantly alter the factors previously described.

Nope.

There are two main traps - the past-future-dilemma: it is easy to get explain how we got here, but we want factors and scores helping us out of here - so the process should start here but be forwardlooking, but in a "realistic way", and history then gives guidance. The second trap/solution is that in many cases the starting Point (level of development) of both the SDGs involved might be what gives most guidance, but not Always - and that can give an assymtry in that a large part of the map is using that "simple" method, and the part something else (other kinds of factors). It might be OK, it might even be good, but it might give different Dynamics over time: As/if hings (SDG-levels and/or factor-values) improve, there might be a need to look at them again and scrap them to find better ones - and it might be that the two different types have different over-time-quality (and there might be other problems too). But at the moment that is PhD-stuff - lets do some maps now.

It hasn't been clear throughout the process how the results will be synthesized and presented. It seems almost like an impossible task to pull all of the different results together, especially considering the many disparate ways of working with the interactions depending on the people involved

I very much liked learning more about other topics and the work colleagues do at SEI.

No

When contributing to a project as an "expert" it would be nice to what the product i am contributing to is. This was not clear and I get nervous that what i provided will be used in the wrong way.

Annex 4 – SDG Expert Panel – summary of factors, factor themes and indicators

Figure 3 Factors identified by SDG Expert Panel across all SDG interactions (total=750, T indicates row and column totals)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	T
1	0	4	2	2	2	5	1	3	1	8	1	4	3	2	3	2	1	44
2	3	1	4	1	3	3	2	2	2	2	1	1	3	3	3	1	1	36
3	2	5	0	4	3	2	0	1	1	1	1	2	1	0	1	1	1	26
4	3	3	3	0	4	3	3	3	1	5	1	5	3	1	2	3	3	46
5	4	3	6	3	0	4	2	3	3	4	1	2	2	1	2	7	2	49
6	5	6	3	3	4	1	5	2	5	7	4	3	2	3	4	3	2	62
7	4	3	4	2	2	2	0	2	2	3	1	6	2	2	2	1	3	41
8	4	3	1	2	4	5	2	0	3	5	1	2	2	2	4	6	3	49
9	3	4	5	2	2	4	3	1	0	2	1	3	5	3	2	2	3	45
10	5	3	3	3	2	5	1	5	2	0	8	4	5	3	1	3	2	55
11	1	2	5	1	1	6	1	1	1	5	0	3	1	1	1	1	1	32
12	4	4	3	2	3	5	3	3	3	1	1	0	3	4	4	1	1	45
13	3	2	2	3	1	1	4	3	3	3	1	1	0	4	2	3	1	37
14	5	2	0	1	1	2	1	3	1	4	1	4	2	0	2	2	2	33
15	3	3	2	1	1	5	2	2	1	4	1	4	2	3	0	2	3	39
16	7	3	2	4	6	3	4	4	3	2	1	4	6	3	3	0	7	62
17	3	4	1	3	4	4	3	1	6	2	1	1	1	7	3	5	0	49
T	59	55	46	37	43	60	37	39	38	58	26	49	43	42	39	43	36	

Table 4 Global datasets used in assigning proxy indicators to factors

Dataset	URL reference
FAOSTAT	http://www.fao.org/faostat/en/
AQUASTAT	http://www.fao.org/nr/water/aquastat/data/query/index.html?lang=en
WDI (World Bank Development Indicators)	https://data.worldbank.org/indicator/
WDI (World Bank Development Indicators)	https://databank.worldbank.org/
WB GOV360	https://govdata360.worldbank.org/
Findex (World Bank Global Financial Index)	https://globalfindex.worldbank.org/
SDG Tracker	https://sdg-tracker.org/
Natural Resource Management Index (NRMI)	https://sedac.ciesin.columbia.edu/data/collection/nrmi
WRI Water Risk Atlas	https://wri.org/applications/aqueduct/water-risk-atlas
WRI Aqueduct Maps	https://www.wri.org/aqueduct/data
FAO databases	http://www.fao.org/land-water/databases-and-software/aquastat/en/

UN SDG Indicators	https://unstats.un.org/sdgs/indicators/database/
Global Freshwater Database	https://gemstat.org/
Global Ecosystems Data	https://www.usgs.gov/centers/gecsc/science/global-ecosystems-data?qt-science_center_objects=4#qt-science_center_objects
PREDICTS: Global database of terrestrial species' responses to human pressures	https://www.predicts.org.uk/
UN Data	http://data.un.org/Explorer.aspx
Harvard Power Sector Reform Tracker (PSRT)	https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/M7SY6X
IEA Policies Database	https://www.iea.org/policies
World Social Protection Report Data	https://www.social-protection.org/gimi/Wspr.action
World Social Protection Data	https://www.social-protection.org/gimi/WSPDB.action?id=32
ILO Social Security Database	https://www.ilo.org/sesame/IFPSES.SocialDatabase
Migration Data Portal	https://migrationdataportal.org/data
Migration Data Portal (Stata data)	http://www.impic-project.eu/data/
UN ESA Population Policies	https://esa.un.org/PopPolicy/dataquery.aspx?MainMenu
Integration of Immigrants	https://www1.compareyourcountry.org/indicators-of-immigrant-integration
DEMIG	https://www.migrationinstitute.org/data/demig-data/demig-policy-1/download-the-data/demig-policy-data-downloads
STATUS OF RATIFICATION INTERACTIVE DASHBOARD	https://indicators.ohchr.org/
World Justice Project Rule of Law Index	https://worldjusticeproject.org/our-work/research-and-data/global-insights-access-justice-2019
WHO Global Health Observatory data repository	https://apps.who.int/gho/data/node.home
UN Drug and Crimes Database	https://dataunodc.un.org/
ILO Employment Protection Database	https://www.ilo.org/dyn/eplex/termmain.byCountry?p_lang=en
Protected areas	https://www.protectedplanet.net/
Transboundary FreshWater Dispute Database	https://tfddmgmt.github.io/tfdd/map.html
Ocean Health Index	http://www.oceanhealthindex.org/region-scores
Global Lakes and Wetlands Database	https://www.worldwildlife.org/pages/global-lakes-and-wetlands-database
FAO food loss and waste database	http://www.fao.org/food-loss-and-food-waste/flw-data/en/
BP Statistical Review of World Energy	https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html
IEA Renewables Information Statistics	https://www.oecd-ilibrary.org/energy/data/iea-renewables-information-statistics/world-renewable-and-waste-energy-statistics_data-00549-en?parentId=http%3A%2F%2Finstance.metastore.ingenta.com%2Fcontent%2Fcollection%2Frenewab-data-en

