The Q approach to consensus building: integrating diverse perspectives to guide decision-making

Authors:
Jonas Geschke[1,*], Davnah Urbach[2,3], Graham W. Prescott[1,4], Markus Fischer[1,2]

Affiliations:
[1] Institute of Plant Sciences, University of Bern, Bern, Switzerland
[2] Global Mountain Biodiversity Assessment (GMBA), University of Bern, Bern, Switzerland
[3] Centre Interdisciplinaire de Recherche sur la Montagne, University of Lausanne, Lausanne, Switzerland
[4] The Biodiversity Consultancy Ltd, Cambridge, UK
[*] Corresponding author: jonas.geschke@ips.unibe.ch

Abstract:

1. Decision-making processes are complex and time-intensive, particularly when a consensus needs to be achieved amongst more than two parties. Discussions and negotiations must consider all relevant stakeholders and their individual perspectives on the decision to be taken. Methods for identifying, understanding, and acknowledging divergent perspectives can support successful consensus building. A tool pointing to those perspectives that have a consensus is missing though.
2. Here, we propose a policy support tool to statistically guide the processes of consensus building around sets of goals or statements, using the Q method. Priority rankings of the goals or statements are used to analyze group perspectives. Our Q approach then expands Q method by consolidating the group perspectives and producing a novel consensus priority score indicating the level of consensual preference or priority for each goal or statement.
3. We demonstrate the applicability of our Q approach in a hypothetical prioritization example involving the Sustainable Development Goals (SDGs). Although all 193 United Nation's member states have agreed upon the 17 SDGs, the implementation of sustainable development measures often requires the prioritization of one or more goals. In the example, we use 40 individual stakeholder perspectives to identify which SDGs should be prioritized to successfully achieve the 2030 Agenda. This is, to satisfy most of the 40 people the best way possible, SGDs 4, 8, and 3. It is important to note that every individual perspective matters.
4. The Q approach to consensus building provides a transparent and replicable method to calculate consensus priority scores for goals or statements of interest and identify those that have medium to high consensus. The approach can be applied to a wide range of situations where diverse perspectives and objectives need to be reconciled and synthesized at a range of scales. It can thereby be applied in consensus building processes from subnational to international levels.

Keywords:
environmental governance; sustainability governance; policy-making; consensus building; priority setting; Q methodology

Introduction

The environmental and societal crises we currently experience are multidimensional and syndemic (IPCC 2018; IPBES 2019; Horton 2020; IPBES 2020; Pörtner *et al.* 2021; IPCC 2022). Tackling them requires bold political will and concerted action between and within countries. However, concerted action by a wide range of actors, for example national states, is ambitious, especially as multilateral environmental and sustainability governance processes typically require consensus decision-making. Following Carter (2018) and in the context of sustainability (Bressen 2007; Agrawal *et al.* 2022) and transformative (Pascual *et al.* 2022) governance, we define consensus decision-making as a process involving collaborative discussions and negotiations of interests, values, and ideas among two or more parties, with the aim to achieve an agreement that balances all perspectives consulted. Building consensus between parties therefore is a discursive and time-intensive process. But it is worth the effort, as bringing together different values can be a leverage point for decision-making and successful governance processes (Horcea-Milcu 2022; IPBES 2022). One prominent example from the environmental domain is the United Nations Framework Convention on Climate: At its 2009 Conference of the Parties in Copenhagen, both preparatory and procedural causes led to limitations in the consensus building process and a failure to agree upon a climate treaty (Winkler & Beaumont 2010). Among others, one underlying reason was "the lack of a common system of values and norms, a high threshold for decision-making" (Winkler & Beaumont 2010, p. 642). It took another six years for the climate negotiations to be successful and the Paris Agreement to find consensus (Bernardo *et al.* 2021).

The scientific literature on consensus building and consensus models is diverse but has typically focused on the parties themselves rather than shared preferences among them (see, e.g., Regan, Colyvan & Markovchick-Nicholls 2006; Vogel & Lowham 2007; Still & Gordon 2009; Allen, Metternicht & Wiedmann 2018; Xue *et al.* 2020). But methods for the analysis of decision-relevant preferences and values are available, too. This can for example be cost-benefit analyses, the Delphi technique, focus group discussions, interviews, multi-criteria decision analysis, nominal group technique, or the Q methodology (for more information on the different methods, see, e.g., Martin 2015; Mukherjee *et al.* 2018; IPBES 2022). Here, we focus on the Q methodology (hereafter Q method), as it provides a complete and detailing workflow to explore both the differences and commonalities of preferences of parties.
Q method has originally been developed for psychology to study subjectivity (Stephenson 1935). In recent years, Q method has increasingly been applied (Sneegas *et al.* 2021) to study, for example, ecosystem services (Armatas, Venn & Watson 2016; Hermelingmeier & Nicholas 2017), bio-economy (D'Amato *et al.* 2019), landscape management (Hamadou *et al.* 2016; Langston *et al.* 2019), nature conservation (Bredin *et al.* 2015; Berry *et al.* 2018; Zabala, Sandbrook & Mukherjee 2018), and sustainable development (Barry & Proops 1999; Moser & Baulcomb 2020) including the SDGs (Eppinga, Mijts & Santos 2022).
In Q method, quantitative inferences are drawn from a set of qualitative sorts of statements representing the totality of a societal discourse (Brown 1993; Watts & Stenner 2012; Akhtar-Danesh 2017b; Akhtar-Danesh 2017a): First, the discourse on the topic of interest is analyzed and written up as statements describing all potential viewpoints in the discourse. This set of statements, called concourse, is reduced to a subset of the statements, called Q sample, that is representative for all statements. Then, a group of respondents specifically selected to represent

the full societal discourse on the topic of interest is asked to rank the statements in a gradient from disagreement to agreement, based on their subjective opinion. The ranking is typically accompanied or followed by an interview or a survey to gain insights into the views of the respondents. The respondents' rankings are called Q sorts and used as input for the statistical analysis. Based on a by-person correlation, a factor extraction and rotation are conducted to retrieve group perspectives from the respondents' rankings.

The main purpose of Q method therefore is to identify group perspectives that are as distinct as possible, to analyze their patterns, and to explore underlying causes why the perspectives are as they are (Watts & Stenner 2005). It also investigates common viewpoints within the perspectives, however, does not necessarily offer constructive solutions towards potentially uniting divergent perspectives. This means Q method reveals if two or more group perspectives have a common ranking of a statement, may it be a lower, medium, or higher ranking. It does, however, not result in how far all statements have a shared assessment among all group perspectives.

To address this, we propose a science-based policy support tool to statistically guide priority setting and processes of consensus building around sets of goals or statements. We call it 'the Q approach to consensus building' (hereafter the Q approach). Using Q method, individual perspectives (of the relevant parties) on the goals or statements of interest are analyzed, and representative group perspectives are used for integrating perspectives across scales. A statistic expanding Q method directly points to those goals or statements that have a consensual preference or priority, which can guide decision-making processes. A mapping of group perspectives can further support the identification of minority perspectives and other barriers that may arise throughout discussions or negotiations.

With the Q approach, we foster the use of Q method in environmental and sustainability decision-making, but also beyond. In fact, any decision-making process can possibly benefit by shedding light on diverse perspectives and integrating them in a target-oriented manner. After introducing the methodological workflow of the Q approach, we demonstrate its applicability using a hypothetical case example of prioritizing the Sustainable Development Goals (SDGs). The R functions to run the Q approach are provided as supplementary (see Data and code accessibility for details).

The Q approach to consensus building

In our Q approach to consensus building, we apply a range of major and minor differences to the typical application of Q method (Tab. 1). For example, instead of reducing a pre-analyzed set of statements for the analysis, we use the full set of statements for ranking. The term "statement" here stands for anything that can be ranked. For consensus building in negotiation processes (e.g., in the negotiation of the post-2020 global biodiversity framework, see Discussion) this could be different options to formulate a target. For priority setting in nature conservation or the implementation of sustainable development measures, it could be targets as they stand (e.g., the SDGs, or, looking into the future, the global biodiversity targets succeeding the Aichi Targets). While in Q method only selected participants are invited for the ranking, the Q approach is intended to be open for participation by all relevant parties, i.e. stakeholders. Moreover, in contrast to Q method that ranks in an agreement gradient, the Q approach has the statements ranked in a preference or priority gradient. Also, as decision-makers in real-world settings may have to prioritize goals that are part of a set of goals individually all important, we feel that a question about the "importance" of statements (which is sometimes used in existing Q method studies) is subtly but meaningfully different from the question about the "preference" one has for a different statement or the "priority" a statement needs. This distinction is also made in IPBES (2015, p. 18), where a "preference refers to the […] importance attributed to one entity relative to another one." After collecting the parties' individual perspectives, we apply the standard Q method statistics and complement them by a consensus priority score for each statement (see details below). With this additional score, we aim to analyze the commonalities of group perspectives and to identify those statements having a shared medium to high ranking.

Tab. 1: Major (blue) and minor (green) differences between a typical Q method application and the Q approach to consensus building.

Step of the analysis and methodological decision therein	**Q method**	**The Q approach to consensus building**
Aim of the analysis	Identification and understanding of distinct group perspectives	Identification of goals or statements in a decision-making process that have a consensual preference or priority
Collection of statements to be sorted/ranked	A large, pre-analyzed set of statements (the concourse) is reduced to a representative subset of statements (the Q sample)	A meaningful number of statement options (e.g., versions of a draft text), or the full set of statements (if already agreed upon)
Selection of participants	Selection of respondents to represent the full discourse on the topic of interest	Any stakeholder relevant for the discussion or negotiation process
Data collection: the sorting/ranking of statements	In a forced quasi-normal distribution, ranging from "disagree" to "agree"	In a forced quasi-normal distribution, ranging from "lowest preference or priority" to "highest preference or priority"
Identification of how many factors to analyze	Based on a principal component analysis	Based on five hierarchical criteria (see figure 1e and details in the text)
Factor rotation	Varimax method (typically)	Quartimax method
Statistics used for the interpretation of results	Factor eigenvalue, factor flagging, and z-scores	Factor eigenvalue, factor flagging, z-scores, and consensus priority scores
Identification of consensus between factors	Using the z-scores: consensus statements that have a similar ranking (vs. significantly divergent statements)	Using the consensus priority scores: statements that have a shared medium to high ranking
Analysis across scales/levels	Not done (typically)	Using resulting lower-level group perspectives as input for upper-level analysis

For presenting the workflow of the Q approach to consensus building (*a.* to *i.*, Fig. 1), we use both the Q method technical terms, i.e. Q sorts and factors, and a Q approach synonym (Tab. 2). For example, we use the term "ranking(s)" for one or more Q sorts, which can be individual rankings of or group perspectives, and the term "group perspective(s)" for one or more factors. The workflow is implemented using the R packages "qmethod" (Zabala 2014; Zabala & Pascual 2016) and, for network visualizations, "igraph" (Csárdi & Nepusz 2006).

