Operations Research Perspectives 16 (2026) 100390

Contents lists available at ScienceDirect

Operations Research Perspectives

journal homepage: www.elsevier.com/locate/orp

Simple Group Choice: A minimal-information multi-criteria group decision aiding approach

River Huang

Laboratory for Energy Systems Analysis, PSI Centers for Nuclear Engineering & Sciences and Energy & Environmental Sciences, Forschungsstrasse 111, Villigen, 5232, Switzerland

ARTICLE INFO

Keywords:
Multi-criteria decision aiding
Group decision-making
Binary preferences
Uncertainty
Robust optimization

ABSTRACT

In many real group decision-making problems, decisions must be made with limited data and busy Decision Makers (DMs). In such settings, asking for detailed scores or trade-off weights can be unrealistic and can obscure, rather than clarify, how a recommendation emerges. We propose Simple Group Choice (SGC), a deliberately lightweight approach to multi-criteria group choice that relies only on two sets of binary questions. SGC then aggregates by simple counting: wins on a criterion are weighted by how many people consider that criterion key and are summed across the group. The result is a transparent support score for each alternative and a top choice set, accompanied by compact diagnostics that help interpret and screen. Because incomplete answers are common, we extend SGC to uncertainty by treating missing entries as unknown binary bits and examining all feasible completions. This yields clear, robust recommendations without introducing parametric assumptions. A didactic urban-logistics study illustrates the approach under complete information and under uncertainty. Overall, SGC offers a practical, explainable, and low-burden tool for contexts like shortlisting in early-phase group decisions, and a natural front end to more detailed analyses when needed.

1. Introduction

Complex decisions often involve multiple Decision Makers (DMs) and multiple criteria that must be considered simultaneously. This is common in domains ranging from strategic planning to engineering and public policy, where a group of DMs must evaluate alternatives based on various (and sometimes conflicting) criteria [1]. Multiple Criteria Decision Aiding (MCDA) provides a systematic framework for such problems, helping decision-makers structure the problem and evaluate options against all relevant criteria. The benefits of applying MCDA methods in group settings include improved decision quality, more transparent communication of preferences, and greater commitment to the final decision.

Despite the proliferation of MCDA methods, many classical approaches require extensive information from decision makers, such as utility values or detailed pairwise comparisons [2–4]. Eliciting this level of information can be difficult, time-consuming, and cognitively demanding for participants [5]. In practice, DMs might be reluctant or unable to provide fine-grained preference data, especially in early decision phases or when expertise is limited. This is not just a practical inconvenience: elicitation itself has cognitive and time costs, and more elaborate processes can accumulate biases and reduce the reliability of the elicited inputs [6]. Group Decision-Making (GDM) adds further complexity, as multiple individual preference structures must be reconciled into a collective choice. Indeed, interest in GDM has grown significantly in the last decade due to the need to aggregate inputs from various experts for complex problems [7]. Yet, practitioners may find the plethora of methods daunting or overly complex to apply. There is a demand for simpler, more intuitive approaches that require minimal information from participants while still producing rational outcomes.

Social choice theory, which studies methods for aggregating individual preferences into a collective decision, offers insights that can complement MCDA in group contexts [8]. Voting rules such as majority voting, approval voting, Borda count, etc., demonstrate how preferences can be combined with limited information. Notably, approval voting allows voters to simply approve or reject options, a binary input, and selects the option with most approvals [9]. This idea of using binary preferences (approve or not) aligns with the concept of minimal information input. In fact, earlier research has applied approval-voting concepts to multi-criteria decision-making: multicriteria approval methods were introduced as a way to simplify preference elicitation in MCDA [5]. Such methods require only yes/no judgments and have been shown to satisfy desirable properties while greatly reducing cognitive effort for decision makers. Our work builds on this vein, proposing a new streamlined approach for group MCDA problems.

The Simple Group Choice (SGC) method is proposed, aiming to integrate principles from social choice (in particular, approval-style

E-mail address: river.huang@psi.ch.

https://doi.org/10.1016/j.orp.2026.100390
Received 24 October 2025; Received in revised form 28 February 2026; Accepted 21 April 2026
Available online 23 April 2026

preference elicitation) with MCDA and group decision frameworks. Each DM provides a binary evaluation of each alternative on each criterion. This minimal information structure forgoes detailed scoring or weighting, yet it is sufficient to discern preference patterns. By aggregating these binary preferences, SGC identifies the alternatives that are collectively preferred by the group. The method deliberately emphasizes simplicity and transparency: it can be explained in straightforward terms, and the computation essentially reduces to counting preferences, making it easy for DMs to understand how the outcome is obtained. Crucially, in practical participatory settings, incomplete responses are frequent. DMs may be uncertain about whether a criterion is key, may abstain from identifying a single most-preferred alternative, or may lack sufficient information at early project stages. Classical MCDA methods typically require complete preference structures or impose probabilistic assumptions to handle missing data. In contrast, we aim to preserve the binary, low-burden nature of the elicitation while explicitly accounting for incomplete inputs. This motivates the uncertainty extension developed in this paper, where missing entries are treated as unknown binary variables and analyzed through feasible completion and robustness concepts. This yields robust outputs, such as necessary and possible winners, that connect naturally to established robust-decision-aiding notions [10].

This paper makes three contributions. First, we define SGC, a transparent group-choice rule requiring only minimal binary/categorical elicitation; second, we provide diagnostics for screening and interpreting results; third, we develop a nonparametric robustness extension for incomplete inputs. The proposed techniques are illustrated on an urban-logistics shortlisting problem.

The rest of this paper is organized as follows. In Section 2 we review relevant literature on MCDA methods for choice problems and discuss relevant concepts from social choice and GDM theory that inspire the SGC method. Section 3 introduces the formal framework of the SGC method. Section 4 proposes extensions to the basic SGC: we incorporate uncertainty modeling to handle ambiguous or probabilistic judgments, introduce effort-aware promotion of alternatives, and outline a robust optimization variant of SGC that secures the decision against worst-case deviations in preferences. In Section 5, a didactic case study is presented to illustrate the method, and we discuss the results. Finally, Section 7 concludes the paper, highlighting the contributions and advantages of SGC as a practical group decision tool, and suggesting avenues for future research.

2. Literature review

This paper addresses a specific and common setting in applied MCDA: a *group choice* problem in which a panel must shortlist one or a few alternatives under limited time, limited data, and heterogeneous viewpoints. This problem is particularly present in the early stages of decision-making. The central methodological challenge lies in balancing expressive models, which require rich and calibrated preference information, with lightweight procedures that reduce the burden of elicitation and remain explainable to participants. We therefore focus this review on three closely related strands: choice-oriented MCDA with limited compensation, minimal-information preference elicitation, and aggregation under group and uncertainty.

Classical families of MCDA — value/utility models (e.g., Multi-Attribute Utility Theory (MAUT)), outranking approaches (e.g., ELECTRE, PROMETHEE), and pairwise-comparison methods (e.g., Analytic Hierarchy Process (AHP)) — are well established and widely applied for multi-criteria comparison [4,11–14]. However, in the *choice problematic* the aim is not necessarily to produce a complete ranking, but to justify retaining a small subset of broadly satisfactory options, while allowing incomparabilities when evidence is insufficient or stakeholder views diverge [15,16]. In practice, many scoring and ranking models rely on compensatory trade-offs and require either cardinal value functions, weights, or detailed pairwise intensities [17]. Outranking reduces full compensation, but typically introduces additional calibration parameters (e.g., indifference/preference/veto thresholds) whose elicitation and justification can be demanding in participatory settings [4,12]. These requirements are amplified in group contexts because individual inputs must be made comparable and then aggregated, which can add negotiation overhead and additional assumptions [1].

To reduce burden, a stream of work studies MCDA procedures based on ordinal, qualitative, or coarse judgments rather than precise cardinal scores [18]. A practically attractive case is to constrain inputs to binary judgments, such as acceptability or "pass/fail" criteria, which aligns with screening logic and satisficing behavior, especially in the early decision-making phase. In MCDA, binary inputs have been used both as a preliminary filter and as a core representational choice to avoid spurious precision when respondents cannot reliably provide fine-grained ratings [19,20]. The most direct connection to the present paper is Multicriteria Approval, which adapts approval-voting ideas to a multi-criteria setting by combining dichotomous per-criterion judgments with criterion-importance information [5]. Its main contribution is to demonstrate that meaningful discrimination among alternatives can be obtained from sparse inputs, while reducing cognitive effort relative to full scoring. However, this method focuses on the single DM setting and does not consider the complexities of group decision-making.

Group MCDA introduces two more design questions: how to aggregate across decision makers at the criterion level, and how to translate the aggregated information into an alternative-level recommendation [21]. Common approaches include:

1. aggregation operators that fuse individual evaluations before applying an MCDA aggregator [22],
2. Consensus Reaching Processs (CRPs) that iteratively seek agreement through feedback and revision [23–25],
3. weighting and fusion schemes that assign differential influence to DMs [7],
4. stochastic analysis to find the acceptability of ranking among DMs [26,27]

In parallel, social choice theory provides well-understood aggregation rules for individual preferences over alternatives (e.g., majority, Borda, approval) and clarifies the trade-offs among fairness and consistency properties [28–30]. Approval voting is particularly relevant because it relies on binary input and tends to select broadly acceptable compromise options [9]. Nonetheless, importing social-choice rules into MCDA is not immediate: criteria create an additional structure, and any method must specify how criterion-level judgments and importance views are represented and combined before an alternative-level group choice is formed [1].

Minimal-information designs are more robust to respondent fatigue and easier to administer, which can increase the feasible size of the DM pool and facilitate broader participation in the decision process. The trade-off is practical: missing, incomplete, or ambiguous responses may become more frequent. DMs may be uncertain about particular judgments. For example, they may be unsure which criterion is more important or unable to assess an alternative's performance on a given criterion. Because minimal-information designs rely on a smaller set of inputs, such uncertainty can have a proportionally larger impact on the final recommendations. A substantial literature handles uncertainty in MCDA via intervals, feasible weight sets, or stochastic acceptability analysis, yielding rankings or acceptability indices under incomplete information [31,32]. Related robust preference–disaggregation approaches derive recommendations that hold across sets of compatible models rather than a single calibrated parameterization [10]. These frameworks provide valuable concepts (acceptability, necessary/possible outcomes), but they are often developed for cardinal performance tables and weight uncertainty, and they typically require either probabilistic assumptions or a more elaborate preference model than binary judgments.