WB Rural Access Index	https://datacatalog.worldbank.org/dataset/rural-access-index-rai
Ecosystem services valuation database	https://www.es-partnership.org/services/data-knowledge-sharing/ecosystem-service-valuation-database/
Conflict, War and Terrorism: Datasets	https://guides.ucf.edu/war/wardata
Conflict Analysis and Resolution Datasets	https://infoguides.gmu.edu/conflict/data
FAO Gender and Land Rights Database	http://www.fao.org/gender-landrights-database/data-map/statistics/en/
Database for International Treaty on Plant Genetic Resources	https://ssl.fao.org/glis/
Database on Freshwater Fish Species	https://www.nature.com/articles/sdata2017141
Region-Based Marine Trophic Index	http://www.seaaroundus.org/regional-mti-tools/
Fishing Vessels Criminal Records	https://spyglass.fish/
IUU Fishing Index	http://iuufishingindex.net/
Global Fishing Watch	https://globalfishingwatch.org/map/
Global Reporting Initiative	https://database.globalreporting.org/SDG-12-6/Global-Tracker
Berkeley Global Temperature Data	http://berkeleyearth.org/data/
Climate Watch	https://www.climatewatchdata.org/
International Environmental Data Sources	https://libguides.princeton.edu/envirecon/InternationalData
IEA World Energy Database	https://www.oecd-ilibrary.org/energy/data/iea-world-energy-statistics-and-balances/world-energy-balances_data-00512-en
The Shift Data Portal	https://theshiftdataportal.org
Global Public Procurement Database	https://www.globalpublicprocurementdata.org/gppd/
WB What A Waste Global Database	https://datacatalog.worldbank.org/dataset/what-waste-global-database
UN Global Material Flows Database	https://www.resourcepanel.org/global-material-flows-database
WB Global Indicators of Regulatory Governance	https://rulemaking.worldbank.org/
WB Trade and Competitiveness Data	https://tcdata360.worldbank.org/topics
IRENA Data and Statistics	https://www.irena.org/Statistics
RE Explorer	https://www.re-explorer.org/
Climate Equity Calculator	https://calculator.climateequityreference.org/
Global Petrol Prices	https://www.globalpetrolprices.com/electricity_prices/
Public Transport	http://brtdata.org/indicators/systems/modal_split___public_transport
Corruption Perceptions Index	https://www.transparency.org/en/cpi
Global Health Data	http://ghdx.healthdata.org/gbd-results-tool
Global Dams Database	https://datacatalog.worldbank.org/dataset/global-dams-database
World Bank World Integrated Trade Solution (WITS)	https://wits.worldbank.org/country-indicator.aspx?lang=en

Climate change laws and policies	https://climate-laws.org/legislation_and_policies

Process of assigning factor themes and factor indicators

The following section provides examples of how factors were interpreted and assigned indicators by the project team.

In the first example shown in Table 5, the factor title and justification provided by the expert panel clearly stated the interaction between SDG 9 and SDG 2, that is, the need for more investment in agricultural research. In this case, and in the case of many other factors, the justification lays out the specific SDG targets that the interaction describes which provides further context behind the factor. The project team noted that this factor was high quality and detailed and then developed a plain language interpretation of the factor to provide clarity in the subsequent decisions that were made for the factor theme and indicator. In this case, since the interaction mostly focused on the agricultural sector, the factor theme "Agriculture" was assigned. The team then looked for relevant quantitative indicators from global datasets to represent the interaction that was taking place and found two indicators including the number of agricultural researchers and the amount of expenditures according to the Agricultural Science and Technology Indicators (ASTI).

Table 5: Example of a factor with sufficient detail

Source	Parameter	Detail
Expert Panel	**From SDG**	9
Expert Panel	**To SDG**	2
Expert Panel	**Factor Title**	Research budgets, e.g., specifically for agriculture, but also level of investments.
Expert Panel	**Factor Justification**	9.5 (research) but also basics in 9.1, 9.2, interacts with 2A. both contains increasing research and investments.
Expert Panel	**Condition**	-
Expert Panel	**Modifications**	Investors, research orgs.
Project Team	**Comments/Quality**	Detailed, SDG indicators provided.
Project Team	**Interpretation**	Improving scientific research, includes agricultural research
Project Team	**Factor Theme**	Agriculture
Project Team	**Factor Indicator(s)**	ASTI-Researchers; ASTI-Expenditures

As shown in the second example (Table 6), in some cases, the project team did not assign a factor theme or indicator because the factor (as denoted by the "factor title") describes the SDG itself, and not necessarily the interaction between a set of SDGs.

Table 6: Example of a factor with sufficient detail, but no theme or indicator assigned

Source	Parameter	Detail
Expert Panel	**From SDG**	14
Expert Panel	**To SDG**	7
Expert Panel	**Factor Title**	Expanded area of marine protected areas with strict protection
Expert Panel	**Factor Justification**	14.5 "By 2020, conserve at least 10 per cent of coastal and marine areas, consistent with national and international law and based on the best available scientific

Source	Parameter	Detail
		information" to 7.2 "By 2030, increase substantially the share of renewable energy in the global energy mix"
Expert Panel	**Condition**	More marine areas that are strictly protected will increase the difficulty of exploiting off-shore fossil fuels. | However, as long as the economic revenues of exploiting off-shore fossil resources are higher than the incentives to protected the sea, the likelihood of this happening is not very high.
Expert Panel	**Modifications**	-
Project Team	**Comments/Quality**	Detailed, SDG indicators provided.
Project Team	**Interpretation**	Conserving coastal and marine areas may contribute to increasing renewable energy in the global energy mix
Project Team	**Factor Theme**	No factor
Project Team	**Factor Indicator(s)**	-

As shown in the third example (Table 7), the project team did not assign a factor theme or indicator because the factor is too specific for there to be a relevant indicator. Furthermore, the interpretation was strongly influenced by the SDG targets noted in the factor's justification due to the vague nature of the factor itself.

