

Beneficiary Satisfaction with an Agroecology-Based Empowerment Program of a Hydropower Company in Asahan, Indonesia

Zainuddin Iqbal Sidabutar*, Subandi*, Dewi Hajar**

*(Community Development Division, PLTA PT Indonesia Asahan Aluminium (Persero), P.O. Box 1 Porsea, Paritohan, Pintu Pohan Meranti District, Toba Regency, North Sumatra 22388, Indonesia

Email: iqbal@inalum.id; subandi@inalum.id

** (Environmental Engineering, Diponegoro University, Semarang City, Central Java, Indonesia

Email: 123dewihajar@gmail.com

Corresponding author: Zainuddin Iqbal Sidabutar (iqbal@inalum.id)

Abstract:

This study aims to evaluate how satisfied the beneficiaries of RANUM ASAHAN (Ruang Agroekologi dan Nilai Unggul Masyarakat Asahan) are with the program, a community empowerment initiative of the hydropower plants of PT Indonesia Asahan Aluminium (Persero) in Aek Songsongan Village, Asahan Regency, by using the Community Satisfaction Index (Indeks Kepuasan Masyarakat, IKM) based on 19 items. The study employs a quantitative approach using an August 2026 interview survey of 38 beneficiaries. Rating counts were reconstructed from the published percentages and tested item by item against the satisfaction threshold (3.00) with Holm-adjusted t-tests and exact sign tests. The results indicate that all 19 items exceed the threshold (adjusted $p \leq 0.0002$). The highest scores are found in environmental quality (3.74), method appropriateness (3.66), and company responsiveness (3.66), and thirteen items received no rating below satisfied. However, the lowest scores are observed in income gain and facilitator competence (both 3.39), and all nine less-satisfied ratings (1.2% of 722) fall on six items: relevance, skills, income, readiness to run the program without the company, and participation in planning and in monitoring. The reported total of 88.17 is “very satisfied” on the report’s scale and grade A under the 2004 guideline, but it lies 0.14 points under grade A of the 2017 guideline, and the difference is not statistically significant ($p = 0.60$). This study concludes that the beneficiaries are highly satisfied, with the environment, the method, and the company’s responsiveness as the main strengths and income and self-reliance as the main gaps. To improve the program, it is recommended to record production and income for each activity, to prepare a handover plan for each activity, and to state the interpretation scale and the confidence intervals in future surveys.

Keywords — community satisfaction index, community empowerment, corporate social responsibility, agroecology, hydropower, program evaluation, Indonesia.

I. INTRODUCTION

A. Corporate programs and community perceptions

Corporate social responsibility (CSR) in Indonesia has evolved beyond philanthropy. Companies are now expected to operate efficiently and, at the same time, to improve the welfare of neighboring communities. This expectation is also a legal obligation, because Law No. 40 of 2007 makes social and environmental responsibility an

obligation for companies whose business relates to natural resources [1, 2]. Beyond legal compliance, a community that receives few benefits from a nearby operation raises a company’s costs and weakens its social license to operate [3, 4], and acceptance depends on perceived fairness, on the quality of contact, and on trust [5]. The same applies to hydropower. Dams and power stations depend on rivers, land, and the support of surrounding villages, and practitioners have long argued that host communities should share in the benefits [6].

Empowerment programs, however, frequently fall short of their objectives; they tend to decline when the company withdraws, fail to address actual community needs, or lack monitoring [2]. Geothermal CSR programs in West Java scored well on effectiveness but poorly on sustainability and empowerment [7], and stakeholder groups around Indonesian mines hold divergent views of the same programs [8]. Stakeholder theory explains why such views matter, since an organization is accountable to every group it affects [9], and stakeholder analysis helps companies incorporate these views into their plans [10]. Regular monitoring and evaluation is therefore essential to ensure that a social investment retains its value.

B. An agroecology-based program in Aek Songsongan

PT Indonesia Asahan Aluminium (Persero), through its hydropower plants (PLTA), runs RANUM ASAHAN (Ruang Agroekologi dan Nilai Unggul Masyarakat Asahan, approximately the Asahan Agroecology and Community Value Space) in Aek Songsongan Village, Asahan Regency [11]. The program comprises several activities. According to the survey report, it teaches organic farming, processing of farm produce, composting, waste sorting and a waste bank, homestay management, and disaster mitigation, and it creates income from organic produce, a tower café, homestays, organic fertilizer, and the waste bank. These activities draw on established concepts. Agroecology applies ecological principles to farming for people who have few outside inputs [12], and it functions simultaneously as a science, a movement, and a set of practices [13]. Livelihoods, in turn, are built from several assets and activities rather than a single source of income [14]. Community-based tourism raises questions of who participates and who makes decisions [15], and disaster preparedness draws on the social capital built among neighbours [16]. Small Indonesian programs that pair composting, hydroponics, or a waste bank with company funding illustrate how such components can be combined [17, 18, 19]. A program of this scope requires clear evidence of which components benefit its beneficiaries.

C. Measuring satisfaction with the IKM

Indonesia's Community Satisfaction Index (Indeks Kepuasan Masyarakat, IKM) was designed for public services. It converts four-point ratings into a score between 25 and 100 [20, 21]. Consistent with this origin, most IKM studies examine service counters and licensing offices. A public service mall in Banjarbaru scored 78.8 [22], a licensing office in Tanjungpinang 66.49 [23], an e-KTP service in Madiun about 80.7 [24], and a civil registry in Central Lombok 83.46 [25]. The instrument has also been adapted for CSR programs, including red ginger cultivation [26], village programs [27], and a mining company's empowerment scheme [28]. These adaptations tend to share three methodological limitations. First, they publish indices to two or three decimals from small samples without saying how uncertain the figures are or testing any hypothesis. Second, they apply grading schemes inconsistently, so that the same value may be classified as "satisfied", "good", or "fair". Third, they aggregate away the item-level differences that indicate priorities for improvement. Participation research adds that involvement should be judged stage by stage rather than as a single score [29, 30, 31].

The company report behind this paper follows the same pattern. Its sample of 38 respondents is sufficient for statistical inference, which the report does not perform, and the report contains minor inconsistencies that require attention. The present study reanalyzes the survey to quantify this uncertainty and to translate item results into management decisions.