The above strands suggest a clear gap for minimal-information elicitation. First, Multicriteria Approval is fundamentally anchored in a single DM's criterion-importance structure and does not natively represent how many DMs regard a criterion as key, nor how to aggregate heterogeneous "importance" views in a transparent way [5]. Second, uncertain outcomes are possible which can force additional elicitation without a built-in mechanism to quantify how much information is missing or where clarifications matter most. Third, existing uncertainty-handling frameworks are not tailored to the specific structure of minimal-information, and they do not directly deliver a set-valued group choice based on feasible completions of missing binary entries. These gaps motivate Simple Group Choice (SGC): a group-oriented, minimal-information procedure, which we introduce next.

3. Simple group choice (SGC) method

Classical MCDA families are powerful but often not well-suited to the choice problematic in real group settings. In early-phase, multi-stakeholder decisions, elicitation burden, aggregation across heterogeneous DMs, and the need for transparent justification frequently outweigh the benefits of fine-grained models. We therefore propose SGC, a lightweight, preference-based method that operates on minimal binary inputs at the DM, criterion, alternative level integrating principles from social choice. SGC aggregates these primitives by counting to form clear alternative-level support scores and selects the top options without forcing a full ranking. The design emphasizes transparency, group consistency under unanimous preferences, and ease of deployment; it also admits practical extensions for uncertain inputs, effort-aware improvement of alternatives, and robust selection. Below we introduce the binary structures and aggregation rule that define SGC, before detailing these extensions.

3.1. Problem framework and binary preference structure

Consider a decision problem with a set of m alternatives $\mathcal{A} = \{a_1, \dots, a_i, \dots, a_m\}$ and a set of n evaluation criteria $\mathcal{G} = \{g_1, \dots, g_j, \dots, g_n\}$. We assume there is a group of T DMs $\mathcal{S} = \{s_1, \dots, s_t, \dots, s_T\}$ who are involved in evaluating the alternatives. The goal is to select one alternative (or a subset of alternatives of a given size k) that best fits the group's preferences across all criteria.

For each DM $t \in \{1, \dots, T\}$ and criterion $g_i \in \mathcal{G}$, we record a binary *key-criterion indicator* $r_i^t \in \{0,1\}$ indicating whether g_i is considered key (non-trivial) by t. A convenient mathematical semantics is to posit a latent, nonnegative salience $w_i^t \geq 0$ and a purely notional cut-off $\varepsilon_t > 0$ that separates trivial from non-trivial salience, and define:

$$r_i^t = \begin{cases} \mathbf{1}, & \text{if } w_i^t \geq \varepsilon_t \\ 0, & \end{cases} \qquad i = 1, \dots, n. \tag{1}$$

Here ε_t is not an estimand nor a calibrated parameter; it only formalizes the idea that $r_i^t = 1$ means "key" and $r_i^t = 0$ means "trivial". In vector form $r^t = (r_1^t, \dots, r_n^t)^\top$ and the group matrix $R = [r^1 \ \cdots \ r^T] \in \{0,1\}^{n\times T}$ induce the *criterion support vector*

$$u := R\,\mathbf{1}_T \in \{0, 1, \dots, T\}^n,$$

where u_i is the number of DMs who marked criterion g_i as key.

Optional constraints. To avoid degenerate declarations one may impose $1 \leq \sum_i r_i^t \leq \kappa$ for each t. Equivalently, in a top-κ variant, let ε_t be the κth order statistic of $(w_1^t, \dots, w_n^t)$ so that exactly κ criteria are tagged key. If one prefers to avoid latent variables altogether, the same semantics is captured by treating ε_t as a symbolic boundary ("trivial vs. non-trivial") and eliciting r_i^t directly.

Having collected the key-criterion indicators $R = [r^1 \ \cdots \ r^T] \in \{0,1\}^{n\times T}$, we next elicit, for each DM $t \in \{1, \dots, T\}$ and each criterion $g_j \in \mathcal{G}$, a single most-preferred alternative. Let $\succ_{g_j}^t$ denote DM t's strict preference on $\mathcal{A}$ induced by g_j, and let $i_j^t \in \{1, \dots, m\}$ be a unique top element such that

$$a_{i_j^t} \succ_{g_j}^t a_i, \qquad \forall\, i \in \{1, \dots, m\} \setminus \{i_j^t\}.$$

Because SGC requires only a binary winner/non-winner outcome per (g_j, t), we assume subjective qualitative judgments from the DM are sufficient; if numerical data exist, the same identification can be made by any criterion-specific argmax rule. We encode DM t's selections in a binary *criterion-winner matrix*:

$$P^{(t)} = \left[p_{ij}^t\right]_{i=1,\dots,m}^{j=1,\dots,n} \in \{0,1\}^{m\times n}, \qquad p_{ij}^t = \begin{cases} 1, & i = i_j^t, \\ 0, & \text{otherwise}, \end{cases} \quad \Rightarrow \quad \sum_{i=1}^m p_{ij}^t = 1 \quad \forall j.$$

Stacking $\{P^t\}_{t=1}^T$ yields a three-dimensional *most-preferred-alternative tensor*

$$\mathcal{P} = [P^1 \mid \cdots \mid P^T] \in \{0,1\}^{m\times n\times T},$$

whose (i, j, t) entry records whether a_i is DM t's most-preferred alternative on criterion g_j.

3.2. Choosing the preferred alternatives

Given the key-criterion indicators $R = [r^1 \ \cdots \ r^T] \in \{0,1\}^{n\times T}$ and, for each DM t, the criterion-winner matrix $P^{(t)} \in \{0,1\}^{m\times n}$, SGC aggregates in two transparent steps.

- **Level 1 (cross-view weighting of DM t's winners).** Fix $t \in \{1, \dots, T\}$. Define:

$$M^{(t)} := P^{(t)} R \in \mathbb{N}^{m\times T}, \qquad M_{i t'}^{(t)} = \sum_{j=1}^n p_{ij}^t r_j^{t'} \quad (t' = 1, \dots, T). \tag{2}$$

  Thus $M_{i t'}^{(t)}$ counts, for alternative a_i, how many criteria that DM t' marked as key coincide with the criteria on which DM t declares a_i most preferred. **Interpretation:** "Looking at DM t's 'wins' for a_i, how often do those wins lie on criteria that DM t' actually cares about?"

  It is convenient to summarize keyness by the criterion support vector $u := R\,\mathbf{1}_T$, and then obtain the per-t alternative score:

$$f^{(t)} := P^{(t)} u \in \mathbb{N}^m, \qquad f_i^{(t)} = \sum_{j=1}^n p_{ij}^t u_j.$$

  Interpretation: each time DM t says "a_i is best on g_j", that contribution is weighted by how many DMs consider g_j key (e.g., a win on a criterion marked key by two of three DMs contributes 2).
- **Level 2 (across-DM pooling).** The aggregation across decision makers produces a vector of collective scores:

$$F := \sum_{t=1}^T f^{(t)} = \sum_{t=1}^T P^{(t)} u \in \mathbb{N}^m, \qquad F_i = \sum_{t=1}^T \sum_{j=1}^n p_{ij}^t u_j \quad (i = 1, \dots, m). \tag{3}$$

  Interpretation. F_i counts how often alternative a_i is declared most-preferred on criteria that many DM label as key. Thus, an alternative scores highly if its "wins" occur on widely supported criteria—rewarding breadth of support on what the group deems important.

  We can further define two $[0,1]$-scaled versions of F:

  1. *Keyness-based normalization.* Scale by the best score compatible with the elicited keyness pattern u:

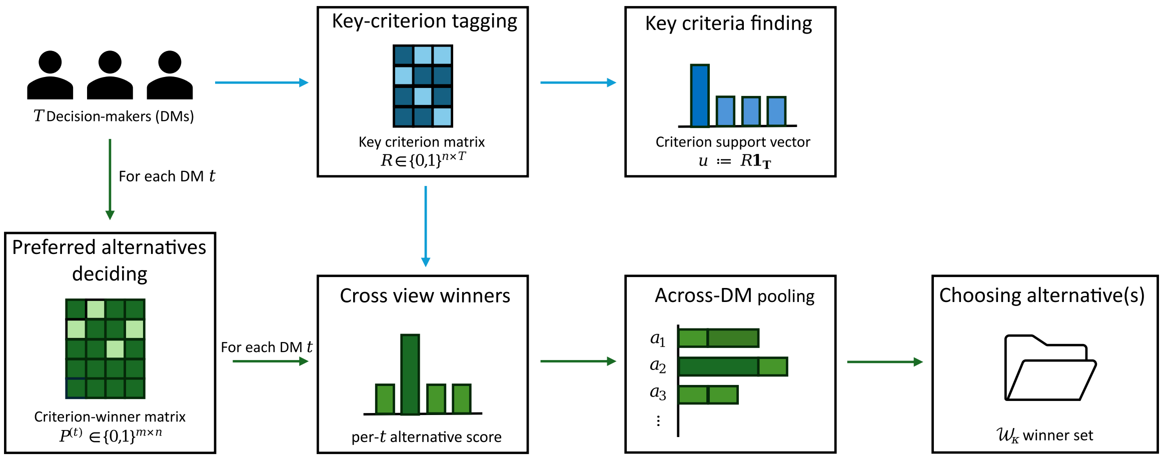


Fig. 1. SGC framework. First, for each DM s_t, elicit binary keyness over criteria, forming $R \in \{0,1\}^{n\times T}$, and per-criterion most-preferred alternatives, forming $P^{(t)} \in \{0,1\}^{m\times n}$. Then, aggregate keyness into the criterion support vector $u = R\mathbf{1}_T$. Next, accumulate alternative scores $F = \sum_{t=1}^{T} P^{(t)}u$. Finally, select the winner set $\mathcal{W}_\kappa$.

$$F_i^{\text{key}} := \begin{cases} \dfrac{F_i}{T\,\sum_{j=1}^{n} u_j}, & \sum_{j=1}^{n} u_j > 0, \\ 0, & \sum_{j=1}^{n} u_j = 0. \end{cases}$$

Interpretation: fraction of the maximum one could achieve if every DM named the same alternative on every criterion, given the current keyness declarations.