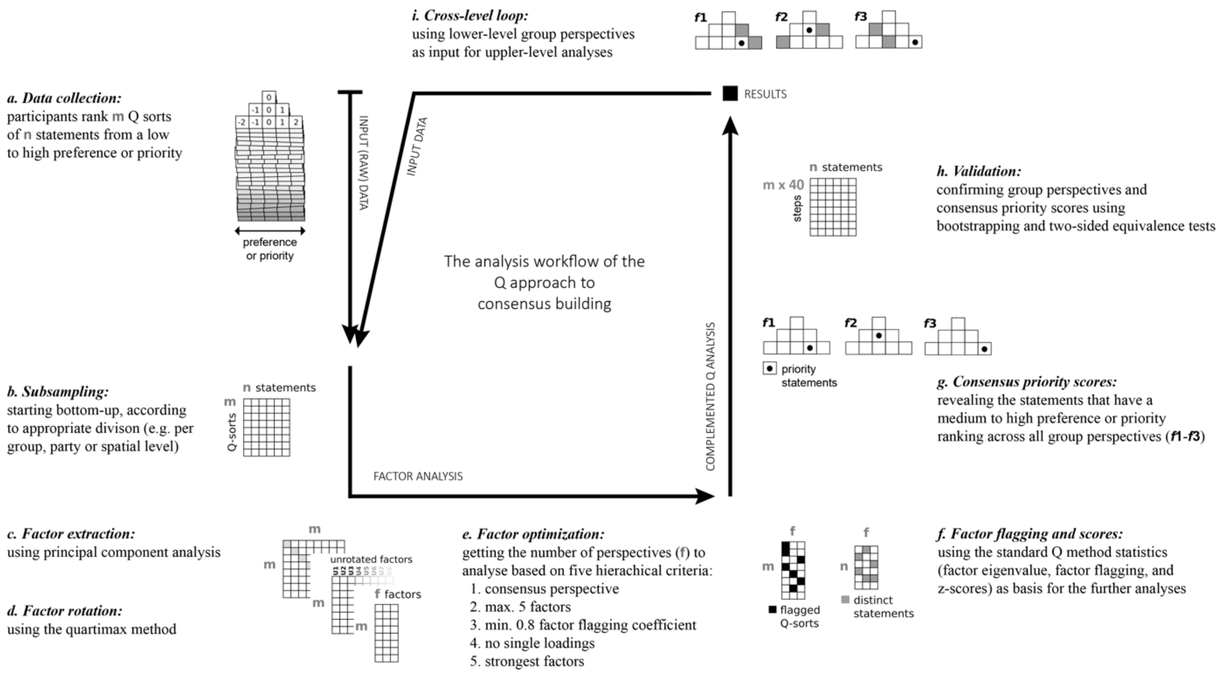

Fig. 1: The analysis workflow of the Q approach to consensus building, anti-clockwise: The data collection (*a*) can be done through, for example, stakeholder workshops or online tools. Subsampling (*b*), factor extraction (*c*) and factor rotation (*d*) are typical steps of a Q method analysis. In the Q approach, we use five hierarchical criteria to optimize the number of factors, i.e. group perspectives, for consensus building (*e*). Factor flagging and scores (*f*) represent the standard Q method statistics. In the Q approach, the consensus priority scores (*g*) are a novel value to guide decision-making processes. Validation of the results (*h*) is done through comparing different group perspective variables and the consensus priority scores. The cross-level loop (*i*) allows re-using lower-level group perspectives in upper-level analyses. The icons are from Zabala and Pascual (2016), but have been partly modified.

Tab. 2: The technical Q method terminologies, what synonyms we use for presenting the Q approach, and what they are.

Q method technical term	Q approach synonym	What this is
Q sort(s)	ranking(s) / individual perspective(s)	The participants' ranking of statements, used as input data for the analysis
factor extraction	-	Based on a by-person correlation, unrotated factors are extracted from the Q sorts using principal components analysis (PCA)
factor rotation	-	After the factor extraction, to receive more structured and better interpretable group perspectives, unrotated factors are rotated using the quartimax method
factor loadings	-	The correlation values of the Q sorts with the rotated factors
single loading	-	A factor that has only one Q sort loading to it
(factor) flagging	agreement	The significantly highest loading of a Q sort tells which factor(s) the Q sort flags/agrees to
factor flagging coefficient	percentage agreement	The percentage of Q sorts flagging to one or more factor(s)
factor(s)	group perspective(s)	The perspectives resulting from the analysis i.e. the factor extraction and rotation. They may be used as input data for upper-level analyses

a. Data collection

All participants, i.e. the parties relevant to the discussion or negotiation, create an individual Q sort, wherein they rank a set of statements along a pre-defined preference or priority gradient. This is typically done through workshops, interviews, or online tools. During the rank, if appropriate time is available, the participants reflect on their specific situation regarding the topic of interest and decide on a lowest and a highest preference or priority statement. A forced quasi-normal distribution centered around zero is used for the ranking, as a free distribution does not have a noticeable contribution to the resulting factors (Brown 1980; Watts & Stenner 2005) and a forced distribution makes both the ranking and the analysis easier (Watts & Stenner 2005). Further, from our point of view, a priorization exercise loses in value if participants can give all statements the same (or almost the same) “high” preference or priority, which could be the case if participants are allowed to rank all statements in just one or a few gradient levels.

b. Subsampling

In case the analysis is intended to contain one level only, the full dataset is used for the analysis. Yet, the Q approach is particularly meant to allow analyses including multiple, usually nested levels. That can be social (e.g., an individual level and one or more group levels) or spatial levels (e.g., local, national and international level). In such cases, one or more lower-level subsamples of the Q sorts are analyzed first. The order in which subsamples are analyzed across levels always must be bottom-up, as the resulting factors are included in or combined for the upper-level analyses (➔ *i. Cross-level loop*). Every subsample, no matter on what level, needs a full analysis along the presented workflow. If necessary, weighting can be applied (e.g., income equity weighting) for a more just valuation of preferences or priorities across scales (IPBES 2022). Subsequently, we use the term “sample” for the Q sorts used in each analysis, irrespective of the analysis level.

c. Factor extraction

First, and in contrast to R statistics, where variables are correlated (Rost 2020), in Q method a by-person correlation is conducted, using the Q sorts (Watts & Stenner 2005; Zabala & Pascual 2016).

Next, in Q method, unrotated factors are extracted from a sample, typically using a principal component analysis (PCA). A PCA provides the maximum variance for all factor from the sample, which can be summed up to a cumulative explaining variance of the factors in the sample (Zabala & Pascual 2016). Typically, the number of factors that are further analyzed is then selected based on certain criteria such as their eigenvalue or their cumulative explaining variance (Akhtar-Danesh 2017b). Another way of getting the number of factors to further analyze is through a scree test (Cattell 1966) of the unrotated factors, which shows how many factors have a proportion of the total variance bigger than an alpha threshold. A threshold of $\alpha = 0.05$, for example, would mean that only unrotated factors representing at least 5% of the total variance are considered in the further analysis. Adding a linear model of the PCA eigenvectors to a scree plot of unrotated factors can visually indicate how many principal components a sample might have. In the Q approach to consensus building, we use such a visualization of unrotated factors for general orientation on how many factors there could be; but in fact pre-analyze an optimal number of group perspectives for guiding decision-making. This is described in *e*.

d. Factor rotation

For the rotation of factors, we use the quartimax method. This is different to the typically used rotation method varimax, which results in factors as distinct from each other as possible (Zabala 2014). Yet it ensures that each Q sort flags to the minimum number of factors, excludes a smaller number of Q sorts in the factors, has a smaller number of distinguishing statements within the factors, and generates a general factor among the Q sorts (Hair *et al.* 2014; Akhtar-Danesh 2017a; Akhtar-Danesh 2017b). This results in a higher consensus amongst participants.

e. Factor optimization

Q method studies typically use PCA measures to determine the number of factors to rotate (see *c*). In the Q approach, however, as we ultimately seek to support consensus building in a concrete decision-making process, we aim at a limited number of factors representing the major group perspectives from the discourse, discussion, or negotiation. Contrasting the cumulative variance criterion, we aim at having as many participants as possible flag (i.e. agree to a group perspective), with the potential to subsequently gain the agreement of the remaining participants. "Group" perspectives to which only one participant agrees are not irrelevant but not necessarily conducive to consensus building.

Therefore, we run and compare multiple factor extractions and rotations to determine how many factors result in the highest consensus, with an optimization towards a minimum number of group perspectives and a maximum statistical agreement of the rankings to the group perspectives. As outlined, we apply the following five hierarchical criteria:

1. First, we test if the sample contains a consensus perspective with a factor flagging coefficient of 1.

   If this is not the case, we continue with the following set of criteria:
2. There should be a maximum of five factors,
3. The factor flagging coefficient should have a minimum of 0.8, meaning that at least 80% of the input rankings should agree to the group perspectives,
4. There should be no single loading, meaning that all group perspectives should have at least two rankings agree, and
5. The first, second, third, etc. factor should always be as strong as possible, meaning that a higher number of rankings agreeing to the first, second, third, etc. group perspective is preferred.

f. Factor flagging and scores

To support the understanding of diverse perspectives and preferences or priorities among participants, and to acknowledge different viewpoints, the standard Q method statistics (as explained in, e.g., Zabala 2014) are applied:

Factor flagging, based on the factor loadings, indicates to which group perspective a ranking significantly agrees. The z-score is a measure to compare the ranking of statements within and across group perspectives. A comparison of z-scores reveals distinguishing and consensus statements in the group perspectives. The factor eigenvalue represents the strength of a group perspective compared to the other group perspectives within an analysis.