D. Objectives and hypotheses

This study aims to analyze how satisfied the beneficiaries of RANUM ASAHAN are with the program, using the IKM approach. Its significance lies in the fact that the program bundles many activities, that the report gives index values without saying how certain they are, and that its sample of 38 respondents is sufficient for statistical inference. To this end, we pursue four objectives: to reproduce and validate the reported indices, to estimate how precisely they are known, to test hypotheses against the satisfaction threshold and against official grade boundaries, and to translate item results into action

groups. Three hypotheses guide the tests. H1 states that each of the 19 items has a mean above the satisfaction threshold of 3.00. H2 states that environmental quality is rated higher than the three outcomes that depend on time and effort, namely skills gain, income gain, and readiness to run the program without the company. H3 states that the total index reaches the 2017 grade A boundary (88.31). H1 is primary; H2 and H3 are secondary. A further descriptive objective is to identify the items on which the few less-than-satisfied ratings cluster.

II. METHODS

A. Study design and respondents

This is a secondary quantitative analysis of a survey that PT SUCOFINDO (Persero) Cabang Semarang carried out for the company [11]. The survey took place in August 2026, and a focus group discussion and questionnaire session was held on 3 August 2026. Respondents were representatives of the groups that receive the program and live in Aek Songsongan Village, Aek Songsongan District, Asahan Regency. The report lists 38 respondents and does not describe how they were chosen, so we treat the sample as a purposive set of group representatives and do not generalize beyond it. The report also gives no information on sex. Interviews followed a questionnaire with closed questions on a four-point scale (1 = dissatisfied, 2 = less satisfied, 3 = satisfied, 4 = very satisfied) [32], together with a few open questions.

B. Items and indices

The report groups 19 items into three aspects. Overall program has seven items: relevance, knowledge gain, skills gain, environmental quality, income gain, continuity, and readiness to run the program without the company. Program management has eight: participation in planning, implementation, and monitoring and evaluation; stakeholder synergy; and four facilitator items (closeness, adaptability, responsibility, competence). Program delivery has four: appropriateness of the method, reliability of the method, fairness of targeting, and responsiveness of the company. The facilitator is the company's community development officer (CDO). The index follows the official conversion [20, 21]:

$$IKM = 25 \times (1/k) \sum \bar{x}_i \quad (1)$$

where $\bar{x}_i$ is the mean rating of item i and k is the number of items. The report states three ways to read its scores: "satisfied" for 50 to 75 and "very satisfied" for 75 to 100; grades D to A of the 2004 guideline (A from 81.26) [20]; and grades D to A of the 2017 guideline (A from 88.31) [21]. All three are reported because they yield different classifications for scores in this range.

C. Reconstruction and validation of item counts

The report publishes each item as a distribution in whole percentages, without counts. With 38 respondents, whole-number percentages map to a unique count for every item, so we recovered all counts (Table III). We validated them in two steps. First, we recomputed each percentage from the counts and matched the report. Second, we recomputed the index of the management aspect and compared it with the published value. It matched to the third decimal (86.924). Rebuilt overall and delivery indices differ slightly from the published ones, and Section III-B reports and discusses the gaps. All statistics use the reconstructed counts.

D. Statistical analysis

Likert-type ratings can be analyzed with parametric methods when many respondents are available [33], and we added rank-based checks. For each item, we tested the mean against 3.00 with a one-sample t -test ($df = 37$) and an exact sign test that compares the number of 4s with the number of ratings below 3 and excludes ties. We report the standardized effect $d_z = (M - 3)/SD$ [34, 35] and 95% confidence intervals. Proportions use Wilson intervals [36, 37]. Both families of 19 tests were adjusted with the Holm procedure [38]. With 38 respondents and $\alpha = 0.05$, the study had 80% power to detect $|d_z| \geq 0.47$ for a single test, and $|d_z| \geq 0.67$ under Holm adjustment over 19 tests; all observed effects exceed the latter. For H2, we used Mann–Whitney tests of the environmental item against each of the three later outcomes, plus the knowledge–income contrast, with Holm adjustment over these four. Because the same respondents answered every item, treating ratings as independent is conservative for these contrasts, and they are treated as exploratory. For H3 and for

benchmarks, we treated items as units and ran one-sample *t*-tests of item means against 81.26 (2004 grade A), 88.31 (2017 grade A), and 84.75 (a published CSR benchmark [26]) for the management aspect ($df = 7$) and all 19 items ($df = 18$), adjusted over six tests. A Kruskal–Wallis test examined whether the 19 item distributions differ. Analyses used Python with SciPy.

E. Action groups

To translate the tests into decisions, we classified items by the number of respondents below “satisfied” and by the share rating “very satisfied”. Group A (sustain) has no rating below 3 and at least half rating 4. Group B (maintain) has no rating below 3 and fewer than half rating 4. Group C (watch) has one or two ratings below 3. Group D (act) has three or more (about 8%). We examine below how the result changes with the cut-off. The grouping is a pragmatic rule, not a validated scale, and its purpose is to direct management attention to items with less-than-satisfied ratings.

F. Ethics and positionality

The analysis used aggregated, anonymized figures only and does not reproduce the names listed in the report. The Declarations section states the authors’ relationship to the company and the ethics statement.

III. RESULTS

A. Respondents

The 38 respondents were mostly of working age (Table I). Most (32 of 38, 84%) were between 26 and 65 years old, five (13%) were under 25, and one was over 65. Most had finished senior high school (27 of the 39 listed, see below). Half were farmers, 30% daily laborers, and 20% traders. Approximately 80% reported a monthly income below IDR 1,000,000, and the rest between IDR 1,000,000 and IDR 3,000,000. This profile is consistent with a program that reaches low-income households engaged in farming and small trade.

TABLE I
PROFILE OF RESPONDENTS (N = 38)

Characteristic	n	% of 38	Reported %
Age (years)			
< 25	5	13.2	13
26–45	22	57.9	58

Characteristic	n	% of 38	Reported %
46–65	10	26.3	26
> 65	1	2.6	3
Education			
Elementary school	4	10.5	10
Junior high	8	21.1	20
Senior high	27	71.1	70
Occupation			
Farmer	19	50.0	50
Daily laborer	11	28.9	30
Trader	8	21.1	20
Monthly income (IDR)			
< 1,000,000	30	78.9	80
1,000,000–3,000,000	8	21.1	20

Note. Source: company survey report [11]. Education counts as printed in the report sum to 39, not 38. Reported percentages are rounded; none was reported for sex.

B. Reported and rebuilt indices

The report presents a total of 88.166 for RANUM ASAHAN, built from three aspect indices: 87.911 for the overall program, 86.924 for management, and 89.662 for delivery. Table II compares them with our reconstruction. The management index is reproduced exactly. The overall index is 0.15 points higher (88.06) and the delivery index 0.47 points higher (90.13). Equal weights give a 19-item total of 88.02, and the mean of the three rebuilt aspects gives 88.37. The report mentions weighting per question in its methods but does not describe the weights, so rounding or weighting may explain the gaps; the available data do not allow these explanations to be distinguished.