Table 7: Example of a factor with some detail, and no theme or indicator assigned

Source	Parameter	Detail
Expert Panel	**From SDG**	6
Expert Panel	**To SDG**	17
Expert Panel	**Factor Title**	Depending on the enabling environment for technical innovations
Expert Panel	**Factor Justification**	6.3 - 6.4, 6A -> 17.6-17.8 Demand for innovative solutions can create new knowledge and capacity that in turn can support global capacity.
Expert Panel	**Condition**	-
Expert Panel	**Modifications**	-
Project Team	**Comments/Quality**	No detail, SDG indicators provided, but description is unclear (innovation in what? Assumed water innovation since it is regarding SDG 6)
Project Team	**Interpretation**	Improving water quality and water use efficiency through innovation facilitated by international cooperation and global partnerships
Project Team	**Factor Theme**	Factor too specific
Project Team	**Factor Indicator(s)**	-

As shown in the fourth example (Table 8), while some detail and SDG targets were provided, the project team was unable to interpret the factor in a way that clearly captured the intent behind the factor, and for this reason, a factor theme and indicator was not assigned.

Table 8: Example of a factor with some detail, and no theme or indicator assigned due to lack of clarity

Source	Parameter	Detail
Expert Panel	**From SDG**	16
Expert Panel	**To SDG**	8
Expert Panel	**Factor Title**	Fragility, vulnerability
Expert Panel	**Factor Justification**	E.g., post-disaster communities | Marginalisation/high vulnerability x with institutional strength/capacity ==> lower score | Less obvious/targetable ==> lower score | 16.5 -> 8.3, 8.4, 8.5, 8.7, 8.8, (8.9), 8.10 | "all their forms" in unclear whether implementable
Expert Panel	**Condition**	-
Expert Panel	**Modifications**	-
Project Team	**Comments/Quality**	Some detail, SDG indicators provided
Project Team	**Interpretation**	Corruption, especially the less obvious forms of corruption, can lead to increased fragility and vulnerabilities to disasters
Project Team	**Factor Theme**	Factor unclear
Project Team	**Factor Indicator(s)**	-

As shown in the fifth example (Table 9), while sufficient detail and SDG targets were provided, the project team did not assign a factor theme and indicator because the factor could not be applied to a specific country, which is necessary for the context mapping.

Table 9: Example of a factor with detail, and no theme or indicator assigned due to lack of application to a specific country

Source	Parameter	Detail
Expert Panel	**From SDG**	16
Expert Panel	**To SDG**	8
Expert Panel	**Factor Title**	Whether or not increased representation of developing countries puts pressure on developed countries to deliver on goal 17
Expert Panel	**Factor Justification**	Broaden and strengthen the participation of developing countries in the institutions of global governance (16.8) could positively affect e.g. targets 17.1, 17.2, 17.3, if increased representation and voice of developing countries do indeed increase the pressure they put on developed countries to deliver on goal 17.
Expert Panel	**Condition**	-
Expert Panel	**Modifications**	-
Project Team	**Comments/Quality**	Detailed, SDG indicators provided.
Project Team	**Interpretation**	Broadening and strengthening the participation of developing countries in the institutions of global governance could increase support for sustainable development from developed to developing countries
Project Team	**Factor Theme**	Factor not country specific
Project Team	**Factor Indicator(s)**	-

Annex 5: Bayesian interpretation of impact scores

The proposed approach to SDG interaction scores is to treat the interaction scores in terms of the (log of the) odds ratio of making progress on an SDG vs. not making progress. The interaction scores are then conditional odds, which suggests an approach based on Bayesian statistical reasoning.

The Bayesian interpretation leads to an interpretation of scores for strategies. The presence of a strategy is further evidence (in Bayesian terms), which changes the assessment of the likelihood of progress on the SDGs. Thus, the strategy does not, by itself, lead directly to fulfilling the SDGs within a given context. Rather, it increases – or, possibly, decreases – the likelihood of making progress when all SDG interactions are taken into account.

In the Bayesian approach, each SDG is assumed to be in one of two mutually exclusive states: "progress" or "no progress". (It is possible to include "regression" as well, but the binary progress/no progress was used for the analysis.) Generically, the state of SDG *j* is denoted by s_j. To streamline the notation, we use the symbol *j* when progress has been made and $\bar{j}$ (with an overbar) when progress has not been made.

The fundamental value of interest is the odds of making progress vs. not making progress, which can be expressed in terms of an odds ratio, R_j,

$$R_j = \frac{P(j)}{P(\bar{j})}. \tag{1}$$

If we are told that we have made progress on SDG *i*, then we should update our assessment of the likelihood of progress on SDG *j*. That can be derived using Bayes' rule, to give

$$\underbrace{\frac{P(j\,|\,i)}{P(\bar{j}\,|\,i)}}_{\text{posterior odds}} = \underbrace{\frac{P(i\,|\,j)}{P(i\,|\,\bar{j})}}_{\text{Bayes factor}}\underbrace{\frac{P(j)}{P(\bar{j})}}_{\text{prior odds}}. \tag{2}$$

For a full assessment, it is also necessary to evaluate how the odds change if we are certain that progress has *not* been made on SDG *i*. We put this in terms of an update of the odds ratio of *not* making progress on *j*,

$$\frac{P(\bar{j}\,|\,\bar{i})}{P(j\,|\,\bar{i})} = \frac{P(\bar{i}\,|\,\bar{j})}{P(\bar{i}\,|\,j)}\frac{P(\bar{j})}{P(j)}. \tag{3}$$

SDG interactions scores as log odds ratios

Because odds ratios can vary over a wide range and they are not symmetric, it is common to use the log (with some base *b*) of the ratio, so we are interested in the quantity

$$Z_j = \log_b \frac{P(j)}{P(\bar{j})} = \log_b P(j) - \log_b P(\bar{j}). \tag{4}$$

The log odds ratio is equal to zero when there are even odds. Unlike the odds ratio (without the log), it is symmetric, in that an increase in the probability of progress has the same effect as a decrease in the probability of no progress.