Based on these measurements, Q method is predestined for the analysis of consensus building processes and has been used to do so. Both Vogel and Lowham (2007) and Eppinga, Mijts and

Santos (2022) ran a cluster analysis on their study participants Q sorts to identify participants with shared believes. Rust (2016) combined Q method and the Delphi method to explore options for consensus building. In a review on the use of Q method in environmental sustainability research, Sneegas *et al.* (2021) found the majority of their literature corpus exploring not only differences but also agreement between perspectives, with varying detail. Therefore, we do not consider the Q approach as a completely new methodology but as a complemented approach to utilize Q method to directly feed into decision-making processes using the consensus priority score, not merely focusing on the respondents (who belongs to what group perspective, and how can the group be characterized) but on the statements that need to be prioritized (described in *g*).

g. Consensus priority score

To complement the standard Q method statistics, we developed the consensus priority score (cp-score) that results in one number per statement, regardless of how many factors are to be unified. All statements have one consensus priority value each, to be understood as a set. Thus, the consensus priority values only exist in association with their respective statement. Larger differences between two consensus priority values also represent a gap in their preference or priority. The consensus priority scores therefore reveal those statements that have a consensual medium to high preference or priority ranking among all group perspectives. Due to its harmonized gradient, consensus priority scores are comparable throughout analyses.

Mathematically, the consensus priority score contains the z-scores per statement per factor and the eigenvalue of each factor. First, to incorporate the strength of factors, for every statement and factor the z-scores are multiplied with the respective factor's eigenvalue, resulting in a weighted z-score per statement and factor: $x_{s,f} = \text{z-score}_{s,f} * \text{eigenvalue}_f$

Next, to get one value per statement, the arithmetic mean of weighted z-scores across the factors is calculated: $y_s = \sum \frac{x_{s,f}}{f}$

In the third and final step, to get a priority gradient between 0 (least preference or priority) and 1 (highest preference or priority), the mean of the weighted z-scores across the statements is normalized: $\text{cp-score}_s = \frac{y_s - \min(y)}{\max(y) - \min(y)}$

A numerical example of the calculation process is provided in table S1.

h. Validation

The factor scores, including the consensus priority score, and their validation offer insights to the discourse represented by the group perspectives and can be used as indicators of how individual statements shape perspectives. To validate the stability of, and gain more insights in, the resulting factors, they are bootstrapped with a minimum of 40 times the number of input rankings as the number of steps (Zabala & Pascual 2016). This allows comparing the position of statements in the factors, evaluating the stability of group perspectives, and provides bootstrapped factor scores for the validation of the consensus priority scores. A similar bootstrapping was used for validation by Eppinga, Mijts and Santos (2022).

The validation of the group perspectives and the consensus priority scores is done in three ways, each complementing the others but the third to be considered most robust:

1. Comparison of the consensus priority scores and the input rankings' means ("input means"): Here, the rankings' mean values per statement are normalized to get a gradient between 0 and 1 that is comparable with the consensus priority scores. Then, the position of statements is compared, and a paired two one-sided equivalence test (TOST) is conducted to test the significance of the consensus priority scores.
2. Comparison of the resulting factors and the consensus priority scores with the bootstrap results ("bootstrap factors" and "bootstrap cp-score"): Here, the position of statements in the resulting factors is compared with the bootstrap factors. Then, the bootstrap factor scores are used to calculate bootstrap cp-scores, and a paired TOST is conducted to test the significance of the consensus priority scores against the bootstrap cp-scores.
3. Bootstrapping of the consensus priority scores ("bootstrapped cp-scores"): Here, the full results of the bootstrap are used to calculate the consensus priority scores for each of the bootstrap steps. Then, the median, mean, and standard deviation are calculated for each of the statements' bootstrap results. Based on the outcome of a Shapiro-Wilk test, a Wilcoxon rank-sum test or t-test is conducted to test the significance of each statements' consensus priority score against the statements' bootstrapped cp-score. Finally, a TOST is conducted to test the significance of the consensus priority scores against the bootstrapped cp-scores. Here, the individual statements' consensus priority scores are validated, allowing further insights to where the statements' preference or priority has consensus or not.

All three ways of validation are positive if the respective TOST null hypothesis of statistical difference is rejected.

Case example: Which SDGs should be prioritized to successfully achieve the 2030 Agenda?

One context in which the Q approach to consensus building could be particularly relevant is sustainable development. To illustrate this, we take a discussion on the Sustainable Development Goals (SDGs) as an example. In the SDGs, all 193 members of the United Nations agreed to a complex set of goals, which together should enable the achievement of the "2030 Agenda for Sustainable Development" (UNGA 2015). With the pledge of leaving no one behind, the ambition of the 2030 Agenda is high and nominally inclusive to all people across the world. As framed in their inception, all SDGs are important. However, they are the outcome of a global political negotiation process (Le Blanc 2015) and "effectively a non-binding set of global aspirations with weak institutional oversight arrangements and high levels of national discretion with respect to priorities" (Hirons 2020, p. 324). Consequently, their political impact has been more discursive than transformative (Biermann *et al.* 2022). This may also be as tools and analytical approaches to implement the SDGs in a coherent, integrated, and dynamic manner have yet to be developed (Allen, Metternicht & Wiedmann 2018).

Nevertheless, the SDGs are the one framework that is globally set and that science, policy, and all other stakeholders such as non-governmental organizations have to work with. Therefore, in the decision-making on and the implementation of sustainable development measures, a subset of SDGs is often prioritized over others (Allen, Metternicht & Wiedmann 2018; Horn & Grugel 2018; Forestier & Kim 2020). This is, among others, due to the wide array of sectors the SDGs address (EAT 2016; Muff, Kapalka & Dyllick 2017; Tremblay *et al.* 2020), the different importance of certain SDGs across spatial scales (Payne *et al.* 2020), and the diversity of stakeholders that need to be consulted and involved, each with potentially different perspectives on which SDGs should or need to be prioritized (Allen, Metternicht & Wiedmann 2018). Stakeholder groups needed to successfully achieve sustainable development include, among others, academia, the private sector, non-governmental organizations, and civil society, and their perspectives must be recognized to enable synergies (IPBES 2022).

In our case example, which is hypothetical, 40 people from two groups (for the purpose of illustrating the application of the Q approach across scales) enter a discussion about which SDGs to focus on in a new project. The groups could be country representatives at a negotiation table, two non-governmental organizations or subgroups within one non-governmental organization (e.g. from different places), two divisions of a company, or whatever else. The Q approach to consensus building can help consolidating the groups' perspectives and point towards what SDGs the project could strategically focus on.
The example is only used for illustrative purposes and for a methodological discussion of the Q approach's workflow. The 40 perspectives used in the example are responses randomly taken from Geschke *et al.* (in preparation). The interpretation of the results remains without contextual background of the perspectives, therefore neglecting information such as the respondents' age, gender, or work experience.

To start, all 17 SDGs are taken as statements and ranked in a gradient ranging from -3 to +3 (-3, -2, -2, -1, -1, -1, 0, 0, 0, 0, 0, 1, 1, 1, 2, 2, 3). The same distribution was used by Eppinga, Mijts and Santos (2022). For the analysis of more or fewer than 17 statements, different distributions may be needed and/or useful. Nevertheless, aiming at a clear priority statement, we recommend having only one statement space in the gradient extremes (in this case -3 and +3). The 40 individual perspectives are then divided into two samples, groups A and B, the rankings of which are provided in table S2. The group perspectives of groups A and B are used for a synthesis analysis (Fig. 2).

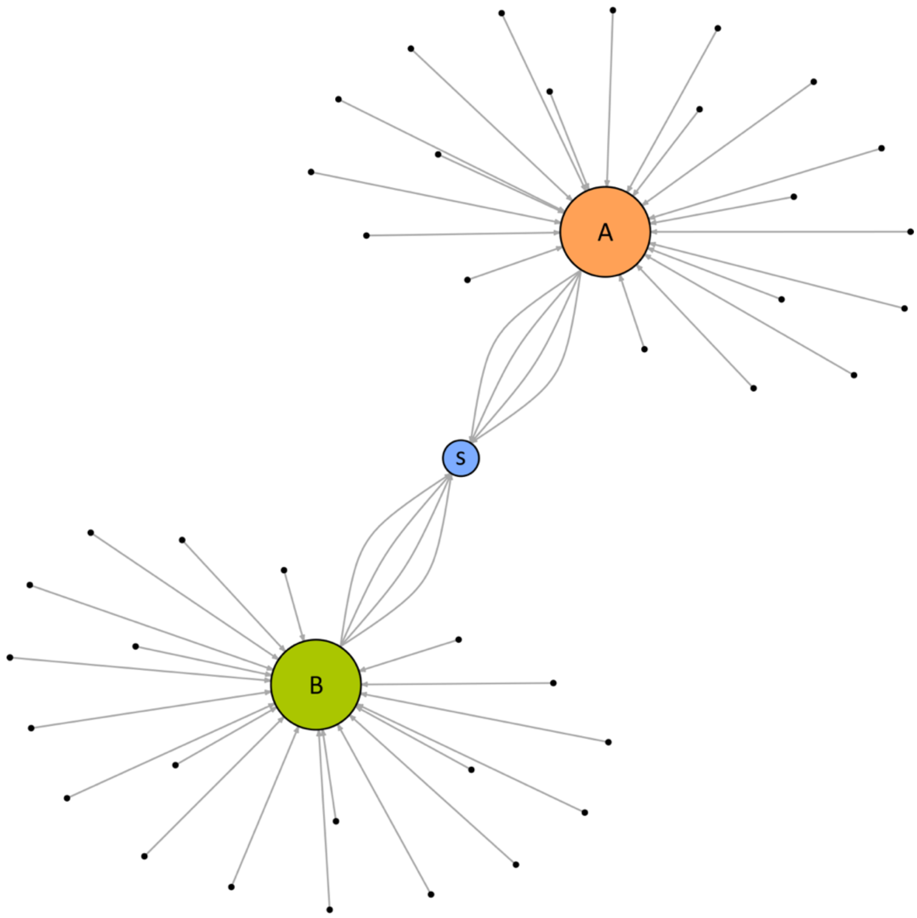

Fig. 2: Conceptual network of how the individual people (black dots) feed into the group analyses (orange circle = group “A”, and green circle = group “B”) that are used for the synthesis analysis (blue circle, “s”). Each arrow throughout the cross-level analysis, i.e. from the orange and green circles to the blue circle, represents one perspective. The circle sizes correspond to the number of incoming perspectives.

Group A

Based on a scree plot of unrotated factors including a linear model of their PCA eigenvalues, group A has 8 major factors within its set of responses. However, the inclusion criteria of the factor optimization (see *e*) lead to only 4 factors being analyzed as group perspectives (A1-A4). 18 of the 20 individual perspectives flag to the group perspectives, resulting in a factor flagging coefficient of 0.9. The group perspectives are shown in figure 3 and numerically provided in tables 3 and S3. The three SDGs with the highest consensus score and on which group A therefore could focus are SDGs 13 *Climate action*, 15 *Life on land,* and 4 *Quality education*. The full list of consensus priority scores is provided in tables 3 and S4.