TABLE II
REPORTED AND RECONSTRUCTED INDICES

Aspect	Items	Reported	Rebuilt	Diff.
Overall program	7	87.91	88.06	-0.15
Program management	8	86.92	86.92	0.00
Program delivery	4	89.66	90.13	-0.47
Total (mean of three aspects)	19	88.17	88.37	-0.21
Total (19 items, equal weights)	19	88.17	88.02	+0.15

Note. Converted indices (25–100). Rebuilt values use the reconstructed counts. Diff. = reported – rebuilt. The reported total (88.166) is the mean of the three reported aspect indices. In its text the report also quotes 85.333 for delivery and 88.134 for the total.

Four further inconsistencies in the report require attention. First, the text under the delivery table quotes an index of 3.413 (85.333), while the table above it shows 3.586 (89.662). Second, the text under the total table gives 88.134, and the table gives 88.166. Third, the education counts (4, 8, and 27) add up to 39, not 38. Fourth, the respondent

roster lists several names twice, four of them identically and one with a different spelling. It cannot be determined whether these are different people with the same name or entries that were recorded twice. If they are duplicates, the sample would be closer to 33 or 34. We use the report's tables and treat the 38 as stated. These points are noted so that the company can verify them before the figures are cited.

C. Item-level results

All 19 items exceed the satisfaction threshold (Table III, Fig. 1). Item means ran from 3.39 to 3.74,

adjusted p values from 5.4×10^{-11} to 1.1×10^{-4} , and d_z from 0.72 to 1.65, which are large effects by conventional benchmarks [35]. The exact sign tests yield consistent results, with the largest adjusted p at 4.4×10^{-4} . H1 is supported for every item. Of the 722 ratings, 385 (53.3%) were "very satisfied", 328 (45.4%) "satisfied", and 9 (1.2%, Wilson 95% CI 0.7% to 2.4%) "less satisfied". No respondent chose "dissatisfied".

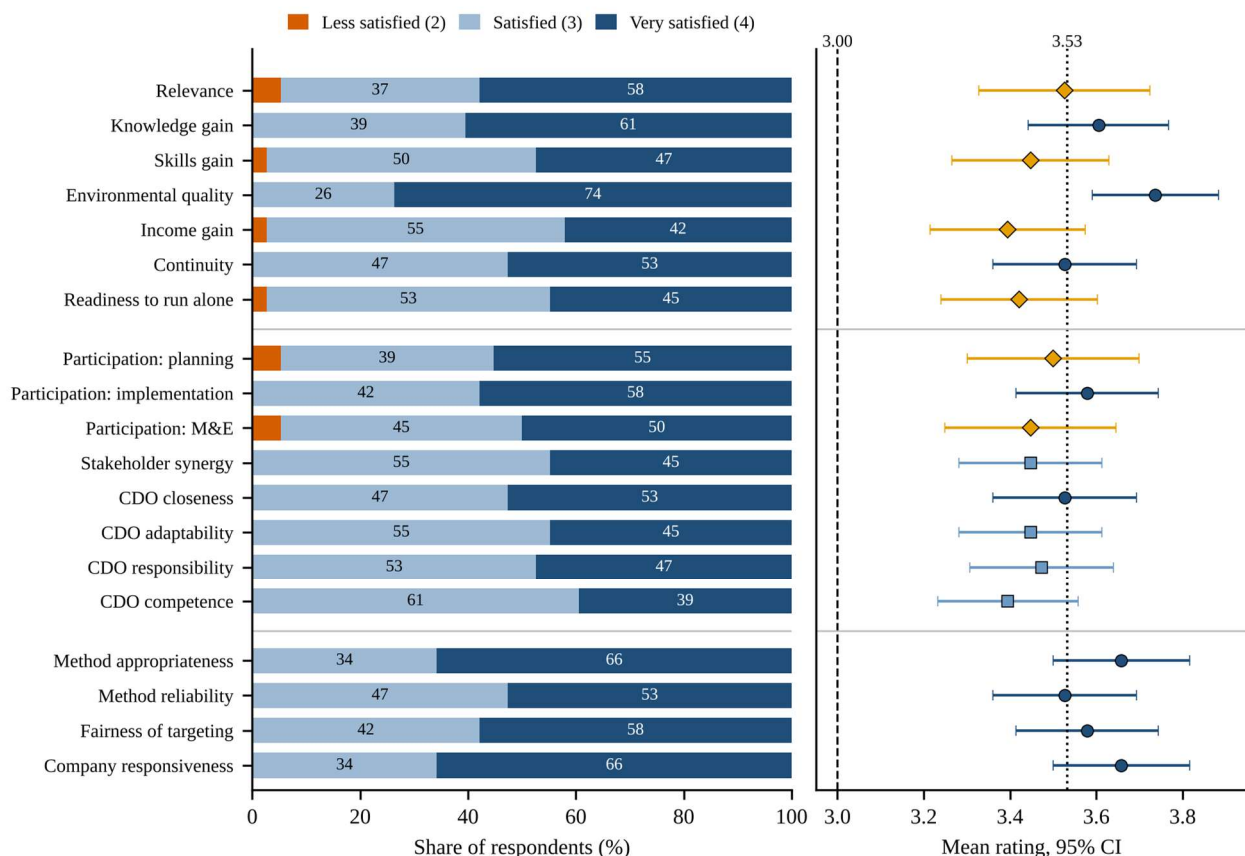


Fig. 1 Rating distribution and mean rating (95% CI) of the 19 survey items (N = 38).