2. *Ideal normalization.* Scale by an absolute upper bound attained when all criteria are key for all DMs and everyone names the same alternative:

$$F_i^{\text{ideal}} := \frac{F_i}{nT^2}, \qquad i = 1, \dots, m.$$

Interpretation: fraction of the theoretical maximum in a fully key, fully aligned scenario. It can be useful for cross-study comparability.

This aggregation enforces cross-DM corroboration: an alternative gains score only when a DM's "win" on a criterion coincides with that criterion being marked key by (possibly many) others. Because $u := R\mathbf{1}_T$ is common to all DMs and each $P^{(t)}$ is column-stochastic ($\sum_i p_{ij}^t = 1$), a unilateral change in some r^s simply adds 1 to u_j and thereby adds the same increment to every term $P^{(t)}e_j$ (the winner on criterion j for each DM t). Hence no single DM can inflate their relative influence by marking many criteria as key: such changes shift weight at the criterion level uniformly across DMs' winner declarations and do not privilege any particular $P^{(t)}$.

Given the collective score vector $F \in \mathbb{N}^m$ with entries F_i for $a_i \in \mathcal{A}$, the winner set of cardinality $\kappa \in \{1,\dots,m\}$ is defined as the indices of the κ largest scores. We can find the κ preferred alternatives in the permutation form.

Alternatively, we can find the alternatives in an optimization form. Let $y \in \{0,1\}^m$ be the selection indicator vector ($y_i = 1$ iff $i \in \mathcal{W}_\kappa$). Then $\mathcal{W}_\kappa$ is obtained by solving the 0–1 program:

$$\max_{y\in\{0,1\}^m} \quad F^\top y \quad \text{s.t.} \quad \mathbf{1}^\top y = \kappa, \tag{4}$$

whose optimal solutions correspond to choosing the κ largest components of F. Computationally, sorting F yields $\mathcal{W}_\kappa$ in $O(m\log m)$ time. And in the most conventional choice problematic form, i.e., for $\kappa = 1$, we have:

$$i^\star \in \arg\max_{i\in\{1,\dots,m\}} F_i.$$

Proof. From the Eq. (3), the contribution of column j to F_i is $\sum_t p_{ij}^t u_j = 0$ when $u_j = 0$, hence F does not depend on $\{p_{ij}^t\}_{i,t}$ for that j. □

Fig. 1 summarizes the SGC workflow from binary elicitation to top-κ selection. A step-by-step algorithmic description of SGC is provided in Algorithm 1 for clarity.

Algorithm 1 Simple Group Choice (SGC)

Input: $\mathcal{A} = \{a_1,\dots,a_m\}$, $\mathcal{G} = \{g_1,\dots,g_n\}$, $\mathcal{S} = \{s_1,\dots,s_T\}$, κ
Input: $R \in \{0,1\}^{n\times T}$ (key-criterion indicators); for each $t = 1,\dots,T$, $P^{(t)} \in \{0,1\}^{m\times n}$ with $\sum_{i=1}^{m} p_{ij}^t = 1$ for all j
Output: Collective scores $F \in \mathbb{N}^m$; winner set $\mathcal{W} \subseteq \{1,\dots,m\}$ with $|\mathcal{W}| = \kappa$
$u \leftarrow R\mathbf{1}_T$
$F \leftarrow \mathbf{0}_m$
for $t = 1$ **to** T **do**
  $F \leftarrow F + P^{(t)}u$
end for
Define decision variables: $y \in \{0,1\}^m$
Optimization model:
  maximize $\sum_{i=1}^{m} F_i y_i$
  subject to $\sum_{i=1}^{m} y_i = \kappa$
Solve the binary program and obtain $y^\star$
$\mathcal{W} \leftarrow \{\, i \in \{1,\dots,m\} \;:\; y_i^\star = 1 \,\}$
return $(F, \mathcal{W})$

3.2.1. *Special cases under the choice rule and the solutions*

Two situations can arise naturally under SGC's choice rule.

1. Ties at the cutoff (yielding more than κ winners): Given $F \in \mathbb{N}^m$, let τ_κ denote the κth order statistic of $\{F_1,\dots,F_m\}$ (the κth largest value). SGC adopts a variable-size, tie-inclusive policy:

$$\mathcal{W}_\kappa^{\text{all}}(F) := \{\, i \in \{1,\dots,m\} : F_i \geq \tau_\kappa \,\}.$$

By construction $|\mathcal{W}_\kappa^{\text{all}}(F)| \geq \kappa$, with strict inequality only when a tie occurs at the cutoff. In practice, when the number of criteria and decision makers is moderate to large and preferences are heterogeneous, exact ties at τ_κ are uncommon; when they do occur, including all tied options preserves neutrality and avoids arbitrary tie-breaking (though a group may subsequently agree on a deterministic tie-break if desired).

2. Too few supported alternatives (yielding fewer than κ winners). Define the set of key criteria $J_{\text{key}} := \{ j : u_j > 0 \}$, i.e., criteria marked key by at least one DM, and the set of alternatives that ever appear as most-preferred on some key criterion (for at least one DM):

$$I_{\text{key}} := \left\{ i : \sum_{t=1}^{T} \sum_{j \in J_{\text{key}}} p_{ij}^{t} > 0 \right\} = \{ i : F_i > 0 \}.$$

If $|I_{\text{key}}| = q < \kappa$, then no size-$\kappa$ winner set can be justified without including alternatives that are never most-preferred on any key criterion for any DM (those with $F_i = 0$). In this case SGC returns only the alternatives actually supported by the elicited preferences, namely the co-maximizers $\mathcal{W}^{\max}(F) = \arg\max_i F_i \subseteq I_{\text{key}}$, hence $|\mathcal{W}^{\max}(F)| \le q < \kappa$. This is not a deficiency but an informative outcome: the group's key-criterion preferences are concentrated on fewer than κ alternatives. Formally, if for some index set $J \subseteq J_{\text{key}}$ and all DMs t there exists $i^{\star}$ with

$$a_{i^{\star}} \succ_{g_j}^{t} a_i \quad \text{for all } j \in J \text{ and all } i \neq i^{\star} \quad \left(\text{equivalently } p_{i^{\star}j}^{t} = 1,\ p_{ij}^{t} = 0 \text{ for } i \neq i^{\star}\right),$$

then $F_{i^{\star}} = \sum_{t=1}^{T} \sum_{j \in J} u_j$ and $F_i = 0$ for $i \neq i^{\star}$, so $\mathcal{W}^{\max}(F) = \{i^{\star}\}$ for any $\kappa \ge 1$. In words, all most-preferred judgments on key criteria point to the same alternative (or to a small set of co-maximizers), and SGC appropriately returns this concentrated set rather than "filling up" alternatives to κ.

3.3. Diagnostic metrics for SGC (screening, agreement, and stability)

SGC's binary inputs enable simple, informative diagnostics at the criterion, alternative, and group levels. Metrics can be proposed to help screen "key" criteria for downstream MCDA, interpret breadth and strength of support for alternatives.

Let $\mathcal{P} \in \{0,1\}^{m \times n \times T}$ be the most-preferred-alternative tensor with entries $\mathcal{P}_{ijt} = 1$ iff DM t selects a_i as most preferred on g_j (otherwise 0). The keyness counts are $u = R\mathbf{1}_T \in \{0, \dots, T\}^n$ with $u_j = \sum_{t=1}^{T} r_j^t$. Define the winner counts per criterion

$$c_{ij} := \sum_{t=1}^{T} \mathcal{P}_{ijt} \in \{0, \dots, T\}, \qquad c_{ij}^{\text{key}} := \sum_{t:\, r_j^t = 1} \mathcal{P}_{ijt} \in \{0, \dots, u_j\},$$

i.e., c_{ij} counts all DMs picking a_i on g_j, while c_{ij}^{key} counts only those who marked g_j as key. We have the following diagnostics:

1. **Keyness rate:**

$$\rho_j := \frac{u_j}{T} \in [0,1]. \tag{5}$$

This is the share of DMs who mark criterion g_j as key. For preliminary screening, retain $\{ g_j : \rho_j \ge \tau_\rho \}$ for a chosen threshold $\tau_\rho \in (0,1)$ (e.g., $\tau_\rho = 0.5$).

2. **Key-agreement on** g_j:

$$d_j := \begin{cases} \dfrac{\max_i c_{ij}^{\text{key}}}{u_j}, & u_j > 0, \\ 0, & u_j = 0, \end{cases} \qquad d_j \in [0,1]. \tag{6}$$

This is the fraction of key-declaring DMs who agree on the modal most-preferred alternative for g_j; larger values indicate that g_j is more *decisive*.

3. **Unweighted breadth (coverage) of an alternative:**

$$B_i := \left| \{ j : c_{ij} > 0 \} \right| \tag{7}$$

counts how many criteria a_i wins with at least one DM. A key-only variant is $B_i^{\text{key}} := |\{ j : c_{ij}^{\text{key}} > 0 \}|$.

4. **Overall key-agreement.**

$$\bar{d} := \frac{1}{|J_{\text{key}}|} \sum_{j:\, u_j > 0} d_j, \qquad J_{\text{key}} := \{ j : u_j > 0 \}, \tag{8}$$

summarizes, on average, how decisive the key criteria are. (Optionally, a weighted variant $\bar{d}_w := \frac{\sum_j u_j d_j}{\sum_j u_j}$ emphasizes criteria marked key by more DMs.)

5. **Dispersion of keyness across criteria.** Let $\tilde{u}_j := u_j / \sum_{\ell} u_{\ell}$ over j with $u_j > 0$. The entropy

$$H_u := - \sum_{j:\, u_j > 0} \tilde{u}_j \log \tilde{u}_j \tag{9}$$

measures how concentrated (H_u small) or diffuse (H_u large) keyness is across criteria. A convenient interpretability aid is the "effective number" $N_{\text{eff}} := \exp(H_u)$.