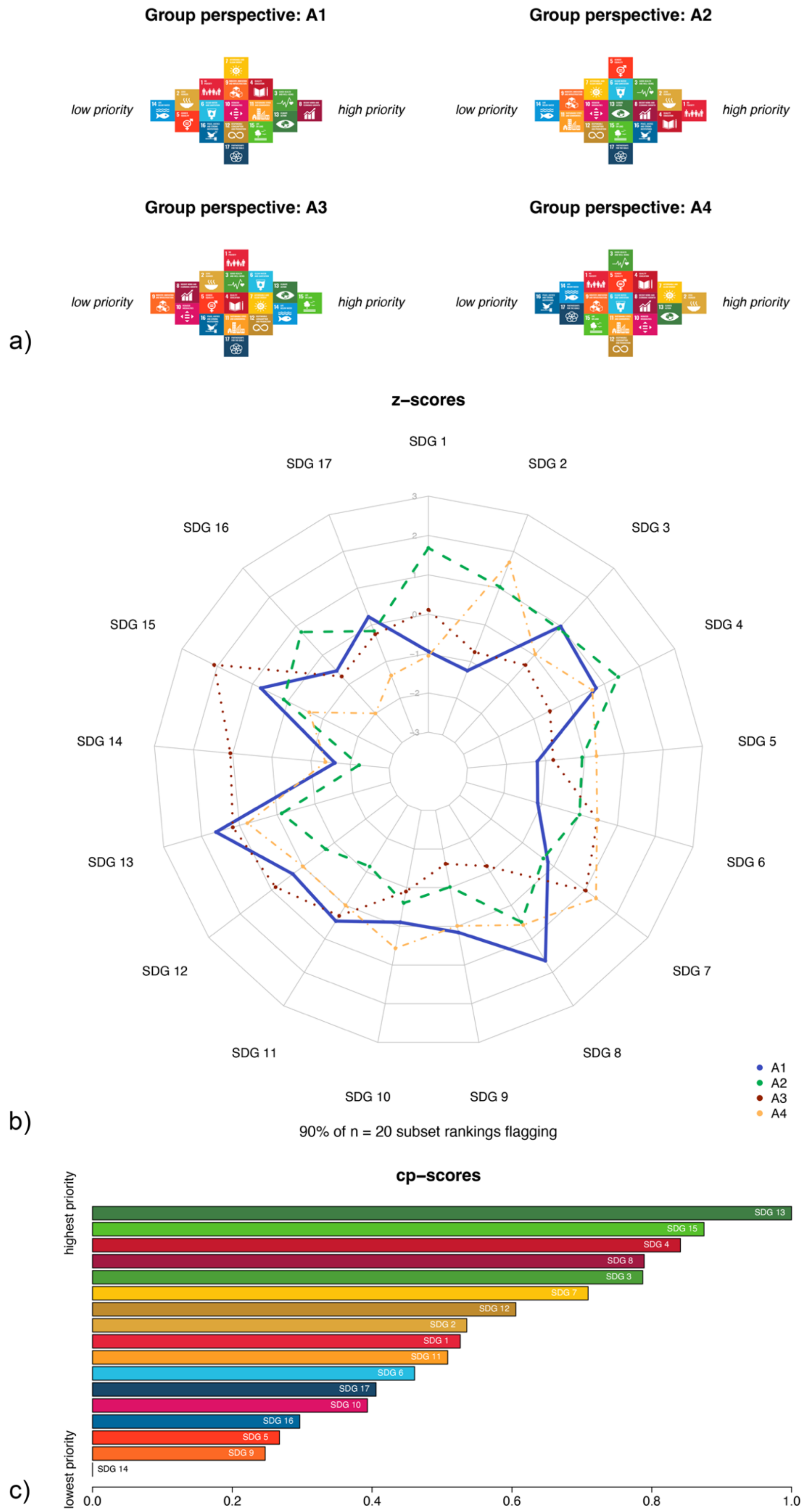

Fig. 3: The results of the Q approach to consensus building for group A, with a) the 4 group perspectives, b) a bar plot of the consensus priority scores (cp-scores) of the group, colored in the SDG color each bar represents, and c)

a radar chart providing further insights to the different group perspectives, showing the z-scores per SDG and perspective.

Factor A1-A4 = group perspectives A1-A4. The line widths within the radar chart represent the factor eigenvalue, i.e. the perspective strength. The line colors and dash types represent each perspective.

Next, we validate the results of group A along the three ways explained in workflow step *h* (Tab. 3): When comparing the consensus priority scores with the input means (Tab. 3a, see *h*.1 for details), the three top consensus priority SDGs match the highest means. Compared with the input means, the consensus priority scores have a significant equivalence (p = 1.39474e-17, ***). The comparison of the group perspectives with the bootstrap results (Tab 3b, see *h*.2 for details) reveals that perspectives A1 and A2 each have two SDGs with a change of one position in the bootstrap factors (i.e. they swap their ranking position). Group perspectives A3 and A4 have four SDGs with a change of one position. This indicates that the perspectives are relatively stable. Looking at the consensus priority scores, the top three consensus priority scores are among the five highest bootstrap cp-scores, allowing us to trust their high priority assessment. When statistically validating the consensus priority scores of group A, they appear equal to the bootstrap cp-scores (p = 2.95819e-23, ***). Bootstrapping the consensus priority scores (Tab 3c, Fig. 4, see *h*.3 for details) results in all the individual SDGs' consensus priority scores being significant except SDG 1 *No poverty* (p = 0.056) and SDG 6 *Clean water and sanitation* (p = 0.72). Their priority position may need a more detailed discussion in group A. Regardless of this, the full set of consensus priority scores is significantly equal to the bootstrapped cp-scores (p = 3.59019e-11, ***).

Tab. 3: Validation of the group perspectives and consensus priority scores, using the group A results as example: a) the input rankings' means (see *h*.1 in the text), b) the bootstrap results of the factors and the consensus priority scores (see *h*.2 in the text), and c) the p-value for each SDG's consensus priority score, based on the bootstrapping results. Significance indicates a positive validation of the SDG's consensus priority scroe (see *h*.3 in the text). The factors (Factor A1-A4 and Bootstrap A1-A4) are in a -3 to +3 gradient, with the SDGs having a position change in the bootstrap results compared to the group A results marked bold. The cp-scores and the input means are in a 0 to 1 gradient, with each the top three SDGs marked bold.
Factor A1-A4 = group perspectives A1 to A4 / cp-scores = (original) consensus priority scores / Bootstrap A1-A4 = group A bootstrap factors 1 to 4 / B cp-scores = bootstrap consensus priority scores / n.s. = not significant

	Group A results					**a)**	**b)**					**c)**
	Factor A1	**Factor A2**	**Factor A3**	**Factor A4**	**cp-scores**	**Input means**	**Bootstrap A1**	**Bootstrap A2**	**Bootstrap A3**	**Bootstrap A4**	**B cp-scores**	**p-value (significance)**
SDG 1	-1	3	0	-1	0.53	0.62	-1	3	0	-1	0.58	0.056 (n.s.)
SDG 2	-2	2	-1	3	0.54	0.38	-2	**1**	-1	3	0.50	0 (***)
SDG 3	2	1	0	0	0.79	0.72	2	**2**	0	**-1**	0.80	0 (***)
SDG 4	1	2	0	1	**0.84**	**0.80**	1	2	**-1**	1	0.82	0 (***)
SDG 5	-2	0	-1	0	0.27	0.18	-2	0	**-2**	0	0.30	0 (***)
SDG 6	-1	0	1	0	0.46	0.54	-1	0	1	0	0.50	0.715 (n.s.)
SDG 7	0	-1	1	2	0.71	0.64	0	-1	1	2	0.67	0 (***)
SDG 8	3	1	-2	1	0.79	0.72	**2**	1	-2	**2**	**0.84**	0 (***)
SDG 9	0	-2	-3	-1	0.25	0.38	0	-2	-3	**0**	0.33	0 (***)
SDG 10	0	-1	-2	1	0.39	0.32	0	-1	**-1**	1	0.43	0 (***)
SDG 11	1	-2	0	0	0.51	0.58	1	-2	0	0	0.51	0.0001 (***)
SDG 12	0	-1	1	0	0.61	0.58	0	-1	1	0	0.58	0 (***)
SDG 13	2	0	2	2	**1.00**	**1.00**	**3**	0	2	**1**	**1.00**	0 (***)
SDG 14	-3	-3	2	-2	0.00	0.00	-3	-3	2	-2	0.00	0 (***)
SDG 15	1	0	3	-1	**0.87**	**0.96**	1	0	3	-1	**0.85**	0 (***)
SDG 16	-1	1	-1	-3	0.30	0.50	-1	1	**0**	-3	0.40	0 (***)
SDG 17	0	0	0	-2	0.41	0.60	0	0	0	-2	0.51	0 (***)