TABLE III
RATING COUNTS AND STATISTICS FOR THE 19 SURVEY ITEMS (N = 38)

Item	2	3	4	Mean	SD	95% CI	Index	d_z	Adj. p (t)	Adj. p (sign)	Group
Overall program											
Program relevance	2	14	22	3.53	0.60	3.33–3.72	88.16	0.87	3.1×10^{-5}	2.5×10^{-4}	C
Knowledge gain	0	15	23	3.61	0.50	3.44–3.77	90.13	1.22	9.0×10^{-8}	3.8×10^{-6}	A
Skills gain	1	19	18	3.45	0.55	3.26–3.63	86.18	0.81	7.8×10^{-5}	3.7×10^{-4}	C
Environmental quality	0	10	28	3.74	0.45	3.59–3.88	93.42	1.65	5.4×10^{-11}	1.4×10^{-7}	A
Income gain	1	21	16	3.39	0.55	3.21–3.57	84.87	0.72	1.1×10^{-4}	4.4×10^{-4}	C
Program continuity	0	18	20	3.53	0.51	3.36–3.69	88.16	1.04	2.3×10^{-6}	2.5×10^{-5}	A
Readiness to run without company	1	20	17	3.42	0.55	3.24–3.60	85.53	0.76	1.1×10^{-4}	4.3×10^{-4}	C

Item	2	3	4	Mean	SD	95% CI	Index	d _z	Adj. p (t)	Adj. p (sign)	Group
Program management											
Participation: planning	2	15	21	3.50	0.60	3.30–3.70	87.50	0.83	6.2×10^{-5}	3.7×10^{-4}	C
Participation: implementation	0	16	22	3.58	0.50	3.41–3.74	89.47	1.16	2.8×10^{-7}	7.2×10^{-6}	A
Participation: M&E	2	17	19	3.45	0.60	3.25–3.65	86.18	0.74	1.1×10^{-4}	4.4×10^{-4}	C
Stakeholder synergy	0	21	17	3.45	0.50	3.28–3.61	86.18	0.89	2.9×10^{-5}	1.4×10^{-4}	B
Facilitator closeness	0	18	20	3.53	0.51	3.36–3.69	88.16	1.04	2.3×10^{-6}	2.5×10^{-5}	A
Facilitator adaptability	0	21	17	3.45	0.50	3.28–3.61	86.18	0.89	2.9×10^{-5}	1.4×10^{-4}	B
Facilitator responsibility	0	20	18	3.47	0.51	3.31–3.64	86.84	0.94	1.3×10^{-5}	7.6×10^{-5}	B
Facilitator competence	0	23	15	3.39	0.50	3.23–3.56	84.87	0.80	7.8×10^{-5}	3.7×10^{-4}	B
Program delivery											
Method appropriateness	0	13	25	3.66	0.48	3.50–3.82	91.45	1.37	6.9×10^{-9}	1.1×10^{-6}	A
Method reliability	0	18	20	3.53	0.51	3.36–3.69	88.16	1.04	2.3×10^{-6}	2.5×10^{-5}	A
Fairness of targeting	0	16	22	3.58	0.50	3.41–3.74	89.47	1.16	2.8×10^{-7}	7.2×10^{-6}	A
Company responsiveness	0	13	25	3.66	0.48	3.50–3.82	91.45	1.37	6.9×10^{-9}	1.1×10^{-6}	A

Note. Columns 2–4: respondents rating the item 2 (less satisfied), 3 (satisfied), or 4 (very satisfied); no respondent rated 1. $d_z = (M - 3)/SD$. Index = $25 \times \text{mean}$. Adj. p: Holm-adjusted over 19 tests against 3.00 (t-test, $df = 37$; exact sign test, 4s against ratings below 3, ties excluded). Group: A sustain, B maintain, C watch (Table V).

The highest means belong to environmental quality (3.74; 74% very satisfied), method appropriateness (3.66), and company responsiveness (3.66). Fairness of targeting and participation in implementation follow at 3.58. All 38 respondents rated the four facilitator items, the method, fairness, responsiveness, knowledge gain, environmental quality, continuity, synergy, and implementation at least “satisfied”; thirteen items in all had no rating below 3. The lowest means are income gain and facilitator competence (both 3.39), followed by skills gain and participation in monitoring and evaluation (both 3.45), and synergy and adaptability (both 3.45). Facilitator competence warrants specific attention. It had no rating below 3, but 61% of respondents chose “satisfied” and only 39% “very satisfied”, the lowest share of top ratings of any item.

The nine ratings below “satisfied” are distributed across six items (Fig. 2): relevance (2), participation in planning (2), participation in monitoring and evaluation (2), skills gain (1), income gain (1), and readiness to run the program without the company (1). Each represents no more than 5.3% of respondents. A Kruskal–Wallis test found no detectable difference among the 19 item distributions ($H = 22.5$, $df = 18$, $p = 0.21$), indicating limited variation; the six items are therefore best interpreted as areas for monitoring rather than as established weaknesses.

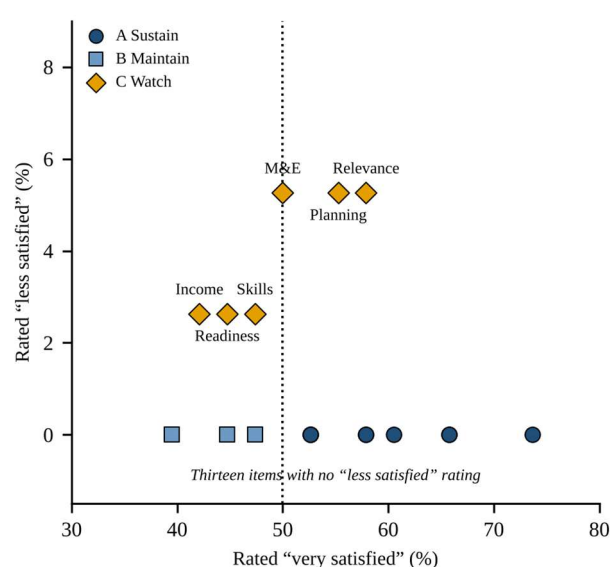


Fig. 2 Share of respondents “very satisfied” against share “less satisfied”, by item.

D. Hypothesis tests and benchmarks

Table IV summarizes the tests. H2 is supported. Environmental quality scored higher than income gain by 0.34 points (adjusted $p = 0.020$), readiness to run the program without the company by 0.32 (0.028), and skills gain by 0.29 (0.035). The knowledge–income contrast (0.21) was not significant (0.095). This pattern is consistent with a program whose most visible outcome is environmental improvement and whose economic outcomes require more time to materialize. Items-as-units contrasts between benefit items (mean 3.55), delivery items (3.61), and facilitator items (3.46) were not significant (exact $p = 0.69$ for both),

so the data do not indicate that delivery is rated higher than outcomes.