4. Extensions: Uncertainty, effort-aware promotion, and robustness

Thanks to SGC's simplicity, elicitation can be implemented as a questionnaire: for each criterion, DMs indicate whether it is key (yes/no) and select a single most-preferred alternative. In practice, incomplete answers are common: a DM may decline to mark a criterion as key, or may be unsure which alternative is most preferred on a criterion ("Not sure"). We extend SGC to handle such cases by treating each missing entry as an *unknown binary bit*, denoted x, which must ultimately resolve to 0 or 1 but is currently unspecified. This induces a set-valued view of the outcome: every feasible completion of the unknown bits yields a deterministic SGC score vector and winner set. The analyst can then report which alternatives are necessarily selected (appear in the winner set for all completions), which are *possibly* selected (appear for at least one completion), and how large the unresolved space remains. When uncertainty is excessive, one can target clarifications, remove very uncertain inputs, or include additional DMs, all while preserving the binary nature of SGC.

Definition 1 (*Feasible Completion and Induced Outcome*). Let $R \in \{0,1,x\}^{n \times T}$ and $P^{(t)} \in \{0,1,x\}^{m \times n}$ for $t = 1, \dots, T$, with "x" an unknown binary. A *feasible completion* is a pair $\omega = (\hat{R}, \{\hat{P}^{(t)}\}_{t=1}^{T})$ obtained by replacing every x with a value in $\{0,1\}$ such that $\sum_{i=1}^{m} \hat{p}_{ij}^{t} = 1$ for all (j,t). The induced SGC objects are

$$u^{(\omega)} = \hat{R}\,\mathbf{1}_T, \qquad F^{(\omega)} = \sum_{t=1}^{T} \hat{P}^{(t)} u^{(\omega)}, \qquad \mathcal{W}_{\kappa}^{(\omega)} = \mathrm{Top}_{\kappa}\big(F^{(\omega)}\big).$$

Let Ω denote the (finite) set of all feasible completions.

Thus, the SGC outcome under missing bits is described by the set of score vectors $\{F^{(\omega)}\}_{\omega \in \Omega}$ and the family of winner sets $\{\mathcal{W}_{\kappa}^{(\omega)}\}_{\omega \in \Omega}$.

Exact enumeration is feasible when the number of unknown bits is small. Otherwise, alternative-wise score bounds are useful:

$$F_i^{\min} := \min_{\omega \in \Omega} F_i^{(\omega)}, \qquad F_i^{\max} := \max_{\omega \in \Omega} F_i^{(\omega)}.$$

These are obtained by binary linear optimization. Introduce binary variables for each missing keyness r_j^t and for each missing winner indicator p_{ij}^t with the one-hot constraints, and linearize the product

$$F_i = \sum_{j=1}^{n} \sum_{t=1}^{T} \sum_{t'=1}^{T} p_{ij}^{t} r_j^{t'} \quad \text{via} \quad z_{ijtt'} \in \{0,1\},\ z_{ijtt'} \le p_{ij}^{t},\ z_{ijtt'} \le r_j^{t'},\ z_{ijtt'} \ge p_{ij}^{t} + r_j^{t'} - 1,$$

so that $F_i = \sum_{j,t,t'} z_{ijtt'}$. Then solve one Mixed Integer Linear Programming (MILP) to minimize F_i and one to maximize F_i (for each i or all i at once). These bounds give immediate, robust information without probabilities.

4.1. Selection indices under missing bits

From Ω (exact or sampled) we define inclusion indices and ambiguity measures.

1. *Inclusion indices (set-based).* For each alternative a_i,

$$\nu_i := \min_{\omega\in\Omega} \mathbf{1}\{\, i \in \mathcal{W}_\kappa^{(\omega)} \,\} \in \{0,1\}, \qquad \pi_i := \max_{\omega\in\Omega} \mathbf{1}\{\, i \in \mathcal{W}_\kappa^{(\omega)} \,\} \in \{0,1\}.$$

Here, $\nu_i = 1$ means a_i is selected for every admissible completion (necessary winner); $\pi_i = 1$ means it is selected for at least one completion (possible winner). When full enumeration is large, a descriptive frequency from uniformly sampled completions $\{\omega_s\}_{s=1}^{N_\Omega}$ is

$$\hat{\Pi}_i := \frac{1}{N_\Omega} \sum_{s=1}^{N_\Omega} \mathbf{1}\{\, i \in \mathcal{W}_\kappa^{(\omega_s)} \,\} \in [0,1].$$

2. *Score and rank intervals.* The intervals $[F_i^{\min}, F_i^{\max}]$ and $[\mathrm{rank}_i^{\min}, \mathrm{rank}_i^{\max}]$ summarize how much each alternative's standing can vary across completions. If $F_i^{\min}$ exceeds the κth largest possible rival score, a_i is necessarily selected; if $F_i^{\max}$ is below that threshold, a_i is necessarily excluded.
3. *Agreement index.* Let $\mathcal{U}_\kappa := \bigcup_{\omega\in\Omega} \mathcal{W}_\kappa^{(\omega)}$ and $\mathcal{I}_\kappa := \bigcap_{\omega\in\Omega} \mathcal{W}_\kappa^{(\omega)}$. The agreement index is

$$\mathrm{Agr}_\kappa := \frac{|\mathcal{I}_\kappa|}{|\mathcal{U}_\kappa|} \in [0,1],$$

so that $\mathrm{Agr}_\kappa = 1$ iff all completions yield the same winner set (full agreement).

4.2. Robust SGC under missing bits

A conservative decision selects only necessary winners,

$$\mathcal{W}_\kappa^{\mathrm{nec}} := \bigcap_{\omega\in\Omega} \mathcal{W}_\kappa^{(\omega)},$$

possibly with $|\mathcal{W}_\kappa^{\mathrm{nec}}| < \kappa$ if the evidence concentrates on fewer alternatives. An optimistic view reports *possible* winners,

$$\mathcal{W}_\kappa^{\mathrm{pos}} := \bigcup_{\omega\in\Omega} \mathcal{W}_\kappa^{(\omega)}.$$

Between these, a robust bounds selection uses score minima,

$$\mathcal{W}_\kappa^{\mathrm{LCB}} := \mathrm{Top}_\kappa(F^{\min}), \qquad F_i^{\min} := \min_{\omega\in\Omega} F_i^{(\omega)},$$

guaranteeing that every chosen alternative attains at least the shown score regardless of how unknown bits are resolved.

5. Case study: Group decision in European urban logistics

To illustrate the SGC in action, we present a didactic case study based on a real-world inspired urban logistics decision problem. The context is a medium-to-large European city that is updating its urban freight strategy in response to increasing last-mile delivery activity, persistent congestion in dense districts, limited curb-space availability, and tightening environmental and livability objectives. In practice, such cities must often balance competing policy targets: reducing emissions and traffic externalities while maintaining reliable service levels for retailers, residents, and public services [33]. Moreover, freight measures are typically implemented under uncertainty about behavioral responses (e.g., carriers' willingness to cooperate) and with non-trivial distributional effects across neighborhoods and stakeholder group [34].

The decision is coordinated by a working group convened by the municipality and composed of representatives from logistics operators, local businesses, residents' associations, and relevant public authorities. This setting naturally leads to heterogeneous priorities and different interpretations of what "good performance" means on each dimension. In addition, the group faces practical implementation constraints: budget, political feasibility, and organizational capacity make it unrealistic to deploy many measures simultaneously. Consequently, the city seeks to select a small portfolio of interventions to pilot over a fixed planning horizon, with the option of scaling up later. We therefore model the decision as the choice of $\kappa = 2$ strategies out of five candidates, allowing the selected pair to combine complementary effects (e.g., a structural measure on consolidation with an operational measure on last-mile operations).

Table 1
List of alternatives $\mathcal{A} = \{a_1, \ldots, a_5\}$ and criteria $\mathcal{G} = \{g_1, \ldots, g_8\}$.

Code	Description
a_1	Peripheral Distribution Center (PDC) at city outskirts
a_2	Electric cargo bikes for last-mile delivery
a_3	Night-time delivery windows
a_4	Urban consolidation hub in the city
a_5	Crowdshipping platform
g_1	Cost
g_2	Operational feasibility
g_3	Traffic congestion impact
g_4	Emissions reduction
g_5	Service level
g_6	Stakeholder acceptance
g_7	Scalability
g_8	Policy fit

This case study is particularly suitable for the proposed SGC because the elicitation context is characterized by limited time, heterogeneous expertise, and bounded willingness to provide detailed preference information. Instead of requiring full rankings, cardinal weights, or extensive pairwise comparisons, each DM provides a binary indication of whether a criterion is considered critical for their judgment, and for each criterion, a single most-preferred alternative. This type of input is realistic in participatory workshops: stakeholders can typically identify which criteria are non-negotiable for them and can state "the best option under this criterion", even when they cannot (or do not want to) articulate complete orderings or precise trade-offs. The SGC is designed to aggregate exactly this kind of sparse and partially structured information, quantify the extent of agreement/disagreement across DMs, and derive a defensible selection of κ alternatives that balances key concerns while remaining transparent about the drivers of the group outcome.

Five alternatives ($a_1, \ldots, a_5$) and eight criteria ($g_1, \ldots, g_8$) are used as illustrated in Table 1.

The five alternatives reflect commonly discussed families of urban freight interventions. A peripheral distribution center (a_1) relocates transshipment to the outskirts to reduce inner-city truck movements but may increase handling and coordination requirements. Electric cargo bikes (a_2) can lower emissions and reduce space occupancy in the core, yet their applicability depends on shipment profiles and micro-depot access [35]. Night-time delivery windows (a_3) can shift freight flows away from peak hours but raise concerns about noise and enforcement [36]. An urban consolidation hub (a_4) aims to consolidate shipments closer to demand areas, potentially reducing vehicle-kilometers in the center, but it requires strong stakeholder participation and viable business models [37]. A crowdshipping platform (a_5) leverages occasional carriers to increase flexibility, but it raises questions regarding reliability, governance, and acceptance [38].

To evaluate these strategies, we consider eight criteria that capture the main trade-offs faced by the municipality and stakeholders: investment and operating cost (g_1), operational feasibility (g_2), effects on traffic congestion (g_3), expected emissions reduction (g_4), service level and reliability (g_5), stakeholder acceptance (g_6), scalability beyond the pilot phase (g_7), and consistency with the city's policy and regulatory framework (g_8). The criteria set is intentionally broad to mirror the multi-dimensional nature of urban freight planning, where technical performance alone is insufficient for implementation.