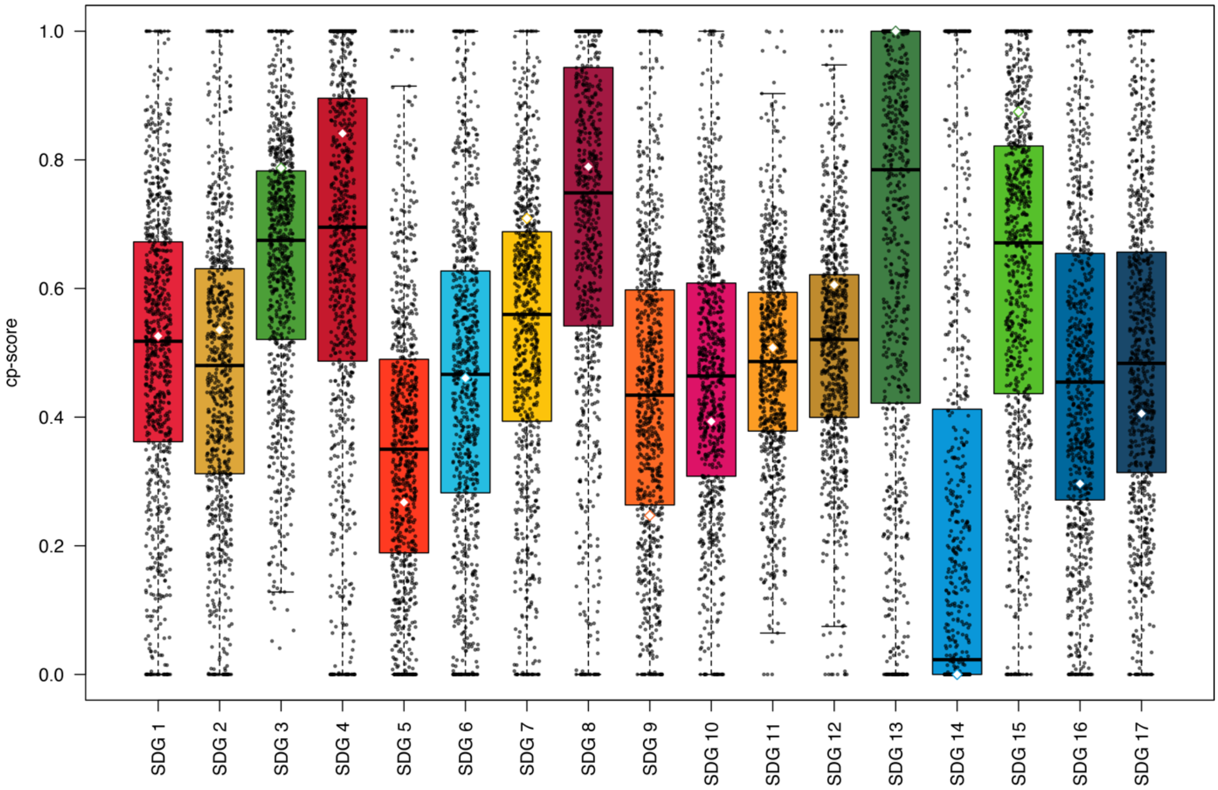

Fig. 4: Boxplots of the bootstrapped consensus priority scores (bootstrapped cp-scores), using the group A results as example. The black dots show the distribution of the bootstrapped cp-scores, each dot representing one bootstrap step. The color-coded white diamonds represent the consensus priority scores of group A. Here, SDGs 13, 15 and 4 have the highest consensus priority scores. The consensus priority score of SDG 15, however, lies outside the first and third quartile of the bootstrapped cp-scores, pointing towards potential discussion needs within group A.

Group B

As with group A, group B has 8 unrotated major factors indicated, but based on the inclusion criteria only 4 group perspectives (B1-B4) are analyzed. With 17 of 20 rankings flagging, they have a factor flagging coefficient of 0.85. Within the group perspectives, there is consensus on the ranking of SDG 16 *Peace, justice and strong institutions*, which has a medium ranking and the fifth lowest consensus priority score. The three top consensus priority SDGs resulting from group B are SDGs 4 *Quality education*, 2 *Zero hunger*, and 13 *Climate action* (Fig. 5). The group perspectives and the full list of consensus priority scores are provided in table S3 and S4.

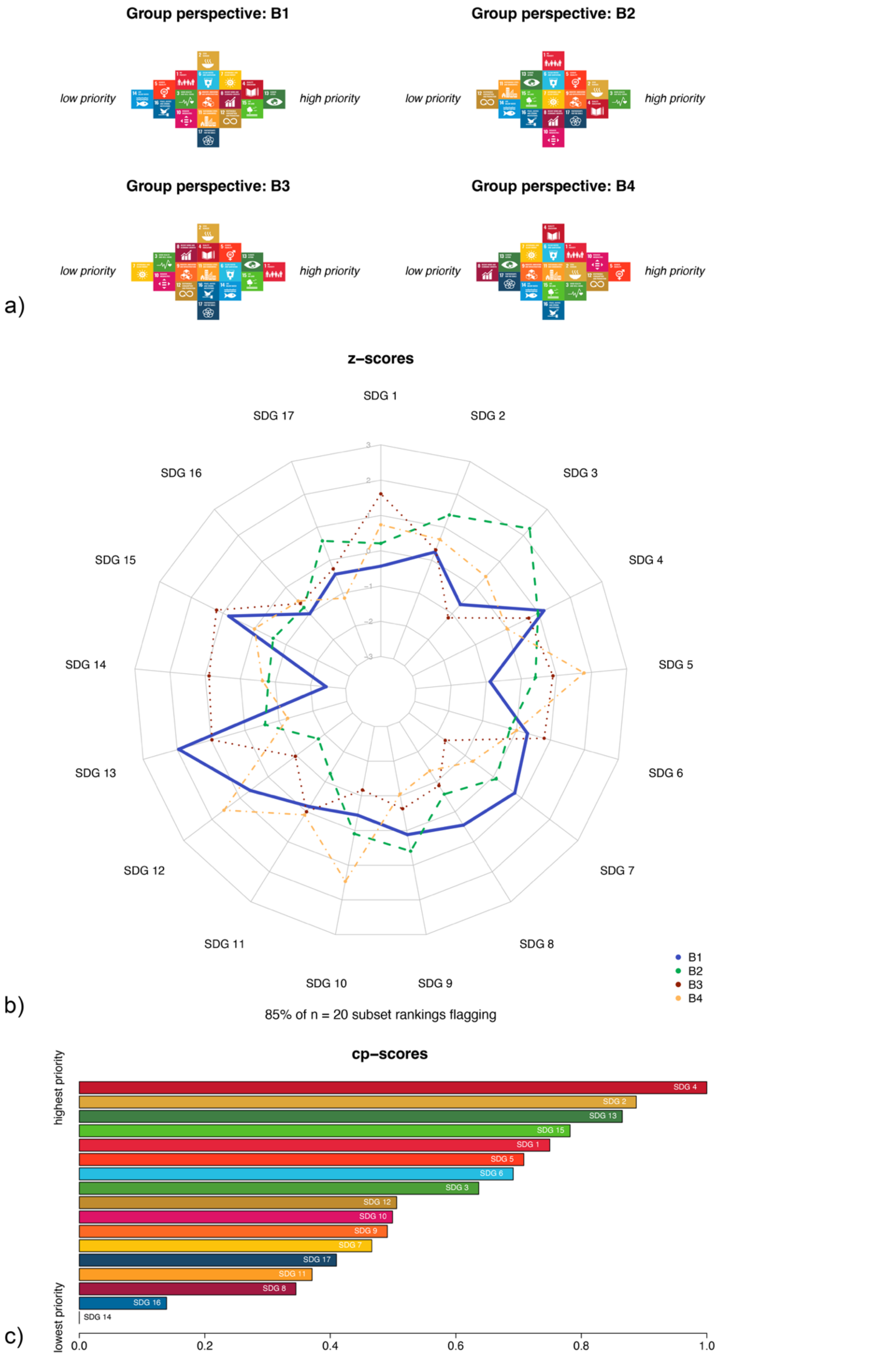

Fig. 5: The results of the Q approach to consensus building for group B, with a) the 4 group perspectives, b) a bar plot of the consensus priority scores (cp-scores) of the group, colored in the SDG color each bar represents, and c)

a radar chart providing further insights to the different group perspectives, showing the z-scores per SDG and perspective.

Factor B1-B4 = group perspectives B1-B4. The line widths within the radar chart represent the factor eigenvalue, i.e. the perspective strength. The line colors and dash types represent each perspective.

In the validation, the comparison of the group perspectives with the input means, group B only has consensus that SDG 14 should have the least priority (cp-score = 0). All other SDGs' consensus priority scores have a change in position compared with the input means, with SDG 2 (second highest cp-score) having its input mean at position 13. Nevertheless, the consensus priority scores are significantly equal to the input means (p = 1.28007e-13, ***). The complexity of the discussion in group B, indicated by the input means, is confirmed by the bootstrap results. Compared with the bootstrap factors, group perspective B1 has two SDGs with a position change value of one. Group perspective B2 instead has three SDGs with a change of one position and one SDG with a change of two positions. Group perspective B3 has four SDGs with a change of one position and one SDG with a change of two positions. Group perspective B4 has ten SDGs with each a change of one position – this is more than half of all SDGs. However, the change in most cases is of one position only, still allowing to consider the priority assessment stable. Also, group perspective B4 is the least strong among the four. Looking at the five highest bootstrap cp-scores, they have the top three consensus priority SDGs among them. Like in group A, this is allowing us to trust their high priority assessment. This is confirmed by the consensus priority scores' test of equivalence, which results in p = 5.82921e-21 (***). When bootstrapping the consensus priority scores, both all individual and the full set of consensus priority scores (p = 6.71306e-11, ***) are significant.

Synthesis consolidating the group perspectives

After the group discussions on the lower level have been analyzed, their group perspectives are brought to an upper-level synthesis analysis (see Fig. 2), supporting the consensus building process which SDGs to focus on the overarching level. Therefore, the four group perspectives from group A and the four group perspectives from group B are taken as sample to be analyzed in the synthesis. While having 6 major factors indicated, the hierarchical criteria of the factor optimization result in 2 group perspectives only (s1 and s2). And indeed, with a factor flagging coefficient of 1, all lower-level group perspectives flag to the two synthesis perspectives (Fig. 6).