TABLE IV
HYPOTHESIS TESTS AND BENCHMARK COMPARISONS

Hyp.	Test and comparison	Result
H1	Item means vs 3.00 (t-test, Holm, 19 tests)	19 of 19 supported; adj. $p = 5.4 \times 10^{-11}$ to 1.1×10^{-4} ; d_x 0.72–1.65
H1	Item 4s vs ratings below 3 (exact sign test, Holm)	19 of 19 supported; largest adj. $p = 4.4 \times 10^{-4}$
H2	Environment vs income (Mann–Whitney, Holm, 4 contrasts)	diff. 0.34; adj. $p = 0.020$
H2	Environment vs readiness to run alone	diff. 0.32; adj. $p = 0.028$
H2	Environment vs skills gain	diff. 0.29; adj. $p = 0.035$
H2	Knowledge vs income (not part of H2)	diff. 0.21; adj. $p = 0.095$
H3	Management index vs 88.31 (2017 grade A)	86.92; $t(7) = -2.75$; adj. $p = 0.057$
H3	19 items vs 88.31 (2017 grade A)	88.02; $t(18) = -0.53$; adj. $p = 0.60$
Bench.	Management index vs 81.26 (2004 grade A)	$t(7) = 11.24$; adj. $p < 0.001$
Bench.	19 items vs 81.26 (2004 grade A)	$t(18) = 12.34$; adj. $p < 0.001$
Bench.	Management index vs 84.75 [26]	$t(7) = 4.31$; adj. $p = 0.011$
Bench.	19 items vs 84.75 [26]	$t(18) = 5.97$; adj. $p < 0.001$
Check	Kruskal–Wallis, 19 item distributions	$H = 22.5$, $df = 18$, $p = 0.21$
Check	Delivery vs benefit items; facilitator vs benefit items (exact Mann–Whitney, items as units)	3.61 vs 3.55, $p = 0.69$; 3.46 vs 3.55, $p = 0.69$

Note. Benchmark and H3 tests treat items as units (one-sample t-tests of item means; $df = \text{items} - 1$) and are Holm-adjusted over six tests; they are indicative. Contrasts for H2 treat ratings as independent, which is conservative, and are exploratory.

H3 can be neither confirmed nor rejected. The 19-item mean converts to 88.02 (95% CI 86.87 to 89.17), which does not differ significantly from 88.31 ($t(18) = -0.53$, $p = 0.60$). The management aspect, at 86.92 (CI 85.73 to 88.12), falls below 88.31 in an unadjusted test ($p = 0.029$), but the Holm-adjusted p is 0.057. Both exceed the 2004 grade A boundary of 81.26 (adjusted $p < 0.001$) and the published CSR benchmark of 84.75 (adjusted $p = 0.011$ for management and < 0.001 for all items). The program therefore lies significantly above the 2004 boundary and the published benchmark and does not differ significantly from the 2017 boundary.

E. Action groups and distance to grade A

Nine items form group A: knowledge gain, environmental quality, continuity, participation in implementation, facilitator closeness, method

appropriateness, method reliability, fairness, and responsiveness. Four items form group B: synergy and the facilitator's adaptability, responsibility, and competence. Six items form group C: relevance, skills, income, readiness, and participation in planning and in monitoring and evaluation. No item reaches group D (Table V). The result is sensitive to the cut-off. A threshold of two respondents moves three items (relevance and the two participation items) into D, and a threshold of one moves all six group C items into group D.

TABLE V
ACTION GROUPS AND SUGGESTED MANAGEMENT RESPONSE

Group	Rule	Items	Suggested response
A Sustain	None below 3; $\geq 50\%$ rate 4	Knowledge; environment; continuity; participation in implementation; CDO closeness; method appropriateness and reliability; fairness; responsiveness	Write down what the CDO and company do (visits, local contacts, targeting) so that these practices are retained when staff change
B Maintain	None below 3; $< 50\%$ rate 4	Stakeholder synergy; CDO adaptability, responsibility, competence	Maintain; offer refresher training for the CDO and local facilitators
C Watch	1–2 respondents below 3	Relevance; skills; income; readiness to run alone; participation in planning and in M&E	Record production, sales, and income by activity; ask each group what it needs to run alone; invite members to planning and review meetings
D Act	≥ 3 respondents below 3 ($\approx 8\%$)	None	Cut-off of 2 would add relevance and both participation items; cut-off of 1 adds all six C items

Note. Cut-offs applied to the reconstructed distributions. M&E = monitoring and evaluation; CDO = community development officer.

Fig. 3 shows where the reported total falls on each scale. The difference in grade is also small in practical terms. Reaching 88.31 across all 19 items needs only nine of the 337 non-top ratings (2.7%) to rise by one step; for the management aspect, 17 of 155 (11%) would be required. The same data may therefore be classified as grade A or grade B within the rounding error of the survey; for this reason, the interpretation scale is examined in detail in the Discussion.

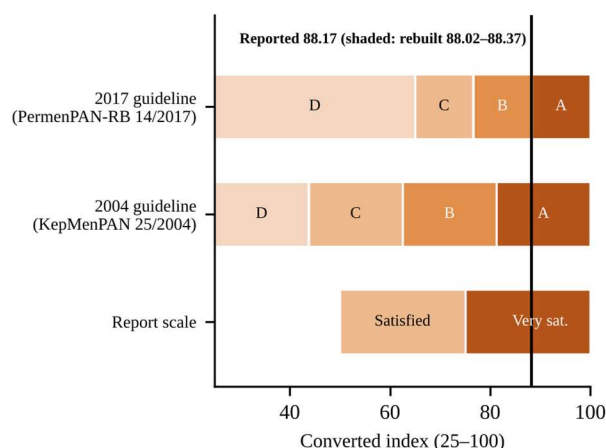


Fig. 3 The reported total (88.17) on three interpretation scales.

IV. DISCUSSION

A. Aspects valued by beneficiaries

The first main finding is that the beneficiaries are highly satisfied. Every item exceeded the threshold with a large effect, and more than half of all ratings were at the top of the scale. The second finding concerns the factors underlying this satisfaction. Environmental quality ranked highest, with 74% of respondents “very satisfied”. In a program built on organic farming, composting, a waste bank, and disaster mitigation, a visible change in the surroundings is the most direct return, and it is the benefit that the report links to every activity of the program. In addition, this benefit is collective, since environmental improvement does not depend on the harvest of individual households.

The delivery items ranked next. Appropriateness of method and responsiveness of the company both scored 3.66, and every respondent rated fairness of targeting at least “satisfied”. This is important because fairness and quality of contact are the elements that Moffat and Zhang [5] place at the center of community acceptance, and because targeting that people see as uneven can erode trust in programs that choose beneficiaries. The facilitator also received favorable ratings. The four facilitator items had no rating below 3, and the report notes that every respondent knew the facilitator well, which supports the reading that regular contact builds trust. At the same time, our contrast tests give no sign that delivery outshines outcomes, so it cannot be concluded that the

company’s mode of delivery distinguishes the program.

