A.1 Component Metrics

Coverage Complementarity Index (CCI):

$$CCI = \frac{|\{c_i : c_i \in C_{methods}\}|}{|C_{total}|} \tag{1}$$

where $C_{methods}$ represents the set of unique coverage types in the method combination, and $C_{total} = 8$ represents the theoretical maximum coverage types in urban safety research (extensive, intensive, observational, narrative, prospective, collective, sensor, aerial).

Triangulation Validity Coefficient (TVC):

$$TVC = \max\left(0, 1 - \frac{CV_{norm}}{0.5}\right) \tag{2}$$

where the coefficient of variation is defined as:

$$CV_{norm} = \frac{\sigma_{explanatory}}{\mu_{explanatory}} \tag{3}$$

and $\sigma_{explanatory}$ and $\mu_{explanatory}$ are the standard deviation and mean of explanatory power values across methods in the combination (45; 46).

Resource Efficiency Ratio (RER):

$$RER = \frac{\sum_{i=1}^{n} E_i}{\sum_{i=1}^{n} R_i} \tag{4}$$

where E_i represents the explanatory power and R_i the resource requirement for method i in an n-method combination (43).

Sample Adequacy Index (SAI):

$$SAI = \min\left(1.0, \frac{\log_{10}(S_{weighted} + 1)}{3}\right) \tag{5}$$

where the weighted sample size is calculated as:

$$S_{weighted} = \frac{\sum_{i=1}^{n} S_i \cdot E_i}{\sum_{i=1}^{n} E_i} \tag{6}$$

and S_i represents the sample size for method i (47).

Convergent Validity Score (CVS):

$$CVS = \max\left(0, \min\left(1, 0.5 + \sum_j \beta_j \cdot I_j - \sum_k \gamma_k \cdot D_k\right)\right) \tag{7}$$

where I_j are binary indicators for documented convergent method pairs with coefficients β_j, and D_k are binary indicators for documented divergent combinations with penalties γ_k, based on empirical findings from the Lomas del Centinela study.

A.2 Composite Integration Index

The Methodological Integration Index is calculated as a weighted linear combination:

$$MII = w_1 \cdot CCI + w_2 \cdot TVC + w_3 \cdot RER_{norm} + w_4 \cdot SAI + w_5 \cdot CVS \tag{8}$$

where weights are based on mixed methods literature (42):

$$w_1 = 0.25 \quad \text{(Coverage complementarity)} \tag{9}$$
$$w_2 = 0.20 \quad \text{(Triangulation validity)} \tag{10}$$
$$w_3 = 0.20 \quad \text{(Resource efficiency)} \tag{11}$$
$$w_4 = 0.15 \quad \text{(Sample adequacy)} \tag{12}$$
$$w_5 = 0.20 \quad \text{(Convergent validity)} \tag{13}$$

and RER_{norm} is the normalized resource efficiency ratio:

$$RER_{norm} = \min\left(1.0, \frac{RER}{RER_{max}}\right) \tag{14}$$

with $RER_{max} = 3.0$ representing the theoretical maximum efficiency based on the method performance data.

A.3 Empirical Calibration Parameters

The convergent validity coefficients are calibrated based on documented research findings:

$$\beta_{video+walks} = 0.15 \quad \text{(89\% human vs 25\% algorithmic accuracy)} \tag{15}$$

$$\beta_{mapping+survey} = 0.10 \quad \text{(Collective knowledge convergence)} \tag{16}$$

$$\gamma_{gps+mapping} = 0.20 \quad \text{(50.9\% coverage gap divergence)} \tag{17}$$

$$\gamma_{drone+isolation} = 0.15 \quad \text{(Aerial-ground truth divergence)} \tag{18}$$

The MII provides a standardized metric (range: 0-1) for comparing methodological integration effectiveness, with higher values indicating superior complementarity, validity, efficiency, and empirical convergence (48).

A.4 Component Metrics

Coverage Complementarity Index (CCI):

$$CCI = \frac{|\{c_i : c_i \in C_{methods}\}|}{|C_{total}|} \tag{19}$$

where $C_{methods}$ represents the set of unique coverage types in the method combination, and $C_{total} = 8$ represents the theoretical maximum coverage types in urban safety research (extensive, intensive, observational, narrative, prospective, collective, sensor, aerial).

Triangulation Validity Coefficient (TVC):

$$TVC = \max\left(0, 1 - \frac{CV_{norm}}{0.5}\right) \tag{20}$$

where the coefficient of variation is defined as:

$$CV_{norm} = \frac{\sigma_{explanatory}}{\mu_{explanatory}} \tag{21}$$

and $\sigma_{explanatory}$ and $\mu_{explanatory}$ are the standard deviation and mean of explanatory power values across methods in the combination (45; 46).

Resource Efficiency Ratio (RER):

$$RER = \frac{\sum_{i=1}^{n} E_i}{\sum_{i=1}^{n} R_i} \tag{22}$$

where E_i represents the explanatory power and R_i the resource requirement for method i in an n-method combination (43).

Sample Adequacy Index (SAI):

$$SAI = \min\left(1.0, \frac{\log_{10}(S_{weighted} + 1)}{3}\right) \tag{23}$$

where the weighted sample size is calculated as:

$$S_{weighted} = \frac{\sum_{i=1}^{n} S_i \cdot E_i}{\sum_{i=1}^{n} E_i} \tag{24}$$

and S_i represents the sample size for method i (47).