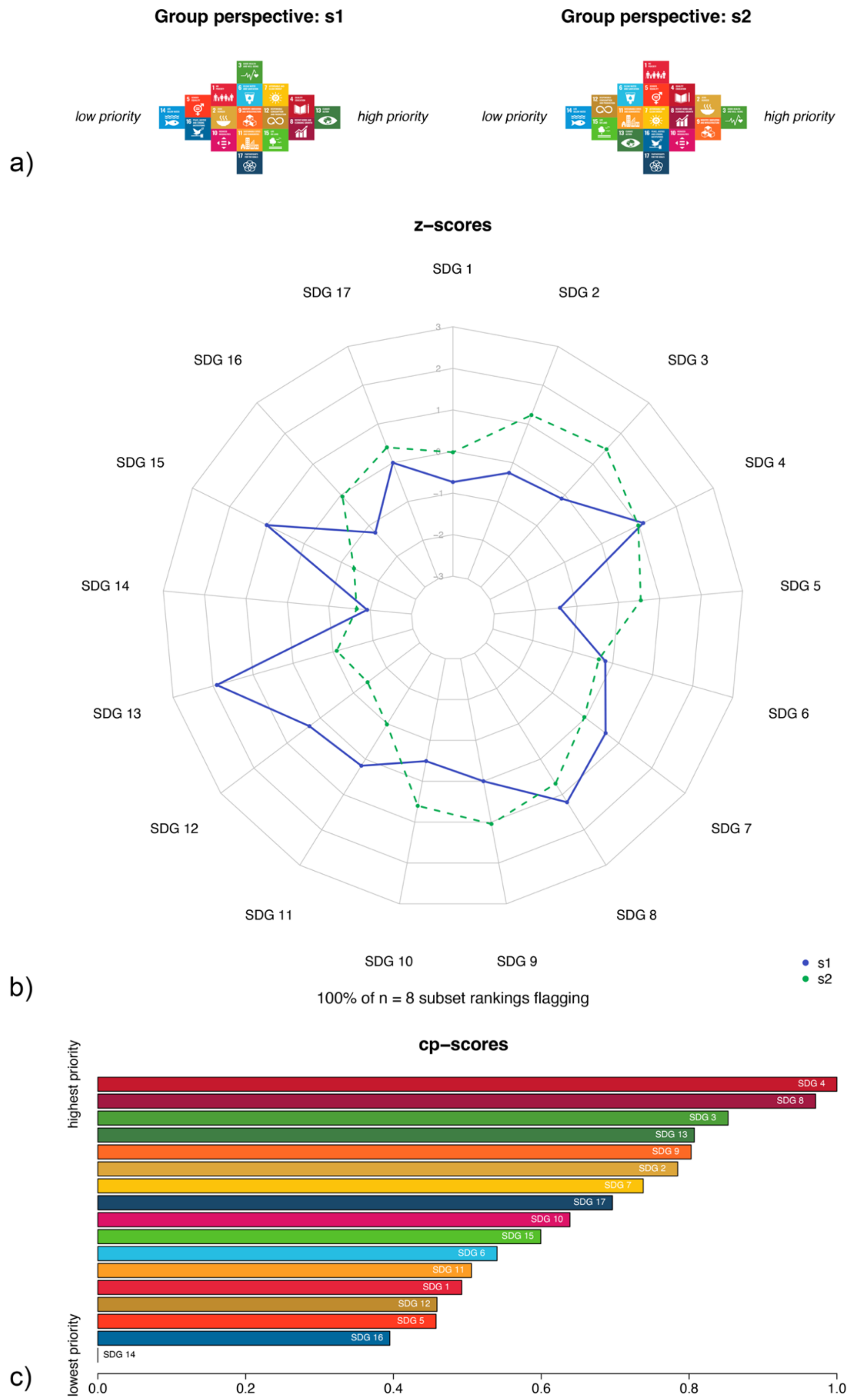

Fig. 6: The results of the Q approach to consensus building for the synthesis analysis (the consolidation of the group perspectives from group A and group B), with a) the 2 synthesis group perspectives, b) a bar plot of the consensus priority scores (cp-scores) of the group, colored in the SDG color each bar represents, and c) a radar chart providing further insights to the different group perspectives, showing the z-scores per SDG and perspective. Factor s1-s2 = group perspectives s1-s2. The line widths within the radar chart represent the factor eigenvalue, i.e. the perspective strength. The line colors and dash types represent each perspective.

For a better understanding and validation of the two synthesis perspectives, we compare them with the input means, having a significant equivalence (p = 1.05332e-11, ***). However, only one of the three top consensus priority SDGs (SDG 4, highest cp-score) also is among the highest five input means. This indicates that in case the project is supposed to focus on more than one priority SDG, there still is some discussion needed. Next, we compare the synthesis results with the bootstrap results. In perspective s1, two SDGs have a position change of one. In perspective s2, six SDGs have a position change of one. Therefore, as there are no major position changes, we consider the synthesis perspectives as stable. Focusing on the consensus priority scores, all top three are among the five highest bootstrap cp-scores. In the lower priority ranked SDGs there are a few position changes compared with the bootstrap cp-scores, but the equivalence test confirms the consensus priority scores and the bootstrap cp-scores are significantly equal (p = 5.60811e-22, ***). Finally looking at the bootstrapped cp-scores, all but SDG 6 (p = 0.06) are individually significant. The full set of consensus priority scores is significant (p = 2.67814e-11, ***). When, due to the relatively small number of input rankings to the synthesis analysis, running the bootstrapping several times, there is a tendency for the consensus priority scores of all SDGs to be individually (and as a set) significant. However, as this is just an example of how the Q approach could be applied, we do not further go into detail of the results here. In real-world analyses, when having such small numbers of input rankings, the bootstrapping should be run with more steps to get robust results.

In decreasing order, to satisfy most of the 40 individual perspectives the best way possible, the project could focus on SGDs 4, 8, and 3, followed by SDGs 13, 9, 2, 7, 17, 10, 15, 6, 11, 1, 12, 5, 16, and 14 (see Fig. 6 and Table S4). This may be a similar or even equal assessment to simply ranking the SDGs by averaging the input ranking values (which we have as a way of validation for the consensus priority scores). However, with the application of the Q approach, we have added value by integrating differently strong group perspectives in the priority gradient. For the discussion among the 40 people, the Q approach therefore can help by making explicit different viewpoints both within the two groups and at the overarching level. At the same time, next to all viewpoint differences, they know they can align their priorities in different ways.

The two synthesis perspectives have statistical consensus in regard of SDGs 4, 6, 7, 8, 14, and 17, two of which are the highest priority scores analyzed (SDGs 4 and 8). Depending on how many SDGs the project could potentially focus on, the highest priority SDGs could be taken, or they can further discuss the role of, for example, SDG 3 and 13. Interestingly, while the consensus priority score of SDG 6 on its own is not significant (as discussed above), the two synthesis perspectives have consensus on its ranking at medium priority.

Next to the content discussion among the 40 people, a network visualization of their individual perspectives flagging (i.e. the people agreeing) to the group perspectives and the synthesis perspectives can point towards individual people or groups of people that have a divergent viewpoint to the ones analyzed (Fig. 7). This is not to point the finger at them. Rather, it can be of great importance for the overall success of the project that their voice is also heard in the discussion and possibly included in the selection of SDGs to focus on. Therefore, the five perspectives not flagging to the group perspectives, and whether they belong to group A or B, are given in table S5. Here, it is important to note that every individual perspective matters.

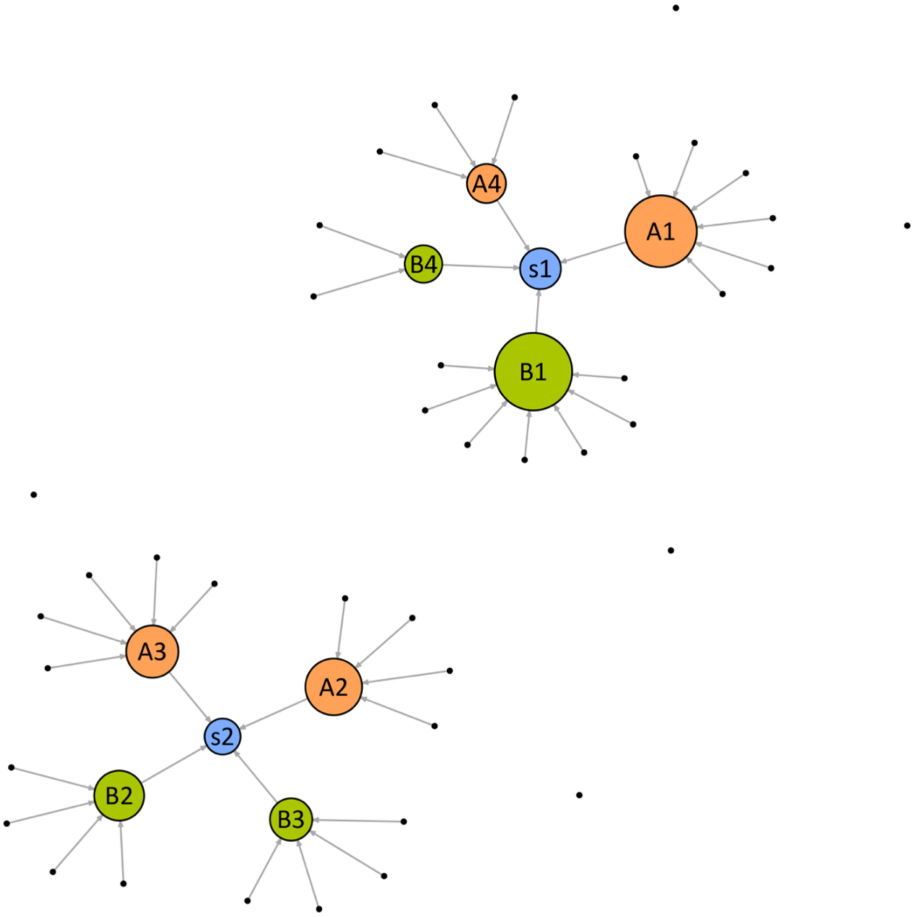

Fig. 7: Factor flagging network illustrating how the individual perspectives (black dots) agree to the different group perspectives; the circle sizes correspond to the strength of the perspectives. “A” (colored orange) and “B” (colored green) indicate to which of the two groups a perspective belongs, “s” (colored blue) represents the synthesis analysis; the adjacent numbers stand for number of the perspective. The circle sizes correspond to the power of the perspectives and are only comparable within the same level of the analysis, i.e. the eight group circles and the two synthesis analysis circles. Solitary black dots represent individual perspectives not flagging, i.e. not agreeing, to a group perspective.

Discussion

Achieving consensus between multiple parties often is a complex and lengthy process (e.g., Bernardo *et al.* 2021). For consensus decision-making, diverse perspectives must be considered, acknowledged, and aligned. However, “there is no single optimal method for making decisions or eliciting views and judgements leading to decisions” (Mukherjee *et al.* 2018, p. 56), nor “[a] single path likely to be universally accepted as superior, and there is no feasible agenda to resolve all conflicts or trade-offs among these pathways [towards sustainability]” (IPBES 2022, p. 26). Individual perspectives, preferences, or priorities depend on personal values and experiences, as well as contextual situations and given framework conditions (Levine, Chan & Satterfield 2015; Chan *et al.* 2016; DeFries & Nagendra 2017). Therefore, a range of methods is applicable for supporting decision-making processes. We chose to work with Q method, as it combines quantitative and qualitative techniques, and therefore enables analyses that go beyond surveys (Kamal, Kocór & Grodzińska-Jurczak 2014) or Likert formats (Cross 2005; Fluckinger & Brodke 2013).

Q method enables a comprehensive and deliberative identification of both diverging perspectives and options upon which to decide, and is particularly applicable in a community setting, i.e. a local, non-governmental, and societal context (Mukherjee *et al.* 2018). Further, and in support of transformative change, IPBES (2022) called for a better integration of diverse values into decision-making through bottom-up governance approaches, knowledge co-production, and participatory and deliberative methods. Q method can play a role here, yet further studies are needed to strengthen the evidence for this (IPBES 2022).

To contribute here, we propose the Q approach to consensus building, expanding Q method by a consensus priority score that can be basis for a constructive dialogue, streamline discussions on the preference or priority of individual goals or statements, and identify options for synergies between parties.

In addition to scientific studies, the Q approach to consensus building has potential to directly support decision-making processes in environmental and sustainability governance, as well as related international agendas in general.