B. Areas requiring improvement

The weaknesses are minor but follow a consistent pattern. Income gain is lowest, and skills gain and readiness to run the program without the company score only slightly higher. These three describe what happens after training: whether it turns into earnings, and whether groups could continue alone. Lower ratings for these outcomes are to be expected. Knowledge can rise within a training session, while income depends on harvests, buyers, prices, and a market for the café, homestays, and fertilizer. Livelihood theory supports this interpretation, because a livelihood rests on several assets and activities and not on a single skill [14]. The report supports this interpretation: respondents hope that the company will continue until the program becomes an independent enterprise and ask for more capacity building and human resource training. The beneficiaries’ requests thus correspond closely to the lower-rated items.

Participation shows a second pattern. Involvement in implementation scored high (3.58, no rating below 3), whereas planning and monitoring and evaluation each received two ratings below 3. Participation theory has long held that taking part in doing a task and taking part in deciding what the task is are different things [29, 30], and community-driven development reviews find that the second is harder to secure [31]. Tosun’s [15] account of tourism development in developing countries makes a related point: involvement tends to be easiest in operations and hardest in decisions. Because RANUM ASAHAN contains a homestay and a café, this observation is directly relevant. The report states that the company’s officers visit the program regularly, which may explain why monitoring is rated reasonably well; the two ratings below 3 for planning and for monitoring suggest that some members do not perceive how their views are incorporated into decisions.

Three caveats apply. First, the counts underlying these observations are very small, one or two respondents per item, and a change in a single

rating would alter the group assignments. Second, the data cannot say whether the six items share the same respondents. Third, the report does not say who the few unsatisfied respondents are. For these reasons the action groups should be regarded as a basis for discussion with members rather than as evidence of causes.

C. Dependence of the classification on the interpretation scale

The reported total of 88.17 is “very satisfied” on the report’s own scale and grade A under the 2004 guideline, but it is grade B under the 2017 guideline, 0.14 points below the boundary. The same data therefore yield three different classifications. The statistical result is more modest than any of them: the confidence interval of the index (86.87 to 89.17) includes 88.31, and the aggregation choice alone decides whether the figure falls above or below. The mean of rebuilt aspect indices gives 88.37, grade A, while the mean of 19 items gives 88.02, grade B. This reflects not a deficiency of the program but the effects of sample size, rounding, and the way a four-point scale compresses differences at the top. We recommend that company reports state the guideline and aggregation used, give the confidence interval, and avoid relying on a single grade as the principal result.

Two further features of the data inform the interpretation of the high scores. The share of top ratings (53%) is far above the 15% that psychometric practice treats as a ceiling effect [39]. That convention comes from health questionnaires, and it is used here only as an indicative reference, but it implies that the instrument has little room to show improvement and little sensitivity to differences among satisfied respondents. Satisfaction ratings are also known to be skewed upward [40], and the present setting may reinforce this tendency. Respondents were interviewed face to face, as group representatives, about a program that supports their livelihoods, and the facilitator knows them well. Social desirability can inflate ratings in such circumstances [41], and satisfaction depends on what people expected as well as on what they received [42, 43]. High satisfaction may therefore reflect program quality, modest

expectations, or courtesy toward a company whose continued support respondents hope to retain.

D. Comparison with other studies

Table VI compares the result with Indonesian IKM studies. The score is higher than the public-service studies, which range from 66.49 to 83.46, and similar to a village CSR study at about 87 [27]. Our items-as-units tests place it above a CSR red ginger program at 84.75 (adjusted $p < 0.001$). Such comparisons are indicative only, since the studies differ in scale, items, sample, and sampling, and the public-service studies measure speed, schedules, and counters, not agroecology and livelihoods. The comparison nevertheless indicates why a program-based IKM needs its own items. Whereas service studies identify weaknesses in process attributes, this survey identifies them in outcomes that require time to develop.

TABLE VI
COMPARISON WITH OTHER INDONESIAN IKM STUDIES

Setting	n	Items	Index
Licensing office, Tanjungpinang [23]	150	14	66.49
Public service mall, Banjarbaru [22]	150	9	78.8
e-KTP service, Madiun [24]	100	9	≈ 80.7
Civil registry, Central Lombok [25]	400	14	83.46
CSR red ginger program, Banjarbaru [26]	n/s	10	84.75
Village CSR programs, Kertajaya [27]	18	8 programs	3.48 of 4 (≈ 87)
Mining company program, Gunung Kidul [28]	135	n/s	3.52 of 5 (70.4%)
This study: RANUM ASAHAN, hydropower company, Asahan	38	19	88.17 (88.02 rebuilt)

Note. n/s = not stated. Scales and cut-offs differ across studies; comparison is indicative. Indices on the converted 25–100 scale unless stated.

E. Implications for program management

The results suggest five actions. First, record production, sales, and income for each activity (organic produce, tower café, homestay, fertilizer, and waste bank) so that the next survey can compare ratings of income with figures from group records. Second, ask each group what it needs to run on its own and set a handover plan for each activity, with a local second-line facilitator trained alongside the company’s CDO. Third, hold a planning meeting in each group at least once a year and report back on suggestions, with a target of no respondent below 3 on participation in planning.

Fourth, extend the program to more beneficiaries and build the capacity of the groups as organizations, which are the report's own recommendations. Fifth, maintain the effective practices identified in group A, including regular visits, collaboration with local stakeholders, and fair targeting, and write them down so that they are retained when staff change. For the survey itself, the company should state the interpretation scale and report confidence intervals, publish counts with percentages, and check the roster and totals before release.

F. Strengths and limitations

The strengths of this study are an explicit test against a threshold, hypotheses stated before testing, adjustment for multiple testing, and a validation of the reconstruction against a published index. Several limitations remain. The sample is small and was not drawn at random from a known list, so the intervals describe sampling error only if respondents are treated as a random draw. Counts were reconstructed from rounded percentages, although the unique mapping and the exact match for management make major errors unlikely. The overall and delivery indices could not be reproduced, and the report contains inconsistencies. Items were treated as equally weighted. The contrasts for H2 treat ratings as independent, which is conservative but approximate. A cross-sectional survey of beneficiaries cannot show change over time, and there is no comparison group. Finally, the authors' relationship to the company is a possible source of bias, which we declare below.

V. CONCLUSIONS

This study evaluated the satisfaction of 38 beneficiaries of RANUM ASAHAN, an agroecology-based empowerment program of a hydropower company in Aek Songsongan Village. All 19 items exceeded the satisfaction threshold, with the highest means for environmental quality, method appropriateness, and company responsiveness, and the lowest for income gain and facilitator competence. All nine less-than-satisfied ratings fell on six items about relevance, skills, income, readiness to run the program alone, and

participation in planning and monitoring. The reported total of 88.17 is "very satisfied" and grade A in 2004, but sits just below grade A in 2017, and it does not differ significantly from that boundary. We conclude that the beneficiaries are highly satisfied, and that further improvement depends on converting knowledge into income and in preparing groups to continue without the company. The company should record income by activity, set handover plans, give members a clear role in planning, and report the scale and the uncertainty with the score. Future work should add a longitudinal design, respondent-level data, a comparison group, and independent checks of income.