Six DM $\mathcal{S} = \{s_1, \ldots, s_6\}$ provide binary keyness on criteria and, for each criterion, name a single most-preferred alternative. The keyness matrix $R = [r^1 \; \cdots \; r^6] \in \{0,1\}^{8\times 6}$ (rows $g_1 \ldots g_8$, columns $s_1 \ldots s_6$) is illustrated in Table 2.

For preferences, each DM s_t supplies a criterion-winner matrix $P^{(t)} \in \{0,1\}^{5\times 8}$ in Table 3, with exactly one "1" per column.

Table 2
Keyness matrix R (rows: g_j, columns: s_t).

	s_1	s_2	s_3	s_4	s_5	s_6
g_1	1	0	1	0	0	1
g_2	1	0	1	1	0	1
g_3	1	1	0	1	1	0
g_4	1	1	0	0	1	0
g_5	0	0	1	1	0	1
g_6	1	1	1	0	1	1
g_7	0	0	1	0	0	1
g_8	1	0	0	1	1	0

Table 3
Preference matrices $P^{(1)}$–$P^{(6)}$(DMs s_1 to s_6).

DM 1 s_1

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
a_1	0	0	0	0	0	0	0	0
a_2	0	0	0	1	0	1	0	1
a_3	0	1	0	0	1	0	0	0
a_4	0	0	1	0	0	0	0	0
a_5	1	0	0	0	0	0	1	0

DM 2 s_2

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
a_1	0	0	0	0	0	0	0	0
a_2	0	0	1	1	0	1	1	1
a_3	0	1	0	0	1	0	0	0
a_4	0	0	0	0	0	0	0	0
a_5	1	0	0	0	0	0	0	0

DM 3 s_3

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
a_1	0	0	0	0	0	0	0	0
a_2	0	0	0	1	0	0	0	0
a_3	0	1	1	0	1	1	0	1
a_4	0	0	0	0	0	0	0	0
a_5	1	0	0	0	0	0	1	0

DM 4 s_4

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
a_1	0	0	0	0	0	0	0	0
a_2	0	0	0	1	0	1	0	0
a_3	1	1	0	0	0	0	1	1
a_4	0	0	1	0	1	0	0	0
a_5	0	0	0	0	0	0	0	0

DM 5 s_5

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
a_1	0	0	0	0	0	0	0	0
a_2	0	1	1	1	0	1	1	1
a_3	0	0	0	0	0	0	0	0
a_4	0	0	0	0	1	0	0	0
a_5	1	0	0	0	0	0	0	0

DM 6 s_6

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
a_1	0	0	0	0	0	0	0	0
a_2	0	0	0	1	0	0	0	0
a_3	1	1	0	0	1	0	1	1
a_4	0	0	1	0	0	1	0	0
a_5	0	0	0	0	0	0	0	0

Using the elicited R and $\{P^{(t)}\}_{t=1}^{6}$, the SGC aggregation in (3) yields the collective score vector

$$F = \left[0,\ 63,\ 60,\ 23,\ 16\right]$$

for $(a_1, \dots, a_5)$. With $\kappa = 2$, the resulting winner set is $\mathcal{W}_2 = \{a_2, a_3\}$; these two alternatives are close to each other and clearly ahead of the remainder. Fig. 2 displays the decomposition of F into the stacked contributions $f^{(t)} = P^{(t)}u$ by DMs. The visual confirms that support for a_2 and a_3 is broad-based across stakeholders rather than driven by a single DM, whereas a_4 and a_5 receive thinner and more concentrated support. No contributions accrue to a_1, which does not win on any criterion for any DM.

For comparability, two normalizations are reported. The keyness-based normalization scales by the best score compatible with the observed keyness pattern, $F_i^{\text{key}} = F_i / \left(T \sum_j u_j\right)$, which here evaluates to

$$F^{\text{key}} = [0.000,\ 0.389,\ 0.370,\ 0.142,\ 0.099]$$

using $T = 6$ and $\sum_j u_j = 27$. The ideal normalization scales by the absolute upper bound nT^2 ($n = 8$, $T = 6$), giving $F^{\text{ideal}} = [0.000, 0.219, 0.208, 0.080, 0.056]$. In both scales, a_2 leads narrowly over a_3, and both dominate the remainder.

At the criterion level, the keyness counts are $u = (3, 4, 4, 3, 3, 5, 2, 3)$, corresponding to rates

$$\rho = u/T = (0.50,\ 0.67,\ 0.67,\ 0.50,\ 0.50,\ 0.83,\ 0.33,\ 0.50).$$

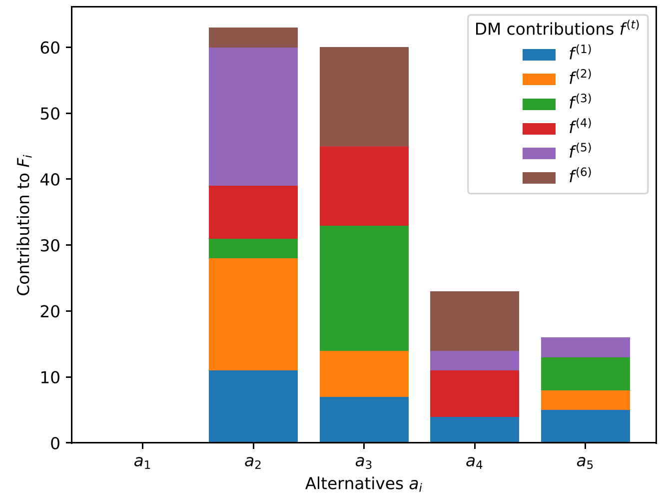


Fig. 2. Decomposition of F by DM: bars for each alternative a_i stack the vectors $f^{(t)}$, and the total heights equal F_i.

Stakeholder acceptance g_6 is the most salient (marked key by five of six DMs), while feasibility g_2 and congestion g_3 are strongly salient (four of six). Scalability g_7 is least salient (two of six). The dispersion of keyness, measured by entropy $H_u = -\sum_j \tilde{u}_j \log \tilde{u}_j$ with $\tilde{u}_j = u_j / \sum_\ell u_\ell$, equals $H_u \approx 2.05$ (effective number $N_{\text{eff}} = \exp(H_u) \approx 7.7$), indicating that perceived importance is fairly spread across criteria rather than concentrated on just a few.

Agreement among those who marked a criterion as key is summarized by $d_j = \max_i c_{ij}^{\text{key}} / u_j$. The profile

$$d = (2/3,\ 1,\ 1/2,\ 1,\ 2/3,\ 3/5,\ 1/2,\ 2/3)$$

shows two decisive criteria with full agreement ($d_j = 1$) on the most-preferred alternative: feasibility g_2 and emissions g_4. While congestion g_3 and scalability g_7 exhibit the most divided views ($d_j = 1/2$). The average agreement across key criteria is $\bar{d} \approx 0.70$, suggesting a generally coherent pattern with a few contested aspects.

Breadth of support across criteria, $B_i = |\{j : c_{ij} > 0\}|$, equals

$$B = (0,\ 6,\ 7,\ 3,\ 2).$$

Alternative a_3 wins on the largest number of distinct criteria, closely followed by a_2, whereas a_4 and a_5 win on few criteria and a_1 on none. Because SGC weights wins by keyness counts u_j, coverage on highly salient criteria is particularly influential. In practical terms, a_2 benefits from unanimity on emissions (g_4) and strong support on the most salient acceptance criterion (g_6), while a_3 is anchored by near-unanimity on feasibility (g_2) and solid support on service (g_5) and policy fit (g_8). Taken together, these indices corroborate the SGC outcome: $\{a_2, a_3\}$ forms a robust compromise pair, drawing wins on the criteria that the group considers most important and where agreement is highest, whereas areas of lower agreement are simultaneously of lower salience and thus exert limited influence on the final choice.

5.1. Case with uncertainty

To assess the method under incomplete inputs, we assume DM present uncertainty in which some entries of the keyness matrix R and the winner matrices $\{P^{(t)}\}$ are marked as x (unknown). The data is presented in Tables 4 and 5.

Unknown entries are treated as binary bits to be resolved subject to SGC's one-hot constraints. We then generate $10,000$ feasible completions (uniform over the admissible set) and compute, for each completion ω, the score vector $F^{(\omega)}$ and winner set $\mathcal{W}_\kappa^{(\omega)}$ with $\kappa = 2$. The dispersion of winner sets across completions is summarized

Table 4
Keyness matrix R with uncertain inputs.

	s_1	s_2	s_3	s_4	s_5	s_6
g_1	x	0	x	0	0	1
g_2	1	0	1	1	0	1
g_3	x	1	x	x	1	0
g_4	1	1	x	0	x	0
g_5	x	0	1	x	0	x
g_6	1	1	x	0	1	1
g_7	0	0	1	0	0	1
g_8	1	0	0	1	1	0

Table 5
Preference matrices $P^{(1)}$–$P^{(6)}$ with uncertain inputs.

DM 1 s_1

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
a_1	0	x	0	0	x	0	0	0
a_2	0	x	0	1	x	1	0	1
a_3	0	x	0	0	x	0	0	0
a_4	0	x	1	0	x	0	0	0
a_5	1	x	0	0	x	0	1	0

DM 2 s_2

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
a_1	0	0	0	0	0	0	0	0
a_2	0	0	1	1	0	1	1	1
a_3	0	1	0	0	1	0	0	0
a_4	0	0	0	0	0	0	0	0
a_5	1	0	0	0	0	0	0	0

DM 3 s_3

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
a_1	0	x	x	x	x	0	0	0
a_2	0	x	x	x	x	0	0	0
a_3	0	x	x	x	x	1	0	1
a_4	0	x	x	x	x	0	0	0
a_5	1	x	x	x	x	0	1	0

DM 4 s_4

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
a_1	x	x	0	0	0	0	0	0
a_2	x	x	0	1	0	1	0	0
a_3	x	x	0	0	0	0	1	1
a_4	x	x	1	0	1	0	0	0
a_5	x	x	0	0	0	0	0	0

DM 5 s_5

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
a_1	0	0	x	0	0	0	0	0
a_2	0	1	x	1	0	1	1	1
a_3	0	0	x	0	0	0	0	0
a_4	0	0	x	0	1	0	0	0
a_5	1	0	x	0	0	0	0	0

DM 6 s_6

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
a_1	0	x	0	0	x	0	0	0
a_2	0	x	0	1	x	0	0	0
a_3	1	x	0	0	x	0	1	1
a_4	0	x	1	0	x	1	0	0
a_5	0	x	0	0	x	0	0	0

by the ambiguity index $\mathrm{Agre}_\kappa = |\cap_\omega \mathcal{W}_\kappa^{(\omega)}|/|\cup_\omega \mathcal{W}_\kappa^{(\omega)}|$, which here is $\mathrm{Agr}_2 \approx 0.75$, indicating rather low variation in the selected pair across completions.