As a recent governance application of the Q approach, we could have imagined the current negotiations on the post-2020 global biodiversity framework, which are facing significant challenges and delays. With 196 member countries in the Convention on Biological Diversity (CBD), diverse perspectives from all over the world must be brought together, including all challenges associated with this. Document CBD/WG2020/3/5 presents an intermediate status of the post-2020 framework draft text, including a composite text integrating text proposals made by the member countries (CBD 2021). Almost impossible to read as continuous text, the composite texts showcase the complexity of bringing the post-2020 framework to a consensus. While past negotiations have shown that consensus building is achievable and agreement upon international frameworks and agendas such as the Aichi Biodiversity Targets, the Paris Agreement, or the Sustainable Development Goals (SDGs) is possible, all those negotiations have in fact been heated. In the post-2020 global biodiversity framework negotiations, the pre-negotiations (i.e. the meetings of the Open-Ended Working Group that happened before the Conference of the Parties) in the end took place in five rounds instead of three as originally planned. Methodological guidance towards consensus building, such as from the Q approach, might have been useful here. Specifically, as an idea for similar processes in the future, different formulations of a target or bracketed text proposals could be ranked, and the negotiation chairs could use the Q approach results to prepare consolidated options for the draft framework texts.

An application of the Q approach therefore is particularly suited for processes in environmental and sustainability governance but not limited to such, as Q method generally works well in situations where conflict is high and an understanding of linkages between perspectives is needed (Mukherjee *et al.* 2018). Q method also gives well interpretable results even with smaller sample sizes (Zabala, Sandbrook & Mukherjee 2018), which allows for its application in smaller negotiation or discussion cases, too. It further enables the uncovering and acknowledgement of hidden and marginalized perspectives (Ockwell 2008; Mazur & Asah 2013), which is critical for transformative governance (Pascual *et al.* 2022), and thus supports inclusiveness in consensus building processes. At the same time, depending on the context the Q approach is applied in (e.g., who is participating, or what kind of statements is assessed), it

is important to consider and account for potential ethical-political implications of the analysis (West & Schill 2022).
In addition to the social network analysis we conducted, a more detailed analysis of the parties sharing a perspective can be done using cluster analysis (Vogel & Lowham 2007; Eppinga, Mijts & Santos 2022), and this could be a future extension to the Q approach. While the ranking of goals or statements might be intellectually challenging and thus needs adequate time (Barry & Proops 1999; Watts & Stenner 2005; Mukherjee *et al.* 2018), we think that the intensive reflection on the goals or statements and their relative ranking can itself contribute to a better understanding of divergent perspectives and support compromises in consensus building. One may argue that a preference or priority cannot be negative (as the ranking distribution is centered around zero). However, in certain world regions certain statements may not need priority or are not relevant at all and therefore can be ranked negatively. For example, in regard of the rankings used for our example, marine biodiversity (SDG 14 *Life below water*) is typically not relevant in mountain regions. Consequently, it is assessed lowest priority for sustainable mountain development (Geschke *et al.* in preparation). To obviate the potential mental barrier of decision-makers that prioritization cannot be negative, we normalized the consensus priority scores to be in a 0 (no priority at all) to 1 (highest priority of all) gradient and harmonized across analyses. The consensus priority score also is a key difference to Eppinga, Mijts and Santos (2022), who use the average participant rankings, i.e. what we call input means, as relative importance of SDGs. In the Q approach, we use the normalized input means in the validation of the consensus priority scores only (see *h*.1).
With the consensus priority scores, our Q approach simplifies the interpretation of the Q method results, which is not straightforward and rather time-consuming (Watts & Stenner 2005; Mukherjee *et al.* 2018), and provides a direct reference supporting decision-making processes in regard of how a consensus-based decision could be reached. A profound interpretation and nuanced consideration of marginalized perspectives that may be identified by the analysis, however, still requires adequate time.

Conclusions

The Q approach to consensus building and its consensus priority score are a proposal for a standardized, transparent, and replicable way to guide the process of decision-making. They are not intended to provide the one and only solution for a consensus but to help having one's own interests reflected while accepting the views of others, and thus support compromising in terms of reaching consensus. The Q approach is constructed to function for different forms and numbers of statements and to allow a broad participation by stakeholders, to support a ministry or convention party develop their negotiation position, or to point towards opportunities for consensus building in multilateral environmental and sustainability governance processes. It is by no means limited to discussions between academics and/or non-governmental organizations, nor to the SDGs. With such flexibility in use, the Q approach is applicable for a wide range of discussion and negotiation settings, and practical applications under real-world conditions are highly encouraged.
With an application of the Q approach to consensus building, stakeholder perspectives can not only be valued and recognized but decision-making can become more inclusive and meaningful through integrating diverse perspectives.

Author contributions

JG developed the methodological idea and workflow, collected, and analyzed the data (see Geschke *et al.* (in preparation) for the full study on SDG priorities in the worlds' mountains), and led the writing of the manuscript. DU and GWP contributed to the workflow development. All authors contributed to the writing of the manuscript.

Data and code accessibility

The data used in the SDG case example is freely accessible in table S2 as supplementary material (supplementary_tables.pdf). Also, the R functions to run the Q approach to consensus building are provided as supplementary (qapproach_functions.R). Table S6 outlines the input needed for and output given by the functions. For the future, JG plans to develop an online tool allowing easy access to the approach and direct application to support governance processes. // The supplementary materials will be made available once the manuscript has been peer-reviewed or upon request.

References

Agrawal, A., Brandhorst, S., Jain, M., Liao, C., Pradhan, N. & Solomon, D. (2022) From environmental governance to governance for sustainability. *One Earth,* **5,** 615-621.

Akhtar-Danesh, N. (2017a) A Comparison between Major Factor Extraction and Factor Rotation Techniques in Q-Methodology. *Open Journal of Applied Sciences,* **07,** 147-156.

Akhtar-Danesh, N. (2017b) An Overview of the Statistical Techniques in Q Methodology: Is There a Better Way of Doing Q Analysis? *Operant Subjectivity,* **38,** 29-36.

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

Armatas, C., Venn, T. & Watson, A. (2016) Understanding social–ecological vulnerability with Q-methodology: a case study of water-based ecosystem services in Wyoming, USA. *Sustainability Science,* **12,** 105-121.

Barry, J. & Proops, J. (1999) Seeking sustainability discourses with Q methodology. *Ecological Economics,* **28,** 337-345.

Bernardo, C., Wang, L., Vasca, F., Hong, Y., Shi, G. & Altafini, C. (2021) Achieving consensus in multilateral international negotiations: The case study of the 2015 Paris Agreement on climate change. *Science Advances,* **7,** eabg8068.

Berry, P.M., Fabók, V., Blicharska, M., Bredin, Y.K., Llorente, M.G., Kovács, E., Geamana, N., Stanciu, A., Termansen, M., Jääskeläinen, T., Haslett, J.R. & Harrison, P.A. (2018) Why conserve biodiversity? A multi-national exploration of stakeholders' views on the arguments for biodiversity conservation. *Biodiversity and Conservation,* **27,** 1741-1762.

Biermann, F., Hickmann, T., Sénit, C.-A., Beisheim, M., Bernstein, S., Chasek, P., Grob, L., Kim, R.E., Kotzé, L.J., Nilsson, M., Ordóñez Llanos, A., Okereke, C., Pradhan, P., Raven, R., Sun, Y., Vijge, M.J., van Vuuren, D. & Wicke, B. (2022) Scientific evidence on the political impact of the Sustainable Development Goals. *Nature Sustainability*.

Bredin, Y.K., Linnell, J.D.C., Silveira, L., Tôrres, N.M., Jácomo, A.A. & Swenson, J.E. (2015) Institutional stakeholders' views on jaguar conservation issues in central Brazil. *Global Ecology and Conservation,* **3,** 814-823.

Bressen, K.T. (2007) Consensus Decision-Making: What, Why, How. *The change handbook: The definitive resource on today's best methods for engaging whole systems,* **495,** 212-217.

Brown, S.R. (1980) *Political subjectivity*. Yale University Press, New Haven and London.

Brown, S.R. (1993) A Primer on Q Methodology. *Operant Subjectivity,* **16,** 91-138.

Carter, G.J. (2018) Multilateral consensus decision making: How Pacific island states build and reach consensus in climate change negotiations. Doctor of Philosophy, Australian National University.

Cattell, R.B. (1966) The Scree Test For The Number Of Factors. *Multivariate Behavioral Research,* **1,** 245-276.

CBD (2021) Report of the Open-Ended Working Group on the Post-2020 Global Biodiversity Framework on its third Meeting (Part I).

Chan, K.M., Balvanera, P., Benessaiah, K., Chapman, M., Díaz, S., Gómez-Baggethun, E., Gould, R., Hannahs, N., Jax, K., Klain, S., Luck, G.W., Martín-López, B., Muraca, B., Norton, B., Ott, K., Pascual, U., Satterfield, T., Tadaki, M., Taggart, J. & Turner, N. (2016) Opinion: Why protect nature? Rethinking values and the environment. *Proceedings of the National Academy of Sciences of the United States of America,* **113,** 1462–1465.

Cross, R.M. (2005) Exploring attitudes: the case for Q methodology. *Health Educ Res,* **20,** 206-213.

Csárdi, G. & Nepusz, T. (2006) The igraph software package for complex network research. *InterJournal, complex systems,* **1695,** 1-9.

D'Amato, D., Droste, N., Winkler, K.J. & Toppinen, A. (2019) Thinking green, circular or bio: Eliciting researchers' perspectives on a sustainable economy with Q method. *Journal of Cleaner Production,* **230,** 460-476.

DeFries, R. & Nagendra, H. (2017) Ecosystem management as a wicked problem. *Science,* **356,** 265.

EAT (2016) Keynote Speech: Prof. Johan Rockström & CEO Pavan Sukhdev. EAT Stockholm Food Forum.

Eppinga, M.B., Mijts, E.N. & Santos, M.J. (2022) Ranking the sustainable development goals: perceived sustainability priorities in small island states. *Sustainability Science*.

Fluckinger, C.D. & Brodke, M.R.H. (2013) Positive Reactions to a Q Sort for Personality Assessment. *Operant Subjectivity,* **36,** 335-341.

Forestier, O. & Kim, R.E. (2020) Cherry-picking the Sustainable Development Goals: Goal prioritization by national governments and implications for global governance. *Sustainable Development,* **28,** 1269-1278.

Geschke, J., Urbach, D., Prescott, G.W. & Fischer, M. (in preparation) Life on land, quality education, and good health and well-being are decisive for sustainable development in mountains.

Hair, J.F., Black, W.C., Babin, B.J. & Anderson, R.E. (2014) *Multivariate Data Analysis*. Pearson Education Limited, Harlow, UK.