If we are told that we have made progress on SDG *i*, then we should update our assessment of the likelihood of progress on SDG *j*. That can be derived using Bayes' rule, to give

$$\log_b \frac{P(j\,|\,i)}{P(\bar{j}\,|\,i)} = \log_b \frac{P(i\,|\,j)}{P(i\,|\,\bar{j})}\frac{P(j)}{P(\bar{j})} = \log_b \frac{P(i\,|\,j)}{P(i\,|\,\bar{j})} + \log_b \frac{P(j)}{P(\bar{j})}. \tag{5}$$

In the Bayesian interpretation of the SDG interaction scores, the score is the update to the log odds ratio for progress on SDG *j* when given the information that progress has been made on SDG *i*. That is given by the term

$$\alpha_{ij} = \log_b \frac{P(i \mid j)}{P(i \mid \bar{j})}. \tag{6}$$

This is the log of the “Bayes factor”. Bayes factors are used to consistently update odds ratios given new information.
The odds ratio itself is found by taking the inverse of the logarithm. For a base b, that is given by

$$\frac{P(i \mid j)}{P(i \mid \bar{j})} = b^{\alpha_{ij}}. \tag{7}$$

Taken literally, a judgement that an interaction “cancelling” should mean $P(i \mid j) = -\infty$. However, in practice it only needs to be “large enough”. Table 10 shows the odds ratios for $b = 10$.

Table 10: Odds ratios for SDG scores when b = 10

Judgement	α_{ij}	Odds		
Cancelling	−3	1	:	1000
Counteracting	−2	1	:	100
Constraining	−1	1	:	10
Consistent	0	1	:	1
Enabling	+1	10	:	1
Reinforcing	+2	100	:	1
Indivisible	+3	1000	:	1

Using factors for a context

Contexts are defined in terms of “factors” that can explain variation in the strength of SDG interactions. The construction of factor maps requires that a set of factors be identified and assigned to distinct “levels”. The indicator itself my take on a very large number of possible values (for example, GDP per capita in USD to two decimal places), but for the analysis the values are binned into a small number of levels – perhaps up to five (for the analysis, quintiles were used).
The quantity of interest now is the odds ratio of making progress on SDG j or not when there is observed to be progress on SDG i, conditional on the collection of factors $\mathbf{F} = \{F_1, F_2, \ldots, F_N\}$ taking on a particular set of values. The factors are assumed to act independently in the sense that the joint distribution of $\mathbf{F}$, i, and j can be written as a product,

$$P(i, j, \mathbf{F}) = \prod_{n=1}^{N} P(i, j, F_n). \tag{8}$$

Each term on the right-hand side can be written in at least two ways,

$$P(i, j, F_n) = P(j \mid i, F_n) P(i, F_n) \quad \text{and} \quad P(i, j, F_n) = P(i \mid j, F_n) P(j \mid F_n) P(F_n). \tag{9}$$

Dividing one by the other, for the case of progress or no progress on SDG j, and cancelling common terms gives

$$\frac{P(j \mid i, F_n)}{P(\bar{j} \mid i, F_n)} = \frac{P(i \mid j, F_n) P(j \mid F_n)}{P(i \mid \bar{j}, F_n) P(\bar{j} \mid F_n)}. \tag{10}$$

The score in the factor map is defined as

$$\alpha_{ij}(F_n) \equiv \log_{\sqrt[N]{b}} \frac{P(i \mid j, F_n)}{P(i \mid \bar{j}, F_n)} = \frac{1}{N} \log_b \frac{P(i \mid j, F_n)}{P(i \mid \bar{j}, F_n)}. \tag{11}$$

The base b is re-scaled to avoid damping or amplifying of scores as the number of factors is increased or decreased. The score in a context map for a context labeled by $\mathbf{F} = \{F_1, F_2, \ldots, F_N\}$ is then

$$\alpha_{ij}(\mathbf{F}) = \sum_{n=1}^{N} \alpha_{ij}(F_n) = \frac{1}{N} \sum_{n=1}^{N} \log_b \frac{P(i \mid j, F_n)}{P(i \mid \bar{j}, F_n)}. \tag{12}$$

$$\Delta\alpha_{ij}(\mathbf{F}) \equiv \alpha_{ij}(\mathbf{F}) - \alpha_{ij}. \tag{13}$$

Empirical estimates of SDG scores

A full representation of the system of interacting SDGs is given by all possible interactions, and not only an evaluation of how progress on SDG *i* is associated with progress on SDG *j*. Following the Bayesian interpretation, that requires the log odds of the ratio of *not* making progress to making progress. The logarithm of the inverse of a number is the negative of the logarithm of that number, so

$$\log_b \frac{P(i \mid \bar{j})}{P(i \mid j)} = -\log_b \frac{P(i \mid j)}{P(i \mid \bar{j})} = -\alpha_{ij}. \tag{14}$$

That is, log odds of *not* making progress on SDG *j* vs. making progress the negative of the score.
More complicated is updating the judgement of the odds ratio given the knowledge that progress has *not* been made on SDG *i*. We define

$$\beta_{ij} = \log_b \frac{P(\bar{i} \mid \bar{j})}{P(\bar{i} \mid j)}. \tag{15}$$

This is the update to the odds of not making progress on SDG *j* given that we have not made progress on SDG *i*.
In practice, on the α scores were computed. However, for completeness, this section describes the construction of both the α and β scores. They can be computed from empirical data arranged in a frequency table. First, a criterion must be applied to the empirical data over what corresponds to "progress", denoted by *p*, vs. "no progress", denoted by *n*. Then, for each case (e.g., a country), each SDG is assigned a value of *n* or *p*. A frequency table for SDGs *i* and *j* can then be constructed by adding up the number of times (i, j) is observed (that is, both are in the state *p*), $(i, \bar{j})$ is observed, $(\bar{\imath}, j)$ is observed, and $(\bar{\imath}, \bar{j})$ is observed. That gives a frequency table

$$F_{ij} = \begin{pmatrix} N_{i,j} & N_{i,\bar{j}} \\ N_{\bar{i},j} & N_{\bar{i},\bar{j}} \end{pmatrix}. \tag{16}$$