Convergent Validity Score (CVS):

$$CVS = \max\left(0, \min\left(1, 0.5 + \sum_j \beta_j \cdot I_j - \sum_k \gamma_k \cdot D_k\right)\right) \tag{25}$$

where I_j are binary indicators for documented convergent method pairs with coefficients β_j, and D_k are binary indicators for documented divergent combinations with penalties γ_k, based on empirical findings from the Lomas del Centinela study.

A.5 Composite Integration Index

The Methodological Integration Index is calculated as a weighted linear combination:

$$MII = w_1 \cdot CCI + w_2 \cdot TVC + w_3 \cdot RER_{norm} + w_4 \cdot SAI + w_5 \cdot CVS \tag{26}$$

where weights are based on mixed methods literature (42):

$$w_1 = 0.25 \quad \text{(Coverage complementarity)} \tag{27}$$
$$w_2 = 0.20 \quad \text{(Triangulation validity)} \tag{28}$$
$$w_3 = 0.20 \quad \text{(Resource efficiency)} \tag{29}$$
$$w_4 = 0.15 \quad \text{(Sample adequacy)} \tag{30}$$
$$w_5 = 0.20 \quad \text{(Convergent validity)} \tag{31}$$

and RER_{norm} is the normalized resource efficiency ratio:

$$RER_{norm} = \min\left(1.0, \frac{RER}{RER_{max}}\right) \tag{32}$$

with $RER_{max} = 3.0$ representing the theoretical maximum efficiency based on the method performance data.

A.6 Empirical Calibration Parameters

The convergent validity coefficients are calibrated based on documented research findings:

$$\beta_{video+walks} = 0.15 \quad \text{(89\% human vs 25\% algorithmic accuracy)} \tag{33}$$
$$\beta_{mapping+survey} = 0.10 \quad \text{(Collective knowledge convergence)} \tag{34}$$
$$\gamma_{gps+mapping} = 0.20 \quad \text{(50.9\% coverage gap divergence)} \tag{35}$$
$$\gamma_{drone+isolation} = 0.15 \quad \text{(Aerial-ground truth divergence)} \tag{36}$$

The MII provides a standardized metric (range: 0-1) for comparing methodological integration effectiveness, with higher values indicating superior complementarity, validity, efficiency, and empirical convergence (48).

B Method Performance Assessment Framework

Table 1. Method Performance Scoring Rationale

Method	**Resources (0-1)**	**Explanatory Power (0-1)**	**Justification**
Resource Requirements Rationale			
Community Survey	0.25	0.55	Very scalable, distributed during food events, minimal coordination required
Safety Paths App	0.30	0.40	Low initial development time, but adoption and internet connectivity issues reduced effectiveness
Safety Walk-alongs	0.45	0.90	More efficient in group format, moderate coordination, provided unique real-time insights
Children Future Visioning	0.50	0.80	Moderate coordination needed for creative sessions, generated implementable local solutions
Video Analysis	0.60	0.89	Moderate resources: systematic documentation plus computational analysis, achieved 89% accuracy vs 25% algorithmic detection
Collective Mapping	0.65	0.60	High coordination complexity, but missed 50.9% of real-time fear events
Audiovisual Interviews	0.70	0.75	One-on-one intensive approach: 8 interviews = 8 hours, deep personal narratives
Agent-Based Modeling	0.80	0.50	High technical expertise required, predicted 30% reduction in route avoidance scenarios
GPS Sensors & Wearables	0.85	0.25	Days of development time plus technical complexity, but only detected 25% of fear events
Drone Observation	0.90	0.35	Very high cost: equipment, piloting expertise, regulations, limited to accessing "blank zones"

Top 3 Most Efficient Methods: 1. Community Survey: 2.20 insight per resource unit 2. Safety Walk-alongs: 2.00 insight per resource unit 3. Children Future Visioning: 1.60 insight per resource unit

Table 2. Scoring Methodology and Key Research Findings

Scoring Component	Criteria and Evidence
Resource Requirements Weighting	
Time Investment (40%)	Based on reported hours: "8 entrevistas = 8 horas", "días de desarrollo" for GPS sensors
Coordination Complexity (30%)	Implementation challenges: "talleres por coordinación", group vs individual methods
Technical Expertise (20%)	Development requirements: app programming, sensor development, ABM simulation
Equipment Costs (10%)	Hardware needs: GPS devices, drones, cameras vs. paper-based methods
Explanatory Power Weighting	
Detection Accuracy (30%)	Quantitative findings: 25% GPS detection, 89% video analysis accuracy
Unique Insights (25%)	Methods providing "insights únicos que no obtuviste de otros"
Community Utility (25%)	Generated "ideas implementables locales", validated qualitative findings
Policy Relevance (20%)	Intervention potential: 30% route avoidance reduction, spatial correlation analysis
Key Research Calibration Points	
Fear Detection Gap	50.9% of fear signals occurred outside collectively mapped danger zones
Algorithm vs Human	Computer vision: 25% accuracy; Human interpretation: 89% accuracy
Spatial Disconnection	Only 16.9% overlap between fear activations and official crime data
Intervention Potential	Agent-based modeling predicted 30% reduction in female route avoidance
Environmental Correlation	Design features (r=0.73) vs crime density (r=0.34) for danger perception