ACKNOWLEDGMENT

The authors thank the management of PLTA PT Indonesia Asahan Aluminium (Persero) for supporting this evaluation, PT SUCOFINDO (Persero) Cabang Semarang for conducting the survey, and the beneficiaries of RANUM ASAHAN in Aek Songsongan Village who took part.

DECLARATIONS

Ethics statement. This study is a secondary analysis of anonymized, aggregated results from a routine program monitoring-and-evaluation survey. It involved no new contact with respondents and no identifiable personal data, and the names listed in the survey report are not reproduced. Such analysis does not require ethical review.

Informed consent. The survey was administered by PT SUCOFINDO (Persero) for PLTA PT Indonesia Asahan Aluminium (Persero). All respondents gave informed consent to take part in the survey and to the use of the aggregated results for program evaluation and publication. No individual-level data are published.

Conflict of interest. Zainuddin Iqbal Sidabutar and Subandi are employees of PT Indonesia Asahan Aluminium (Persero), the company that runs the RANUM ASAHAN program evaluated here and commissioned the survey. Dewi Hajar declares no competing interests.

Funding. This research received no external funding.

Author contributions (CRediT [44]). Zainuddin Iqbal Sidabutar: conceptualization, methodology, formal analysis, supervision, writing – review and editing. Subandi: investigation, data curation, validation, visualization, project administration. Dewi Hajar: writing – original draft. As corresponding author, Zainuddin Iqbal Sidabutar takes responsibility for the integrity of the data and the accuracy of the analysis.

Data availability. Item-level rating counts and all derived statistics appear in Tables I–V. The survey report and underlying data belong to PT Indonesia Asahan Aluminium (Persero) and are available from the corresponding author on reasonable request, subject to the company's approval.

REFERENCES

- [1] Republic of Indonesia, *Undang-Undang Nomor 40 Tahun 2007 tentang Perseroan Terbatas* [Law No. 40 of 2007 on Limited Liability Companies]. Jakarta, Indonesia, 2007.
- [2] D. Prayogo, "Evaluasi program corporate social responsibility dan community development pada industri tambang dan migas," *Makara Hum. Behav. Stud. Asia*, vol. 15, no. 1, pp. 43–58, 2011, doi: 10.7454/mssh.v15i1.893.
- [3] J. Prno and D. S. Slocumbe, "Exploring the origins of 'social license to operate' in the mining sector: Perspectives from governance and sustainability theories," *Resour. Policy*, vol. 37, no. 3, pp. 346–357, 2012.
- [4] A. B. Carroll, "The pyramid of corporate social responsibility: Toward the moral management of organizational stakeholders," *Bus. Horiz.*, vol. 34, no. 4, pp. 39–48, 1991.
- [5] K. Moffat and A. Zhang, "The paths to social licence to operate: An integrative model explaining community acceptance of mining," *Resour. Policy*, vol. 39, pp. 61–70, 2014.
- [6] C. Wang, "A guide for local benefit sharing in hydropower projects," World Bank, Washington, DC, USA, Social Development Papers No. 128, 2012. [Online]. Available: <http://hdl.handle.net/10986/18366>
- [7] D. Prayogo and Y. Hilarius, "Efektivitas program CSR/CD dalam pengentasan kemiskinan: Studi peran perusahaan geotermal di Jawa Barat," *Masyarakat, J. Sociol.*, vol. 17, no. 1, Art. no. 1, 2012, doi: 10.7454/MJS.v17i1.1189.
- [8] S. Muhammad *et al.*, "Corporate social responsibility programs in mining areas: Insights from stakeholder groups in Indonesia," *Cogent Soc. Sci.*, vol. 10, no. 1, Art. no. 2357675, 2024.
- [9] R. E. Freeman, *Strategic Management: A Stakeholder Approach*. Boston, MA, USA: Pitman, 1984.
- [10] S. Que, L. Wang, K. Awuah-Offei, W. Yang, and H. Jiang, "Corporate social responsibility: Understanding the mining stakeholder with a case study," *Sustainability*, vol. 11, no. 8, Art. no. 2407, 2019, doi: 10.3390/su11082407.
- [11] PT SUCOFINDO (Persero) Cabang Semarang, "Laporan indeks kepuasan masyarakat penerima program pemberdayaan masyarakat PLTA PT Indonesia Asahan Aluminium (Persero) tahun 2026" [Community satisfaction index report], PT SUCOFINDO, Semarang, Indonesia, Sep. 2026, unpublished company report.
- [12] M. A. Altieri, "Agroecology: The science of natural resource management for poor farmers in marginal environments," *Agric. Ecosyst. Environ.*, vol. 93, no. 1–3, pp. 1–24, 2002, doi: 10.1016/S0167-8809(02)00085-3.
- [13] A. Wezel, S. Bellon, T. Doré, C. Francis, D. Vallod, and C. David, "Agroecology as a science, a movement and a practice. A review," *Agron. Sustain. Dev.*, vol. 29, no. 4, pp. 503–515, 2009, doi: 10.1051/agro/2009004.
- [14] I. Scoones, "Livelihoods perspectives and rural development," *J. Peasant Stud.*, vol. 36, no. 1, pp. 171–196, 2009, doi: 10.1080/03066150902820503.
- [15] C. Tosun, "Limits to community participation in the tourism development process in developing countries," *Tour. Manag.*, vol. 21, no. 6, pp. 613–633, 2000, doi: 10.1016/S0261-5177(00)00009-1.
- [16] D. P. Aldrich and M. A. Meyer, "Social capital and community resilience," *Am. Behav. Sci.*, vol. 59, no. 2, pp. 254–269, 2015, doi: 10.1177/0002764214550299.
- [17] S. Cahyati, H. S. P. Siagian, L. Ginting, and H. Shabirah, "Pentingnya kewirausahaan hijau: Edukasi penggunaan pupuk organik cair kulit nanas dan air kelapa," *Abdimas Mandalika*, vol. 4, no. 4, pp. 348–354, 2025, doi: 10.31764/am.v4i4.33379.
- [18] S. S. Wati and M. Alisalman, "Community empowerment in the Probabaya program through hydroponic cultivation in RT 02, Gunung Lingai Village, Samarinda City," *Empowerment: J. Ilm. Prog. Stud. Pendidik. Luar Sekolah*, vol. 14, no. 2, pp. 292–298, 2025, doi: 10.22460/empowerment.v14i2.6450.
- [19] S. Y. Fahreza, D. Triansyah, S. Khumayah, and A. D. Lestari, "Corporate social responsibility through the Berkah Mulya waste bank program," *J. Res. Soc. Sci. Econ. Manag.*, vol. 5, no. 1, pp. 2597–2607, 2025, doi: 10.59141/jrssem.v5i1.970.
- [20] Ministry of State Apparatus Empowerment, *Keputusan Menteri PAN Nomor KEP/25/M.PAN/2/2004 tentang Pedoman Umum Penyusunan Indeks Kepuasan Masyarakat Unit Pelayanan Instansi Pemerintah* [General guidelines for compiling the community satisfaction index of government service units]. Jakarta, Indonesia, 2004.
- [21] Ministry of State Apparatus Empowerment and Bureaucratic Reform, *Peraturan Menteri PANRB Nomor 14 Tahun 2017 tentang Pedoman Penyusunan Survei Kepuasan Masyarakat Unit Penyelenggara Pelayanan Publik*, Berita Negara Republik Indonesia No. 708. Jakarta, Indonesia, 2017.
- [22] R. Subhan, "Indeks kepuasan masyarakat (IKM) pada mal pelayanan publik Kota Banjarbaru tahun 2022," *J. Bisnis Pambang.*, vol. 12, no. 1, pp. 87–94, 2023.
- [23] E. Kartikaningdyah, "Analisis indeks kepuasan masyarakat terhadap pelayanan publik pada BP2T Kota Tanjungpinang," *J. Integrasi*, vol. 4, no. 2, pp. 136–147, 2012.
- [24] N. A. Dinata and D. A. Utami, "Analisis kepuasan masyarakat terhadap pelayanan e-KTP berdasarkan IKM di Kantor Kecamatan Geger Kabupaten Madiun," *Inovant*, vol. 1, no. 2, pp. 312–326, 2023.
- [25] Agus and Ardiyansah, "Community satisfaction level toward public services at the Population and Civil Registration Office of Central Lombok Regency in 2025: Analysis of the community satisfaction index and improvement recommendations," *J. Public Policy Gov. Local Gov.*, vol. 1, no. 1, pp. 57–65, 2026.
- [26] R. Yuliani, Z. Mubarak, Arifah, and Norlena, "Indeks kepuasan masyarakat terhadap program corporate social responsibility (CSR) Bujur Berstamina: Budidaya jahe merah," *JPP: J. Adm. Publik Pambang.*, vol. 5, no. 1, pp. 14–21, 2023.
- [27] W. E. Widodo, S. N. Cholidah, A. P. Isnaeni, K. T. Wibowo, and E. Abriandi, "Mengukur kepuasan masyarakat pada program CSR di Desa Kertajaya: Sebuah analisis compass sustainability method," *J. Pemberdayaan Masy.*, vol. 3, no. 1, pp. 31–54, 2019, doi: 10.14421/jpm.2019.031-02.
- [28] M. Pebriyani, A. A. I. A. Adnyano, and S. A. Rande, "Kajian program PPM community development (COMDEV) di PT. Sugih Alamanugroho di Desa Bedoyo, Kecamatan Ponjong, Kabupaten Gunung Kidul, Daerah Istimewa Yogyakarta," in *Pros. Nas. Rekayasa Teknol. Ind. Inf. XVII (ReTII)*, 2022, pp. 182–188.
- [29] S. R. Arnstein, "A ladder of citizen participation," *J. Am. Inst. Planners*, vol. 35, no. 4, pp. 216–224, 1969.
- [30] A. Cornwall, "Unpacking 'participation': Models, meanings and practices," *Community Dev. J.*, vol. 43, no. 3, pp. 269–283, 2008.