Empirical marginal probabilities are then computed by normalizing down the columns,

$$P(i \mid j) = \frac{N_{i,j}}{N_{i,j} + N_{\bar{i},j}} \equiv \frac{N_{i,j}}{N_j}, \quad P(i \mid \bar{j}) = \frac{N_{i,\bar{j}}}{N_{\bar{i},\bar{j}} + N_{i,\bar{j}}} \equiv \frac{N_{i,\bar{j}}}{N_{\bar{j}}}, \tag{17}$$

and so on. The Bayes factors can then be computed, and the scores estimated as

$$\alpha_{ij} = \log_b \frac{N_{i,j}}{N_{i,\bar{j}}} \frac{N_{\bar{j}}}{N_j}, \tag{18}$$

$$\beta_{ij} = \log_b \frac{N_{\bar{i},\bar{j}}}{N_{\bar{i},j}} \frac{N_j}{N_{\bar{j}}}. \tag{19}$$

The construction of factor maps presumes that a set of factors have already been identified and assigned to distinct "levels". The indicator itself my take on a very large number of possible values (for example, GDP per capita in USD to two decimal places), but for the analysis the values are binned into a small number of levels – perhaps up to five (e.g., by quintiles). The SDG database, which includes all countries, is then restricted to the set of countries where factor *n* takes on the value F_n. The score, conditional on the factor level, is then given by

$$\alpha_{ij}(F_n) = \frac{1}{N}\log_b \frac{P(i \mid j, F_n)}{P(i \mid \bar{j}, F_n)}. \tag{20}$$

These values (as well as the corresponding *β* scores) are calculated as above, but restricted to those countries taking on the particular factor score.

Statistical tests

The methodology is new, and there are no standard statistical tests. For the purposes of this project, a rough-and-ready measure of statistical significance, based on a binomial test, was used. It provides an approximate measure of the probability that the true interaction score is not higher or lower than the estimate.

A sample result is shown in Figure 4, for the middle quintile of Gross National Income (GNI). The results are typical, with most interactions scores significant at least at the 5% level, but many insignificant. The results are strongly influenced by data availability; data for SDG indicators remains incomplete, leading to low significance values for some interactions. Where data is available, a low significance score indicates diverse outcomes across the sample.

	SDG1	SDG2	SDG3	SDG4	SDG5	SDG6	SDG7	SDG8	SDG9	SDG10	SDG11	SDG12	SDG13	SDG14	SDG15	SDG16	SDG17
SDG1	NA	0.06	0.05	0.10	0.05	0.10	0.17	0.05	0.14	0.13	0.13	0.13	0.14	0.12	0.05	NA	0.14
SDG2	0.04	NA	0.07	0.15	0.04	0.05	0.15	0.04	0.08	0.07	0.04	0.03	0.08	0.16	0.07	NA	0.16
SDG3	0.14	0.03	NA	0.04	0.14	0.15	0.17	0.14	0.17	0.14	0.14	0.13	0.20	0.04	0.03	NA	0.22
SDG4	0.01	0.01	0.00	NA	0.01	0.01	0.06	0.01	0.11	0.00	0.01	0.01	0.02	0.06	0.00	NA	0.02
SDG5	0.14	0.15	0.14	0.18	NA	0.15	0.11	0.03	0.17	0.14	0.14	0.13	0.04	0.20	0.14	NA	0.04
SDG6	0.08	0.05	0.04	0.08	0.04	NA	0.05	0.04	0.15	0.07	0.07	0.03	0.08	0.16	0.04	NA	0.15
SDG7	0.01	0.01	0.01	0.01	0.00	0.05	NA	0.00	0.04	0.01	0.00	0.00	0.01	0.02	0.00	NA	0.03
SDG8	0.14	0.15	0.14	0.18	0.03	0.15	0.07	NA	0.04	0.14	0.14	0.13	0.04	0.04	0.14	NA	0.04
SDG9	0.04	0.01	0.01	0.12	0.01	0.08	0.07	0.00	NA	0.01	0.01	0.00	0.03	0.03	0.01	NA	0.07
SDG10	0.06	0.03	0.14	0.04	0.14	0.03	0.17	0.14	0.17	NA	0.03	0.13	0.20	0.20	0.14	NA	0.04
SDG11	0.04	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	NA	0.00	0.00	0.00	0.00	NA	0.00
SDG12	0.18	0.05	0.04	0.08	0.04	0.05	0.11	0.04	0.13	0.04	0.04	NA	0.10	0.10	0.04	NA	0.08
SDG13	0.04	0.01	0.00	0.01	0.00	0.01	0.02	0.00	0.01	0.01	0.01	0.00	NA	0.01	0.00	NA	0.00
SDG14	0.01	0.01	0.00	0.01	0.01	0.08	0.03	0.00	0.01	0.01	0.01	0.00	0.01	NA	0.01	NA	0.00
SDG15	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	NA	NA	0.00
SDG16	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SDG17	0.04	0.08	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	NA	NA

Figure 4: Significance scores for middle quintile for Gross National Income: Green means significance at the 1% level, yellow at the 5% level

Annex 6: Aggregate SDG indicators

Aggregate SDG indicators are constructed from indicators in the UN SDG database. They were then combined by identifying linear combinations of the indicators that tended to vary together, the “principal components” in a principal component analysis (PCA) exercise. The exercise resulted in several principal components (PCs) per SDG (with one exception, where there is only one indicator). The loadings for the first three principal components for SDG 1 are shown in Table 11. Of these, some carry significant information as reflected by the eigenvalue, shown in the scree plot shown in Figure 5. For the analysis, all PCs with eigenvalue one or greater were retained. From Figure 5, eight principal components carry some information for SDG 1, and all were retained.

The use of PCs raises some questions of representation. First, how to interpret a change in the value of a principal component. Second, how to reflect the reality of multiple principal components in a uniform way across SDGs. The resolution to these questions is the main focus of this annex.