Uncertainty at the DM and criterion levels is visualized in Fig. 3. The left panel plots, for each DM, the fraction of unknown keyness entries (horizontal axis) against the fraction of unknown winner declarations (vertical axis). In this instance, s_3 exhibits the largest combined missingness, whereas s_2 has the smallest. The right panel plots, for each criterion g_j, the unknown-keyness rate U_j^{key} versus the unknown-winner rate U_j^{pref} (aggregated over DMs). Criteria g_3 and g_5 show the highest keyness uncertainty ($U_j^{\text{key}} \approx 0.5$), and g_2 shows the highest preference uncertainty ($U_j^{\text{pref}} \approx 2/3$). These two views help target clarifications toward the most leveraged respondents and criteria.

Fig. 4 visualizes the distribution of alternative scores across the $10,000$ sampled completions of the unknown bits. Each violin encodes the empirical density of $F_i^{(\omega)}$ over completions ω, with the extremal dots indicating the estimated lower and upper bounds $\left[F_i^{\min}, F_i^{\max}\right]$ for each i. Two features stand out. First, the mass of the distributions for a_2 is concentrated at substantially higher scores than for the remaining options, with overlap with a_3, a_4 and a_5. Second, the width of each violin reflects sensitivity to unresolved entries: wider shapes (and longer whiskers) signal outcomes that swing more depending on how the missing bits resolve. Overall, the violin plot provides a scale-aware picture of uncertainty: it shows not only who leads on average, but also how far an alternative could plausibly move under admissible completions.

Complementing the score view, Fig. 5 reports the stacked distribution of ranks: for each rank $r \in \{1, \dots, 5\}$, the bar height shows the share of completions in which some alternative attains rank r, and the colored segments attribute that share to each a_i. This representation is scale-free and directly tied to choice: it reveals how frequently an alternative appears as the best, second-best, and so on, regardless of absolute score levels or normalizations. The rank-1 column is dominated by a_2; the rank-2 column is dominated by a_3, with relatively small contributions from a_4 and a_5. It indicates the top of the ordering is typically formed by $\{a_2, a_3\}$. Lower ranks (4–5) are largely occupied by a_1 and a_5.

5.2. Final recommendation

In the *certain* run, the collective scores are $F = \left[0,\ 63,\ 60,\ 23,\ 16\right]$, so for $\kappa = 2$ the SGC recommendation is $\mathcal{W}_2 = \{a_2,\ a_3\}$. Both a_2 and a_3 draw wins on criteria that many DMs mark as key (see Fig. 2), which is also reflected by their normalized scores (both F^{key} and F^{ideal}) being well above the rest. At the criterion level, the keyness counts are $u = (3, 4, 4, 3, 3, 5, 2, 3)$, i.e., rates

$$\rho = u/6 = (0.50,\ 0.67,\ 0.67,\ 0.50,\ 0.50,\ 0.83,\ 0.33,\ 0.50).$$

If this is a preliminary screening process, we can select the key criteria to be analyzed later. Under an inclusive screen $\rho_j \geq 0.5$, the *key* criteria set is $\mathcal{G}_{\text{key}}^{(0.5)} = \{g_1, g_2, g_3, g_4, g_5, g_6, g_8\}$ (exclude g_7). Within the key set, feasibility g_2 and emissions g_4 are the most decisive (highest d_j), while stakeholder acceptance g_6 is the most salient (largest u_j).

In the *uncertain* run, the winner-set ambiguity is low, $\mathrm{Agr}_2 \approx 0.75$, indicating rather low dispersion of $\mathcal{W}_2^{(\omega)}$ across feasible completions. The distributional evidence points to the same short list: the stacked rank distribution (Fig. 5) concentrates rank $1/2$ mass on $\{a_2, a_3\}$. Hence, even under missing inputs, the preferred pair remains $\{a_2, a_3\}$.

6. Method comparison and practical guidance

This section first positions SGC against representative established approaches by comparing the amount of information that must be elicited from the DMs. We compare SGC with two established methods, AHP and PROMETHEE. Both are widely used in group decision-making and are commonly treated as benchmark approaches in applied studies [39,40].

6.1. Elicitation-burden comparison by information counts

Because the input formats of different methods are not directly commensurable (binary judgments, pairwise comparisons, cardinal ratings, and parameter settings), we adopt a simple and transparent proxy: *one elicited response is counted as one information item*. Under this proxy, we compare how the elicitation burden scales with the number of criteria n, alternatives m, and decision makers T.

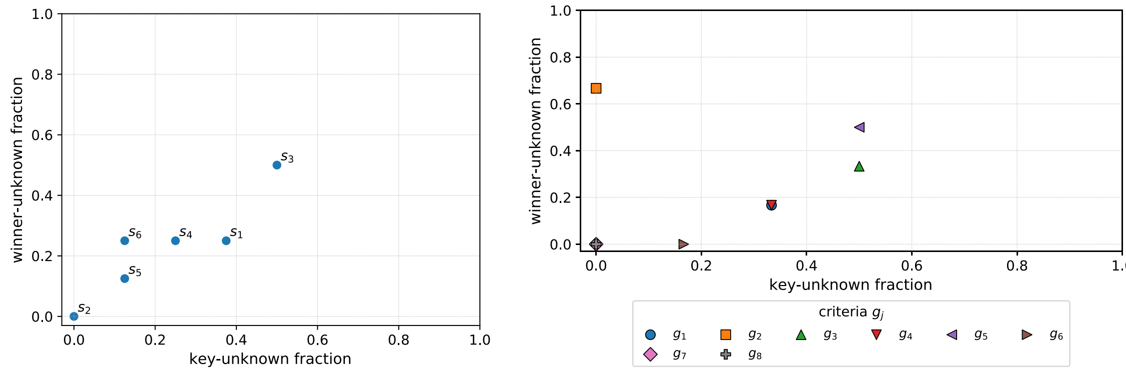
In SGC, each DM provides, for each criterion, a binary keyness declaration and a single most-preferred alternative. Counting each as one response, the elicitation burden is

$$I_{\text{SGC}}(n, m, T) = T\,(n + n) = 2nT. \tag{10}$$

This grows linearly in n and T and *does not grow* with m under the present winner-per-criterion elicitation. This is a valuable feature because the elicitation burden remains stable even as the number of alternatives expands, avoiding participant overload in large option sets.

For AHP, elicitation consists of pairwise comparisons among criteria and, for each criterion, pairwise comparisons among alternatives [2]. The number of pairwise comparisons for criteria is $n(n-1)/2$, and for alternatives it is $m(m-1)/2$ per criterion. Therefore,

$$I_{\text{AHP}}(n, m, T) = T\left(\frac{n(n-1)}{2} + n\,\frac{m(m-1)}{2}\right) = \frac{T}{2}\Big(n(n-1) + n\,m(m-1)\Big).$$



(a) DM uncertainty: keyness-unknown vs winner-unknown fractions.

(b) Criterion uncertainty: U_j^{key} vs U_j^{pref}.

Fig. 3. Uncertainty patterns across decision makers and criteria under missing bits x.

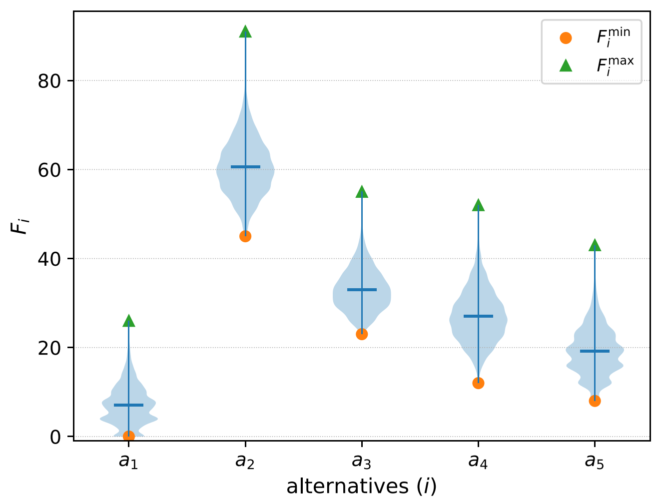


Fig. 4. Score distributions across sampled completions: violin densities of $F_i^{(\omega)}$ for $i \in \{1,\ldots,5\}$ with dots indicating $F_i^{\min}$ and $F_i^{\max}$.

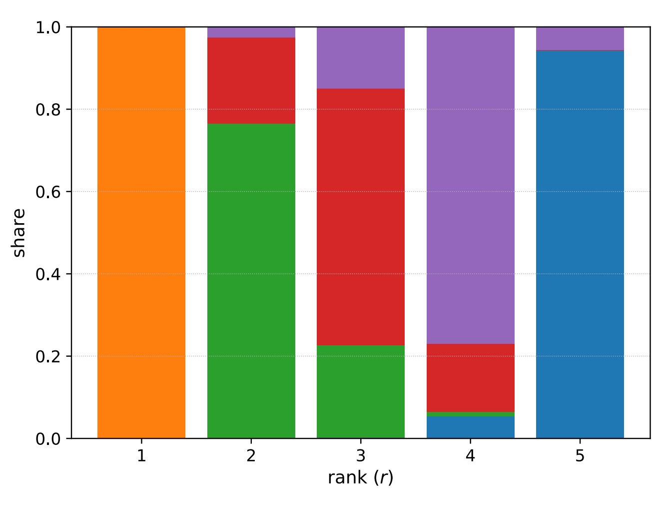


Fig. 5. Stacked rank distribution across sampled completions: for each rank r, the bar shows the share of completions assigning rank r to each alternative a_i.