- [31] G. Mansuri and V. Rao, "Community-based and -driven development: A critical review," *World Bank Res. Obs.*, vol. 19, no. 1, pp. 1–39, 2004.
- [32] R. Likert, "A technique for the measurement of attitudes," *Arch. Psychol.*, vol. 22, no. 140, pp. 1–55, 1932.
- [33] G. Norman, "Likert scales, levels of measurement and the 'laws' of statistics," *Adv. Health Sci. Educ.*, vol. 15, no. 5, pp. 625–632, 2010, doi: 10.1007/s10459-010-9222-y.
- [34] D. Lakens, "Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs," *Front. Psychol.*, vol. 4, Art. no. 863, 2013, doi: 10.3389/fpsyg.2013.00863.
- [35] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ, USA: Lawrence Erlbaum Associates, 1988.
- [36] E. B. Wilson, "Probable inference, the law of succession, and statistical inference," *J. Am. Stat. Assoc.*, vol. 22, no. 158, pp. 209–212, 1927.
- [37] L. D. Brown, T. T. Cai, and A. DasGupta, "Interval estimation for a binomial proportion," *Stat. Sci.*, vol. 16, no. 2, pp. 101–133, 2001, doi: 10.1214/ss/1009213286.
- [38] S. Holm, "A simple sequentially rejective multiple test procedure," *Scand. J. Stat.*, vol. 6, no. 2, pp. 65–70, 1979.
- [39] C. B. Terwee *et al.*, "Quality criteria were proposed for measurement properties of health status questionnaires," *J. Clin. Epidemiol.*, vol. 60, no. 1, pp. 34–42, 2007, doi: 10.1016/j.jclinepi.2006.03.012.
- [40] R. A. Peterson and W. R. Wilson, "Measuring customer satisfaction: Fact and artifact," *J. Acad. Mark. Sci.*, vol. 20, no. 1, pp. 61–71, 1992.
- [41] A. J. Nederhof, "Methods of coping with social desirability bias: A review," *Eur. J. Soc. Psychol.*, vol. 15, no. 3, pp. 263–280, 1985, doi: 10.1002/ejsp.2420150303.
- [42] R. L. Oliver, "A cognitive model of the antecedents and consequences of satisfaction decisions," *J. Mark. Res.*, vol. 17, no. 4, pp. 460–469, 1980.
- [43] G. G. Van Ryzin, "Expectations, performance, and citizen satisfaction with urban services," *J. Policy Anal. Manag.*, vol. 23, no. 3, pp. 433–448, 2004.
- [44] A. Brand, L. Allen, M. Altman, M. Hlava, and J. Scott, "Beyond authorship: Attribution, contribution, collaboration, and credit," *Learn. Publ.*, vol. 28, no. 2, pp. 151–155, 2015, doi: 10.1087/20150211.