Table 11: Loadings for first three principal components for SDG 1

Indicator code	Indicator definition	PC1	PC2	PC3
SD_MDP_ANDI	Average proportion of deprivations for people multidimensionally poor (%)	0.045	-0.211	0.340
SD_MDP_MUHC	Proportion of population living in multidimensional poverty (%)	0.117	-0.198	0.144
SD_MDP_MUHHC	Proportion of households living in multidimensional poverty (%)	0.047	-0.219	0.315
SD_XPD_ESED	Proportion of total government spending on essential services, education (%)	0.015	-0.148	0.050
SG_DSR_LGRGSR	Score of adoption and implementation of national DRR strategies in line with the Sendai Framework	0.174	0.139	0.320
SG_DSR_SILS	Proportion of local governments that adopt and implement local disaster risk reduction strategies in line with national disaster risk reduction strategies (%)	0.175	0.099	0.306
SI_COV_BENFTS	[ILO] Proportion of population covered by at least one social protection benefit, by sex (%)	0.443	0.032	-0.074
SI_COV_CHLD	[ILO] Proportion of children/households receiving child/family cash benefit, by sex (%)	0.432	0.022	-0.128
SI_COV_DISAB	[ILO] Proportion of population with severe disabilities receiving disability cash benefit, by sex (%)	0.347	-0.020	-0.104
SI_COV_LMKT	[World Bank] Proportion of population covered by labour market programs (%)	0.044	-0.369	0.102
SI_COV_PENSN	[ILO] Proportion of population above statutory pensionable age receiving a pension, by sex (%)	0.241	-0.017	-0.085
SI_COV_POOR	[ILO] Proportion of poor population receiving social assistance cash benefit, by sex (%)	0.360	0.084	-0.027
SI_COV_SOCAST	[World Bank] Proportion of population covered by social assistance programs (%)	0.053	-0.451	0.149
SI_COV_SOCINS	[World Bank] Proportion of population covered by social insurance programs (%)	0.058	-0.431	0.084
SI_COV_UEMP	[ILO] Proportion of unemployed persons receiving unemployment cash benefit, by sex (%)	0.114	-0.121	-0.172
SI_COV_VULN	[ILO] Proportion of vulnerable population receiving social assistance cash benefit, by sex (%)	0.447	0.026	-0.109
SI_COV_WKINJRY	[ILO] Proportion of employed population covered in the event of work injury, by sex (%)	0.078	0.128	0.208
SI_POV_NAHC	Proportion of population living below the national poverty line (%)	0.011	-0.270	0.178
SP_ACS_BSRVH2O	Proportion of population using basic drinking water services, by location (%)	-0.056	-0.306	-0.418
SP_ACS_BSRVSAN	Proportion of population using basic sanitation services, by location (%)	-0.016	-0.287	-0.406
VC_DSR_DAFF	Number of directly affected persons attributed to disasters per 100,000 population (number)	0.003	0.029	0.122
VC_DSR_LSGP	Direct economic loss attributed to disasters relative to GDP (%)	0.001	0.022	0.119
VC_DSR_MTMP	Number of deaths and missing persons attributed to disasters per 100,000 population (number)	-0.006	-0.004	0.024

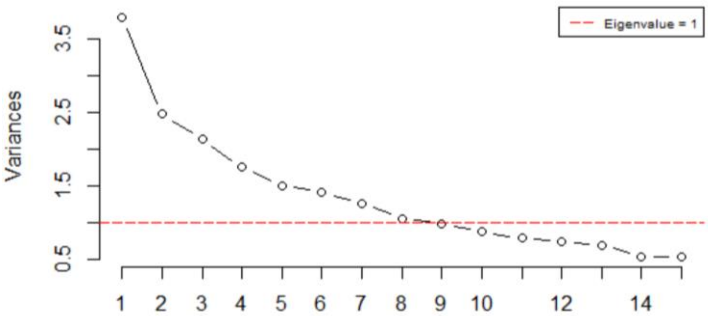


Figure 5: Screeplot for PCs of SDG 1

Interpreting changes in the PCs

For interpreting the PCs, first note that the underlying indicators typically have a “valence”, in that an increase can mean a desirable or undesirable change. Proposed valences for the indicators of SDG 1 are shown in Table 12. So, for example, an increase in “average proportion of deprivations for people multidimensionally poor” would indicate a worsening situation, so it has a valence of −1. In contrast, an increase in the proportion of total government spending on essential services and education would indicate an improvement, so it has a valence of +1.

Table 12: Proposed valences for SDG 1 indicators

Indicator code	Indicator definition	Valence
SD_MDP_ANDI	Average proportion of deprivations for people multidimensionally poor (%)	−1
SD_MDP_MUHC	Proportion of population living in multidimensional poverty (%)	−1
SD_MDP_MUHHC	Proportion of households living in multidimensional poverty (%)	−1
SD_XPD_ESED	Proportion of total government spending on essential services, education (%)	+1
SG_DSR_LGRGSR	Score of adoption and implementation of national DRR strategies in line with the Sendai Framework	+1
SG_DSR_SILS	Proportion of local governments that adopt and implement local disaster risk reduction strategies in line with national disaster risk reduction strategies (%)	+1
SI_COV_BENFTS	[ILO] Proportion of population covered by at least one social protection benefit, by sex (%)	+1
SI_COV_CHLD	[ILO] Proportion of children/households receiving child/family cash benefit, by sex (%)	+1
SI_COV_DISAB	[ILO] Proportion of population with severe disabilities receiving disability cash benefit, by sex (%)	+1
SI_COV_LMKT	[World Bank] Proportion of population covered by labour market programs (%)	+1
SI_COV_PENSN	[ILO] Proportion of population above statutory pensionable age receiving a pension, by sex (%)	+1
SI_COV_POOR	[ILO] Proportion of poor population receiving social assistance cash benefit, by sex (%)	+1
SI_COV_SOCAST	[World Bank] Proportion of population covered by social assistance programs (%)	+1
SI_COV_SOCINS	[World Bank] Proportion of population covered by social insurance programs (%)	+1
SI_COV_UEMP	[ILO] Proportion of unemployed persons receiving unemployment cash benefit, by sex (%)	+1
SI_COV_VULN	[ILO] Proportion of vulnerable population receiving social assistance cash benefit, by sex (%)	+1
SI_COV_WKINJRY	[ILO] Proportion of employed population covered in the event of work injury, by sex (%)	+1
SI_POV_NAHC	Proportion of population living below the national poverty line (%)	−1
SP_ACS_BSRVH2O	Proportion of population using basic drinking water services, by location (%)	+1
SP_ACS_BSRVSAN	Proportion of population using basic sanitation services, by location (%)	+1
VC_DSR_DAFF	Number of directly affected persons attributed to disasters per 100,000 population (number)	−1
VC_DSR_LSGP	Direct economic loss attributed to disasters relative to GDP (%)	−1
VC_DSR_MTMP	Number of deaths and missing persons attributed to disasters per 100,000 population (number)	−1