Hamadou, I., Moula, N., Siddo, S., Issa, M., Marichatou, H., Leroy, P. & Antoine-Moussiaux, N. (2016) Mapping stakeholder viewpoints in biodiversity management: an application in Niger using Q methodology. *Biodiversity and Conservation,* **25,** 1973-1986.

Hermelingmeier, V. & Nicholas, K.A. (2017) Identifying Five Different Perspectives on the Ecosystem Services Concept Using Q Methodology. *Ecological Economics,* **136,** 255-265.

Hirons, M. (2020) How the Sustainable Development Goals risk undermining efforts to address environmental and social issues in the small-scale mining sector. *Environmental Science & Policy,* **114,** 321-328.

Horcea-Milcu, A.-I. (2022) Values as leverage points for sustainability transformation: two pathways for transformation research. *Current Opinion in Environmental Sustainability,* **57**.

Horn, P. & Grugel, J. (2018) The SDGs in middle-income countries: Setting or serving domestic development agendas? Evidence from Ecuador. *World Development,* **109,** 73-84.
Horton, R. (2020) Offline: COVID-19 is not a pandemic. *The Lancet,* **396**.
IPBES (2015) Preliminary guide regarding diverse conceptualization of multiple values of nature and its benefits, including biodiversity and ecosystem functions and services (deliverable 3 (d)). IPBES Secretariat, Bonn, Germany.
IPBES (2019) Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. (eds S. Díaz, J. Settele, E. Brondízio, H.T. Ngo, M. Guèze, J. Agard, A. Arneth, P. Balvanera, K. Brauman, S. Butchart, K. Chan, L. Garibaldi, K. Ichii, J. Liu, S.M. Subramanian, G. Midgley, P. Miloslavich, Z. Molnár, D. Obura, A. Pfaff, S. Polasky, A. Purvis, J. Razzaque, B. Reyers, R.R. Chowdhury, Y.-J. Shin, I. Visseren-Hamakers, K. Willis & C. Zayas). IPBES Secretariat, Bonn, Germany.
IPBES (2020) IPBES workshop on biodiversity and pandemics: workshop report. (eds P. Daszak, J. Amuasi, C.G. das Neves, D. Hayman, T. Kuiken, B. Roche, C. Zambrana-Torrelio, P. Buss, H. Dundarova, Y. Feferholtz, G. Földvári, E. Igbinosa, S. Junglen, Q. Liu, G. Suzan, M. Uhart, C. Wannous, K. Woolaston, P. Mosig Reidl, K. O'Brien, U. Pascual, P. Stoett, H. Li & H.T. Ngo). IPBES Secretariat, Bonn, Germany.
IPBES (2022) Summary for policymakers of the methodological assessment of the diverse values and valuation of nature of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. (eds U. Pascual, P. Balvanera, M. Christie, B. Baptiste, D. González-Jiménez, C.B. Anderson, S. Athayde, R. Chaplin-Kramer, S. Jacobs, E. Kelemen, R. Kumar, E. Lazos, A. Martin, T.H. Mwampamba, B. Nakangu, P. O'Farrell, C.M. Raymond, S.M. Subramanian, M. Termansen, M.V. Noordwijk & A. Vatn). IPBES Secretariat, Bonn, Germany.
IPCC (2018) Global warming of 1.5°C – An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. IPCC Secretariat, Geneva, Switzerland.
IPCC (2022) Climate Change 2022: Impacts, Adaptation and Vulnerability. IPCC Secretariat, Geneva, Switzerland.
Kamal, S., Kocór, M. & Grodzińska-Jurczak, M. (2014) Quantifying Human Subjectivity Using Q Method: When Quality Meets Quantity. *Qualitative Sociology Review,* **10,** 60-79.
Langston, J.D., McIntyre, R., Falconer, K., Sunderland, T., van Noordwijk, M. & Boedhihartono, A.K. (2019) Discourses mapped by Q-method show governance constraints motivate landscape approaches in Indonesia. *PLOS ONE,* **14,** e0211221.
Le Blanc, D. (2015) Towards Integration at Last? The Sustainable Development Goals as a Network of Targets. *Sustainable Development,* **23,** 176-187.
Levine, J., Chan, K.M.A. & Satterfield, T. (2015) From rational actor to efficient complexity manager: Exorcising the ghost of Homo economicus with a unified synthesis of cognition research. *Ecological Economics,* **114,** 22-32.
Martin, L. (2015) Incorporating values into sustainability decision-making. *Journal of Cleaner Production,* **105,** 146-156.
Mazur, K. & Asah, S. (2013) Clarifying standpoints in the gray wolf recovery conflict: Procuring management and policy forethought. *Biological Conservation,* **167,** 79-89.
Moser, D.J. & Baulcomb, C. (2020) Social perspectives on climate change adaptation, sustainable development, and artificial snow production: A Swiss case study using Q methodology. *Environmental Science & Policy,* **104,** 98-106.

Muff, K., Kapalka, A. & Dyllick, T. (2017) The Gap Frame - Translating the SDGs into relevant national grand challenges for strategic business opportunities. *The International Journal of Management Education,* **15,** 363-383.

Mukherjee, N., Zabala, A., Huge, J., Nyumba, T.O., Adem Esmail, B., Sutherland, W.J. & Everard, M. (2018) Comparison of techniques for eliciting views and judgements in decision-making. *Methods in Ecology and Evolution,* **9,** 54-63.

Ockwell, D.G. (2008) 'Opening up' policy to reflexive appraisal: a role for Q Methodology? A case study of fire management in Cape York, Australia. *Policy Sciences,* **41,** 263-292.

Pascual, U., McElwee, P.D., Diamond, S.E., Ngo, H.T., Bai, X., Cheung, W.W.L., Lim, M., Steiner, N., Agard, J., Donatti, C.I., Duarte, C.M., Leemans, R., Managi, S., Pires, A.P.F., Reyes-García, V., Trisos, C., Scholes, R.J. & Pörtner, H.-O. (2022) Governing for Transformative Change across the Biodiversity–Climate–Society Nexus. *BioScience,* **72,** 684-704.

Payne, D., Spehn, E.M., Prescott, G.W., Geschke, J., Snethlage, M.A. & Fischer, M. (2020) Mountain Biodiversity Is Central to Sustainable Development in Mountains and Beyond. *One Earth,* **3,** 530-533.

Pörtner, H.O., Scholes, R.J., Agard, J., Archer, E., Arneth, A., Bai, X., Barnes, D., Burrows, M., Chan, L., Cheung, W.L., Diamond, S., Donatti, C., Duarte, C., Eisenhauer, N., Foden, W., Gasalla, M.A., Handa, C., Hickler, T., Hoegh-Guldberg, O., Ichii, K., Jacob, U., Insarov, G., Kiessling, W., Leadley, P., Leemans, R., Levin, L., Lim, M., Maharaj, S., Managi, S., Marquet, P.A., McElwee, P., Midgley, G., Oberdorff, T., Obura, D., Osman, E., Pandit, R., Pascual, U., Pires, A.P.F., Popp, A., Reyes-García, V., Sankaran, M., Settele, J., Shin, Y.J., Sintayehu, D.W., Smith, P., Steiner, N., Strassburg, B., Sukumar, R., Trisos, C., Val, A.L., Wu, J., Aldrian, E., Parmesan, C., Pichs-Madruga, R., Roberts, D.C., Rogers, A.D., Díaz, S., Fischer, M., Hashimoto, S., Lavorel, S., Wu, N. & Ngo, H.T.z. (2021) Scientific Outcome of the IPBES-IPCC Co-sponsored Workshop on Biodiversity and Climate Change. IPBES Secretariat, Bonn.

Regan, H.M., Colyvan, M. & Markovchick-Nicholls, L. (2006) A formal model for consensus and negotiation in environmental management. *Journal of Environmental Management,* **80,** 167-176.

Rost, F. (2020) Q-sort methodology: Bridging the divide between qualitative and quantitative. An introduction to an innovative method for psychotherapy research. *Counselling and Psychotherapy Research,* **21,** 98-106.

Rust, N.A. (2016) Can stakeholders agree on how to reduce human–carnivore conflict on Namibian livestock farms? A novel Q-methodology and Delphi exercise. *Oryx,* **51,** 339-346.

Sneegas, G., Beckner, S., Brannstrom, C., Jepson, W., Lee, K. & Seghezzo, L. (2021) Using Q-methodology in environmental sustainability research: A bibliometric analysis and systematic review. *Ecological Economics,* **180**.

Stephenson, W. (1935) Technique of Factor Analysis. *Nature,* **136,** 297.

Still, K.L. & Gordon, J.P. (2009) Consensus building through the lens of q-methodology: Defining profiles for effective models of professional development for literacy practitioners. *College Reading Association Yearbook,* **30,** 215-233.

Tremblay, D., Fortier, F., Boucher, J.F., Riffon, O. & Villeneuve, C. (2020) Sustainable development goal interactions: An analysis based on the five pillars of the 2030 agenda. *Sustainable Development,* **28,** 1584-1596.

UNGA (2015) Resolution adopted by the General Assembly on 25 September 2015.

Vogel, J. & Lowham, E. (2007) Building Consensus for Constructive Action: A Study of Perspectives on Natural Resource Management. *Journal of Forestry*.

Watts, S. & Stenner, P. (2005) Doing Q methodology: theroy, method and interpretation. *Qualitative Research in Psychology,* **2,** 67-91.

Watts, S. & Stenner, P. (2012) *Doing Q Methodological Research: Theory, Method and Interpretation*. London.

West, S. & Schill, C. (2022) Negotiating the ethical-political dimensions of research methods: a key competency in mixed methods, inter- and transdisciplinary, and co-production research. *Humanities and Social Sciences Communications,* **9**.

Winkler, H. & Beaumont, J. (2010) Fair and effective multilateralism in the post-Copenhagen climate negotiations. *Climate Policy,* **10,** 638-654.

Xue, Z., Lin, Z., Wang, H. & McClean, S. (2020) Quantifying consensus of rankings based on q-support patterns. *Information Sciences,* **518,** 396-412.

Zabala, A. (2014) qmethod: A Package to Explore Human Perspectives Using Q Methodology. *The R Journal,* **6,** 163-173.

Zabala, A. & Pascual, U. (2016) Bootstrapping Q Methodology to Improve the Understanding of Human Perspectives. *PLOS ONE,* **11,** e0148087.

Zabala, A., Sandbrook, C. & Mukherjee, N. (2018) When and how to use Q methodology to understand perspectives in conservation research. *Conserv Biol,* **32,** 1185-1194.