(11)

This scales as $O(n^2 + nm^2)$ per DM.

For PROMETHEE, we consider a standard workflow in which criteria weights are elicited via a revised Simos-type procedure and PROMETHEE preference functions require criterion-level parameter specification [12,41]. Under a deliberately conservative counting consistent with the case-study questionnaire perspective, we assume each DM provides: (i) a ranking/ordering of the n criteria plus one additional parameter z (counted as $n+1$ items) [42], (ii) ratings of m alternatives on n criteria (counted as mn items), and (iii) at least one preference-parameter value per criterion (counted as n items). This yields

$$I_{\text{PRO}}(n,m,T) = T\,(mn + (n+1) + n) = T\,(mn + 2n + 1). \tag{12}$$

This scales as $O(mn)$ per DM, plus additional linear terms for weights and parameters.

In the case study, $n = 8$, $m = 5$, and $T = 6$. Table 6 reports the implied information counts.

The comparison highlights the main practical advantage of SGC: it replaces the quadratic elicitation growth of AHP and the dense performance-rating and parameter specification required by PROMETHEE with a linear number of short responses, and the inputs are purely binary. This reduction is particularly important when T is not small, since the total elicitation burden scales proportionally with T for all methods.

Fig. 6 visualizes how the total elicitation burden scales as the number of criteria or alternatives increases. The left panel varies n while fixing $m = 5$ and $T = 6$ (case-study values for m and T). The right panel varies m while fixing $n = 8$ and $T = 6$. Consistent with the closed-form counts above, SGC grows linearly in n and is *constant* in m under the winner-per-criterion elicitation, PROMETHEE grows approximately linearly with mn, and AHP grows rapidly due to pairwise comparisons among alternatives repeated across criteria.

6.2. Qualitative comparison: effectiveness and reliability in practice

Beyond elicitation burden, it is important to clarify what SGC delivers relative to established MCDA families and how reliable its recommendations are under the intended use case. We focus on qualitative effectiveness and reliability dimensions that are meaningful across approaches: transparency of the aggregation logic, dependence on calibrated parameters, robustness to incomplete inputs, and suitability for shortlisting rather than full ranking:

- **Transparency and explainability.** SGC produces an alternative-level support score by a simple counting rule: each declared

Table 6
Elicitation burden as information counts for the case study ($n = 8$, $m = 5$, $T = 6$).

Method	Per-DM items	Total items ($\times T$)	Scaling (per DM)
SGC	$2n = 16$	$2nT = 96$	$O(n)$
AHP	$\frac{n(n-1)}{2} + n\frac{m(m-1)}{2} = 108$	648	$O(n^2 + nm^2)$
PROMETHEE	$mn + 2n + 1 = 57$	342	$O(mn)$

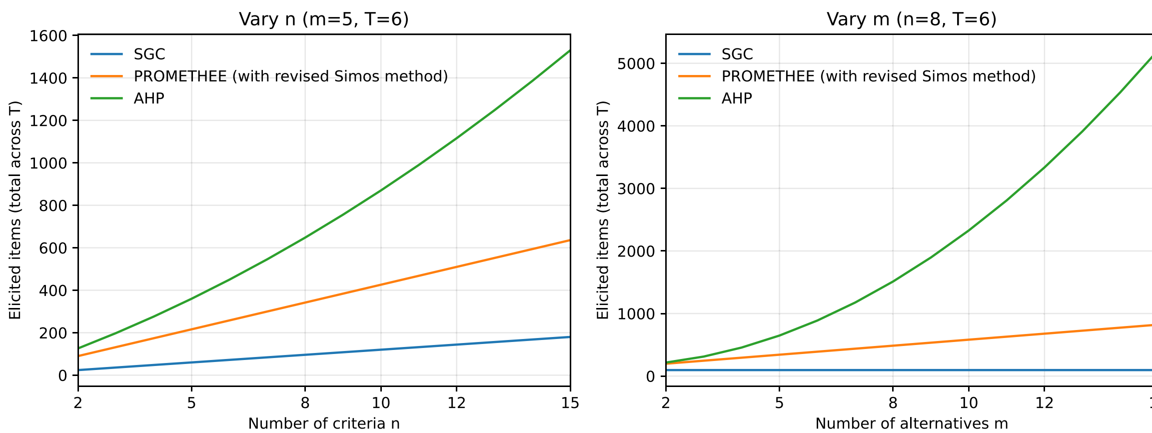


Fig. 6. Total elicitation burden (information counts) versus problem size, using the counting rules in Section 6. Left: vary criteria n with $m = 5$, $T = 6$. Right: vary alternatives m with $n = 8$, $T = 6$.

criterion-level win is weighted by the number of DMs who marked that criterion as key, and these weighted wins are summed across DMs. This provides a direct audit trail: it is always possible to explain a high score as "many stakeholders picked this alternative on criteria that many stakeholders consider key". By contrast, AHP and PROMETHEE can be explainable, but their outputs rely on intermediate constructs (pairwise-comparison matrices and eigenvector weights in AHP; preference functions, thresholds, and flows in PROMETHEE) that are less straightforward to communicate in a shortlisting workshop. In settings where acceptance of the recommendation depends on rapid comprehension by diverse stakeholders, the smaller conceptual distance from elicited inputs to outputs is an effectiveness advantage of SGC.

- **Parameter dependence and calibration risk.** A key reliability consideration is whether the method requires parameters that are difficult to justify or calibrate at early stages. PROMETHEE requires choosing a preference function and specifying criterion-level parameters, and AHP relies on consistency of pairwise comparisons and often requires revising inconsistent matrices. These elements can be well supported when participants are trained and time is available, but in time-constrained group settings they can introduce substantial calibration risk and sensitivity. SGC avoids such calibration by construction: the only required inputs are binary keyness tags and per-criterion winners, and the aggregation does not require threshold tuning, utility scaling, or interpersonal comparability of cardinal weights. This reduces degrees of freedom that can otherwise drive unstable recommendations.
- **Robustness to missing or ambiguous inputs.** Incomplete responses are common in participatory processes. SGC explicitly accommodates this by treating missing entries as unknown binary variables and analyzing all feasible completions, yielding necessary and possible winners, score bounds, and agreement indices. This provides an operational notion of reliability: the recommendation is reliable when the selected alternatives remain winners across a wide range of admissible completions, and the diagnostics identify where additional clarification would matter most. By contrast, AHP and PROMETHEE typically presume complete input structures; missing comparisons or missing performance ratings must be imputed, simplified, or re-elicited, often without a built-in set-based robustness interpretation tied to the missingness. The ability to report what is stable under incompleteness is a practical advantage of SGC.
- **Stability under group heterogeneity.** Group decisions often involve heterogeneous stakeholders whose priorities differ. SGC uses the group keyness vector u to weight criterion-level wins by how broadly the criterion is regarded as key, which tends to reward alternatives that perform well on widely supported criteria and discourages outcomes driven by idiosyncratic concerns. This can improve perceived legitimacy in groups where compromise options are preferred. Importantly, SGC does not claim to replace richer preference models when intensity information is available; rather, it aims to deliver a stable and interpretable shortlist under minimal elicitation, and to make disagreement and uncertainty explicit through diagnostics.

Although applicable in various context, the qualitative comparison suggests that SGC is most effective when the decision phase prioritizes rapid shortlisting, when participants cannot provide reliable cardinal trade-offs, and when transparency and robustness to incompleteness are primary requirements. In later phases, when a smaller set of alternatives remains and more time and data are available, SGC can be used as a structured front end: it narrows the decision space, identifies key criteria with broad support, and highlights contested dimensions that merit deeper modeling in methods such as PROMETHEE or value-based approaches.

6.3. Key parameters and how to justify them in theory and practice

In addition to comparing it to other methods, we discuss the theoretical and practical significance of the key parameters of SGC. Also, applying SGC requires a small number of design choices. These choices are transparent and can be motivated using both theoretical considerations (interpretability, robustness, neutrality) and managerial constraints (resources, governance, decision phase):

- **Portfolio size κ.** The parameter κ specifies how many alternatives are retained. κ reflects the choice problematic (shortlisting rather than full ranking) and prevents over-interpretation of sparse preference information. Managerially, κ is typically fixed by capacity constraints (budget, staffing, time to pilot), governance rules (e.g., "select two measures for the next policy cycle"), or a staged decision process in which κ alternatives proceed to detailed analysis. When the goal is shortlisting, $\kappa > 1$ is natural because it preserves a small set of broadly acceptable options and avoids over-interpreting sparse inputs. When a single commitment decision is required, $\kappa = 1$ yields a single winner. In applications where the desired shortlist size is uncertain, a practical approach is to report results for a small range (e.g., $\kappa \in \{1, 2, 3\}$) and check whether the leading alternatives remain stable across these values.

- **Tie handling at the cutoff.** Because SGC uses discrete counts, ties can arise at the κth score. A tie-inclusive rule (returning all alternatives with $F_i \geq \tau_\kappa$) is theoretically attractive because it preserves neutrality and avoids arbitrary tie-breaking. In managerial terms, tie inclusion is often acceptable in early screening because the output is a shortlist to be further discussed. If a fixed shortlist size is mandatory, a pre-declared secondary rule can be used (e.g., prefer larger key-coverage B_i^{key}, or decide by facilitated deliberation).
- **Keyness elicitation constraints.** In some applications, it is useful to cap the number of criteria each DM can mark as key (e.g., "select up to κ_{key} key criteria"). Regardless of whether the cap is set, keyness declarations r_j^t should be elicited with a clear instruction so that "key" carries a consistent meaning across DMs. Two simple protocols are useful:
  - *Free keyness:* allow any number of criteria to be marked key, but instruct DMs to mark only criteria that are non-negotiable or that they would not be willing to trade away. This encourages sparsity without enforcing an explicit cap.
  - *Budgeted keyness:* ask each DM to select at most K key criteria (or exactly K in a top-K variant). This focuses attention, limits respondent effort, and prevents degenerate declarations. In practice, small values such as $K \in \{3, 4, 5\}$ are often workable; the appropriate K should reflect the breadth of the criteria set and the desired sharpness of the keyness signal.
- **Screening threshold for criteria (diagnostics).** If SGC is used as a front end to more detailed MCDA, the keyness rate threshold τ_ρ (e.g., retain criteria with $\rho_j \geq \tau_\rho$) controls how strict the screening is. Theoretically, larger τ_ρ yields a more consensual "core" criterion set; managerially, it can be aligned with governance conventions. A useful interpretation is: $\tau_\rho = 0.5$ corresponds to a majority of DMs marking the criterion as key, $\tau_\rho = 1$ corresponds to unanimity, and smaller values are more inclusive but may reduce interpretability by admitting criteria that only a small minority consider critical. A practical recommendation is to report the screened set for one baseline value (e.g., $\tau_\rho = 0.5$) and provide a brief sensitivity check at a stricter value (e.g., $\tau_\rho = 2/3$ or 1) when the stakes of screening are high.