Multiplying the weightings from Table 11 by the valences from Table 12 for first eight PCs for SDG 1 gives the result in Table 13. The values in Table 3 have been color-coded so that green means improvement, red means worsening, and the darkeness indicates the strength of the response.

In Table 13, PCs are easiest to interpret if they are dominated by a block of dark green or dark red. That is true only for PC1, which captures a cluster of social benefit coverage indicators, and PC6, which captures a cluster of disaster response indicators. The others are ambiguous.

The ambiguity can be reduced by making linear combinations of indicators. However, a uniform method for combining indicators is needed, one that can be replicated across all SDGs.

Table 13: SDG 1 loadings multiplied by valence

Indicator code	Valence	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
SD_MDP_ANDI	-1	-0.04539635	0.210654	-0.33952	0.421676	-0.17897	0.231458	-0.08353	0.115547
SD_MDP_MUHC	-1	-0.11686382	0.197788	-0.14417	0.313301	-0.06186	0.026251	0.235103	-0.06392
SD_MDP_MUHHC	-1	-0.04718497	0.21887	-0.3154	0.378245	-0.19151	0.214522	-0.10777	0.172207
SD_XPD_ESED	1	0.014965893	-0.14847	0.050387	-0.01467	-0.1314	-0.05458	-0.15682	0.692598
SG_DSR_LGRGSR	1	0.173811045	0.139315	0.32041	-0.12846	-0.26671	0.492588	0.044714	0.02169
SG_DSR_SILS	1	0.175459826	0.098776	0.3057	-0.16905	-0.29004	0.478264	0.109573	-0.00183
SI_COV_BENFTS	1	0.442856094	0.031867	-0.0739	0.068113	0.042481	-0.09017	0.151355	-0.01473
SI_COV_CHLD	1	0.432031078	0.021618	-0.12779	0.070655	0.044621	-0.13723	0.172886	0.02588
SI_COV_DISAB	1	0.347028889	-0.02044	-0.10415	-0.10807	-0.05202	0.080389	-0.22155	0.003147
SI_COV_LMKT	1	0.044490897	-0.36928	0.101525	0.301041	-0.01901	0.112832	0.022409	-0.16441
SI_COV_PENSN	1	0.240756376	-0.01748	-0.08466	-0.02693	0.025293	0.007043	-0.5037	-0.0544
SI_COV_POOR	1	0.359675847	0.084093	-0.02684	0.010631	0.025713	-0.0935	0.168452	0.082301
SI_COV_SOCAST	1	0.052856188	-0.45139	0.149456	0.299824	-0.00696	0.057843	0.050715	-0.0456
SI_COV_SOCINS	1	0.058217269	-0.4314	0.084169	0.350645	-0.03313	0.155154	0.002701	-0.12641
SI_COV_UEMP	1	0.114350349	-0.1211	-0.17171	-0.09439	-0.02762	0.050997	-0.56615	-0.01546
SI_COV_VULN	1	0.446625192	0.026404	-0.10922	0.074192	0.042115	-0.12318	0.192069	0.008701
SI_COV_WKINJRY	1	0.078171678	0.127871	0.207647	0.157862	0.184158	-0.01109	-0.31439	-0.29561
SI_POV_NAHC	-1	-0.01058654	0.269908	-0.17841	-0.0574	-0.01921	0.088124	-0.01276	-0.42039
SP_ACS_BSRVH2O	1	-0.0555412	-0.30608	-0.41834	-0.27966	-0.19587	0.086708	0.155583	-0.05209
SP_ACS_BSRVSAN	1	-0.01563127	-0.28691	-0.40561	-0.28916	-0.18761	0.148641	0.098613	-0.17033
VC_DSR_DAFF	-1	-0.00260316	-0.02869	-0.12243	-0.06203	0.557582	0.367621	0.064784	0.169612
VC_DSR_LSGP	-1	-0.00088176	-0.02197	-0.11878	-0.03827	0.563107	0.37837	0.064087	0.123743
VC_DSR_MTMP	-1	0.005593747	0.004283	-0.02382	0.016841	0.085207	0.096994	-0.04282	-0.27711

Proposed method

We propose the following method. First, some notation. For a given SDG, there are n indicators, indexed by $i \in [1,n]$. *E.g.*, for SDG 1, n = 23; for SDG 2, n = 6. The loadings for PC j are denoted by $\ell_i^{(j)}$. The number of statistically significant PCs as interpreted from the scree plot (or by an equivalent calculation, as this can be automated) is denoted n_{sig}. The valences for the indicators are denoted by $v_i = \pm 1$.

Calculate weights w_j for the significant PCs in the following way:

$$w_j = \frac{\sum_{i=1}^{n} \ell_i^{(j)} v_i}{\sum_{i=1}^{n} \left| \ell_i^{(j)} \right|}. \tag{21}$$

This will be large and positive for unambiguous PCs that tend in a positive direction (as for PCs 1 and 6 in Table 13). It will be small for ambiguous PCs. It will be large and negative for PCs that tend in a negative direction. (There are no good examples in Table 13, but PC2 is close.)
Next, construct a single set of loadings for SDG using the weights,

$$\ell_i^{\text{agg}} = \sum_{j=1}^{n_{\text{sig}}} w_j \ell_i^{(j)}. \tag{22}$$

Note that if we now calculate the total valence with these loadings we get the following:

$$\sum_{i=1}^{n} \ell_i^{\text{agg}} v_i = \sum_{j=1}^{n_{\text{sig}}} \frac{\sum_{k=1}^{n} \ell_k^{(j)} v_k}{\sum_{k=1}^{n} \left| \ell_k^{(j)} \right|} \sum_{i=1}^{n} \ell_i^{(j)} v_i = \sum_{j=1}^{n_{\text{sig}}} \frac{\left(\sum_{k=1}^{n} \ell_k^{(j)} v_k \right)^2}{\sum_{k=1}^{n} \left| \ell_k^{(j)} \right|}. \tag{23}$$

This is guaranteed to be positive, so we can interpret an increase in this aggregate PC as "progress".
The application of this method for SDG 1 is shown in Table 14. As can be seen, the interpretation of this aggregate set of loadings is unambiguously measuring "progress". Some indicators turn out not to be very important: proportion of population (as opposed to households) living in multidimensional poverty; proportion of employed population covered in the event of work injury; and the proportion of population living below the national poverty line[8].

Table 14: Aggregate loadings for SDG 1

Indicator code	Indicator definition	Loadings	With valence
SD_MDP_ANDI	Average proportion of deprivations for people multidimensionally poor (%)	-0.356	0.356
SD_MDP_MUHC	Proportion of population living in multidimensional poverty (%)	-0.013	0.013
SD_MDP_MUHHC	Proportion of households living in multidimensional poverty (%)	-0.316	0.316
SD_XPD_ESED	Proportion of total government spending on essential services, education (%)	-0.058	-0.058
SG_DSR_LGRGSR	Score of adoption and implementation of national DRR strategies in line with the Sendai Framework	0.300	0.300
SG_DSR_SILS	Proportion of local governments that adopt and implement local disaster risk reduction strategies in line with nation	0.286	0.286
SI_COV_BENFTS	[ILO] Proportion of population covered by at least one social protection benefit, by sex (%)	0.322	0.322
SI_COV_CHLD	[ILO] Proportion of children/households receiving child/family cash benefit, by sex (%)	0.297	0.297
SI_COV_DISAB	[ILO] Proportion of population with severe disabilities receiving disability cash benefit, by sex (%)	0.383	0.383
SI_COV_LMKT	[World Bank] Proportion of population covered by labour market programs (%)	0.265	0.265
SI_COV_PENSN	[ILO] Proportion of population above statutory pensionable age receiving a pension, by sex (%)	0.303	0.303
SI_COV_POOR	[ILO] Proportion of poor population receiving social assistance cash benefit, by sex (%)	0.193	0.193
SI_COV_SOCAST	[World Bank] Proportion of population covered by social assistance programs (%)	0.216	0.216
SI_COV_SOCINS	[World Bank] Proportion of population covered by social insurance programs (%)	0.342	0.342
SI_COV_UEMP	[ILO] Proportion of unemployed persons receiving unemployment cash benefit, by sex (%)	0.271	0.271
SI_COV_VULN	[ILO] Proportion of vulnerable population receiving social assistance cash benefit, by sex (%)	0.311	0.311
SI_COV_WKINJRY	[ILO] Proportion of employed population covered in the event of work injury, by sex (%)	0.063	0.063
SI_POV_NAHC	Proportion of population living below the national poverty line (%)	-0.089	0.089
SP_ACS_BSRVH2O	Proportion of population using basic drinking water services, by location (%)	0.144	0.144
SP_ACS_BSRVSAN	Proportion of population using basic sanitation services, by location (%)	0.227	0.227
VC_DSR_DAFF	Number of directly affected persons attributed to disasters per 100,000 population (number)	-0.254	0.254
VC_DSR_LSGP	Direct economic loss attributed to disasters relative to GDP (%)	-0.273	0.273
VC_DSR_MTMP	Number of deaths and missing persons attributed to disasters per 100,000 population (number)	-0.117	0.117

Applying these loadings to the underlying data gives a **single indicator per SDG that is unambiguously measuring progress**.

Complications

One potential complication can be seen when looking at the SDG 2 indicators (Table 15). A valence can readily be assigned to most of the indicators: *e.g.*, the valence for AG_PRD_FIEMSI, prevalence of moderate or severe food insecurity in the adult population, is clearly −1. However, the valence for the first indicator, agriculture value added share of GDP, is less clear. As a measure of economic development, a decline in agriculture value added share is reckoned an improvement, suggesting a negative valence. However, in the context of the other indicators, it is not clear this interpretation is

[8] This is less surprising than it might seem, because national poverty line definitions are far from uniform.

correct. For example, if agriculture's share of GDP is very small, then we do not particularly care what happens to the agricultural orientation of government expenditures.

The procedure followed is to take the "plain reading" of each indicator individually. In that case, AG_PRD_AGVAS gets a negative valence. The result is that it is given a low aggregate weight in any case, as shown in Table 16.

Table 15: SDG 2 indicator set

Indicator code	Indicator definition
AG_PRD_AGVAS	Agriculture value added share of GDP (%)
AG_PRD_FIESMSI	Prevalence of moderate or severe food insecurity in the adult population (%)
AG_PRD_FIESSI	Prevalence of severe food insecurity in the adult population (%)
AG_PRD_ORTIND	Agriculture orientation index for government expenditures
AG_XPD_AGSGB	Agriculture share of Government Expenditure (%)
ER_RSK_LBREDS	Proportion of local breeds classified as being at risk as a share of local breeds with known level of extinction risk (%)

Table 16: SDG 2 aggregate loadings when AG_PRD_AGVAS has a negative valence

Indicator code	Indicator definition	Loadings	With valence
AG_PRD_AGVAS	Agriculture value added share of GDP (%)	-0.037	0.037
AG_PRD_FIESMSI	Prevalence of moderate or severe food insecurity in the adult population (%)	-0.465	0.465
AG_PRD_FIESSI	Prevalence of severe food insecurity in the adult population (%)	-0.445	0.445
AG_PRD_ORTIND	Agriculture orientation index for government expenditures	0.886	0.886
AG_XPD_AGSGB	Agriculture share of Government Expenditure (%)	0.361	0.361
ER_RSK_LBREDS	Proportion of local breeds classified as being at risk as a share of local breeds with known level of extinction risk (%)	-0.061	0.061