7. Conclusion

This paper introduced Simple Group Choice (SGC), a lightweight, preference-based approach to address multi-criteria group choice problem that operates solely on binary inputs: for each criterion, DMs indicate whether it is key and name a single most-preferred alternative. SGC aggregates these primitives transparently by counting: individual winner declarations are weighted by the group's keyness pattern and pooled across DMs to yield an alternative-level score vector, from which the top-κ set is selected. We established elementary properties that justify the rule and showed how SGC can be implemented with minimal elicitation burden and negligible computational overhead. Beyond the base model, we developed diagnostics that help screen criteria and interpret outcomes, and we extended SGC to incomplete inputs via a missing-bit framework that delivers necessary/possible winners, score bounds, inclusion frequencies, and an agreement index. A didactic urban logistics study illustrated how SGC supports transparent shortlisting: in the certain case, a_2 and a_3 are selected; under uncertainty, they remain the leading pair in most completions, while the diagnostic indices pinpoint a small set of high-leverage clarifications that would stabilize the result.

We should note some limitations of SGC. First, binarization discards intensity information about both keyness and within-criterion performance. This is deliberate to minimize elicitation burden and avoid spurious precision. But we should acknowledge it prevents fine-grained trade-off analysis. Second, the current formulation assumes a unique most-preferred alternative per decision maker and criterion (one-hot columns in $P^{(t)}$). In settings where ties or abstentions are natural, a relaxed variant ($\sum_i p_{ij}^t \leq 1$ or allowing multiple top choices) is needed and would slightly modify the counting semantics. Finally, although the cross-view weighting via the group keyness vector mitigates the effect of any one decision maker marking many criteria as key, the winners on criteria whose keyness grows do benefit mechanically; calibration (or normalization) of keyness per respondent may be warranted when standards differ markedly across participants.

Several directions are promising. On elicitation, structured keyness, e.g., hierarchical criteria, budgeted caps on the number of key criteria, or respondent-level normalization, could enhance fairness and comparability without abandoning binary inputs. On modeling, integrating minimal quantitative data when available, e.g., allowing a criterion-specific pre-screen based on performance thresholds, and adding feasibility or resource constraints on the final selection, e.g., knapsack-type multi-winner formulations would broaden applicability. Finally, developing open-source survey and analytics tools that implement SGC end-to-end with questionnaire design, diagnostics, uncertainty analysis, and robust recommendations will facilitate adoption in participatory planning, procurement, and early-phase portfolio down-selection, where transparent and low-burden group choice is most valuable.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] Salo A, Hämäläinen RP. Multicriteria decision analysis in group decision processes. Handbook of group decision and negotiation. Springer; 2010, p. 269–83.
[2] Saaty RW. The analytic hierarchy process—what it is and how it is used. Math Model 1987;9:161–76.
[3] Ferretti V. From stakeholders analysis to cognitive mapping and multi-attribute value theory: An integrated approach for policy support. European J Oper Res 2016;253:524–41.
[4] Govindan K, Jepsen MB. Electre: A comprehensive literature review on methodologies and applications. European J Oper Res 2016;250:1–29.
[5] Fraser NM, Hauge JW. Multicriteria approval: application of approval voting concepts to mcdm problems. J Multi-Criteria Decis Anal 1998;7:263–73.
[6] Riabacke M. A Prescriptive Approach to Eliciting Decision Information [Ph.D. thesis.], Department of Computer and Systems Sciences. Stockholm University; 2012.
[7] Boix-Cots D, Pardo-Bosch F, Pujadas P. A systematic review on multi-criteria group decision-making methods based on weights: Analysis and classification scheme. Inf Fusion 2023;96:16–36.

[8] Kelly JS. Social choice theory: An introduction. Springer Science & Business Media; 2013.
[9] Laslier JF, Sanver MR. Introduction to the handbook on approval voting. Handbook on approval voting. Springer; 2010, p. 1–12.
[10] Greco S, Sł owiński R, Figueira JR, Mousseau V. Robust ordinal regression. Trends Mult Criteria Decis Anal 2010;241–83.
[11] Dyer JS. Multiattribute utility theory (maut). Multiple criteria decision analysis: state of the art surveys. Springer; 2016, p. 285–314.
[12] Brans JP, De Smet Y. Promethee methods. Multiple criteria decision analysis: state of the art surveys. Springer; 2005, p. 187–219.
[13] Vaidya OS, Kumar S. Analytic hierarchy process: An overview of applications. European J Oper Res 2006;169:1–29.
[14] Greco S, Ehrgott M, Figueira JR, editors. Multiple Criteria Decision Analysis: State of the Art Surveys. International series in operations research & management science, volume 233, 2016.
[15] Figueira J, Greco S, Ehrgott M. Multiple criteria decision analysis: state of the art surveys. Springer Science & Business Media; 2005.
[16] Roy B. Paradigms and challenges. Multiple criteria decision analysis: state of the art surveys. Springer; 2016, p. 19–39.
[17] Korhonen P, Moskowitz H, Wallenius J. Multiple criteria decision support-a review. European J Oper Res 1992;63:361–75.
[18] Figueira JR, Greco S, Roy B. Electre methods with interaction between criteria: An extension of the concordance index. European J Oper Res 2009;199:478–95.
[19] Huang H, Mommens K, Lebeau P, Macharis C. The multi-actor multi-criteria analysis (mamca) for mass-participation decision making. In: International conference on decision support system technology. Springer; 2021, p. 3–17.
[20] Huang H. Toward inclusive decision making: A systematic introduction of the mass-participation decision support framework. Soc Sci Humanit Open 2025;12:102093.
[21] Belton V, Pictet J. A framework for group decision using a mcda model: sharing, aggregating or comparing individual information? J Decis Syst 1997;6:283–303.
[22] Pasi G, Yager RR. Modeling the concept of majority opinion in group decision making. Inform Sci 2006;176:390–414.
[23] Herrera F, Herrera-Viedma E, et al. A model of consensus in group decision making under linguistic assessments. Fuzzy Sets and Systems 1996;78:73–87.
[24] Tran TNT, Felfernig A, Le VM. An overview of consensus models for group decision-making and group recommender systems. User Model User-Adapt Interact 2023.
[25] Huang H, De Smet Y, Macharis C, Doan NAV. Collaborative decision-making in sustainable mobility: identifying possible consensuses in the multi-actor multi-criteria analysis based on inverse mixed-integer linear optimization. Int J Sustain Dev World Ecol 2021;28:64–74.
[26] Kadziński M, Rocchi L, Miebs G, Grohmann D, Menconi ME, Paolotti L. Multiple criteria assessment of insulating materials with a group decision framework incorporating outranking preference model and characteristic class profiles. Group Decis Negot 2018;27:33–59.
[27] Huang R, Kadziński M, Siskos E, Burgherr P. Classifying swiss geothermal policies with the group robust flowsort method incorporating imprecise inputs and robustness concerns. Energy Econ 2025;108986.
[28] Tullock G. Problems of majority voting. J Politi. Econ 1959;67:571–9.
[29] McLean I. The borda and condorcet principles: three medieval applications. Soc Choice Welf 1990;7:99–108.
[30] Arrow KJ, Sen A, Suzumura K. In: Handbook of social choice and welfare, vol. 2, Elsevier; 2010.
[31] Lahdelma R, Hokkanen J, Salminen P. Smaa-stochastic multiobjective acceptability analysis. European J Oper Res 1998;106:137–43.
[32] Tervonen T, Figueira J, Lahdelma R, Salminen P. Smaa-iii: A simulation-based approach for sensitivity analysis of electre iii. Real-time and deliberative decision making: application to emerging stressors. Springer; 2008, p. 241–53.
[33] Vanwoensel T, Creten R, Vandaele N. Managing the environmental externalities of traffic logistics: The issue of emissions. Prod Oper Manage 2001;10:207–23.
[34] Onstein AT, Tavasszy LA, Van Damme DA. Factors determining distribution structure decisions in logistics: a literature review and research agenda. Transp Rev 2019;39:243–60.
[35] Narayanan S, Antoniou C. Electric cargo cycles-a comprehensive review. Transp Policy 2022;116:278–303.
[36] Holguín-Veras J, Encarnación T, González-Calderón CA, Winebrake J, Wang C, Kyle S, Herazo-Padilla N, Kalahasthi L, Adarme W, Cantillo V, et al. Direct impacts of off-hour deliveries on urban freight emissions. Transp Res Part D: Transp Environ 2018;61:84–103.
[37] Allen J, Browne M, Woodburn A, Leonardi J. The role of urban consolidation centres in sustainable freight transport. Transp Rev 2012;32:473–90.
[38] Mohri SS, Ghaderi H, Nassir N, Thompson RG. Crowdshipping for sustainable urban logistics: A systematic review of the literature. Transp Res Part E: Logist Transp Rev 2023;178:103289.
[39] Macharis C, Brans JP, Mareschal B. The gdss promethee procedure. J Decis Syst 1998;7:283–307.
[40] Escobar MT, Moreno-Jimenéz JM. Aggregation of individual preference structures in ahp-group decision making. Group Decis Negot 2007;16:287–301.
[41] Figueira J, Roy B. Determining the weights of criteria in the electre type methods with a revised simos' procedure. European J Oper Res 2002;139:317–26.
[42] Huang R, Kadziński M, Figueira JR, Corrente S, Siskos E, Burgherr P. A modular simos-roy-figueira framework for tailored weight elicitation in multi-criteria decision aiding. Expert Syst Appl 2026;131315.