Appendix

Table A1 – Selected questions and responses related to the information provided on OG activities

Indicator code	Question: How frequently did you receive information about the OG activities?	Average score*	SD
INF1	Through documents reserved for OG members	3.69	1.56
INF2	Through direct contact with other OG members	4.03	1.37
INF3	Through documents also distributed externally to the OG	2.57	1.48
INF4	By attending public events (e.g., seminars, media interviews)	2.84	1.55
INF5	By consulting information available online (e.g., website, blog, social network)	2.80	1.49
INF6	Through tools specifically created for communication among OG partners	3.02	1.64

* Scores: 1 = very low frequency; 5 = very high frequency.
Source: Authors' elaboration based on data from the Italian NRN

Table A2 – Selected questions and responses related to interaction between OG members

Indicator code	Question: How frequently did you interact with the other participants?	Average Score*	SD
INT1	With the project leader	4.27	1.08
INT2	With the innovators	4.05	1.14
INT3	With the partners providing technical and informational support	3.80	1.14
INT4	With the partners testing/adopting the proposed innovation	3.73	1.21
INT5	With the partners facilitating/disseminating the innovation	3.64	1.17
INT6	With the farms targeted by the innovation	3.53	1.23
INT7	With other OG participants	3.71	1.16

* Scores: 1 = very low frequency; 5 = very high frequency.
Source: Authors' elaboration based on data from the Italian NRN

Table A3 – Selected questions and responses related to participation in OG activities

Indicator code	Question: How frequently did you participate in OG activities?	Average score*	SD
PAR1	Through plenary meetings	4.09	1.26
PAR2	Through subgroup meetings	3.79	1.38
PAR3	Through remote work (email and social networks)	4.20	1.17
PAR4	Through remote work with collaborative tools (e.g., online platforms, shared folders or documents)	3.61	1.46
PAR5	Through field visits (farms, laboratories, etc.)	3.46	1.40
PAR6	Through other methods	1.45	1.13

* Scores: 1 = very low frequency; 5 = very high frequency.
Source: Authors' elaboration based on data from the Italian NRN

Table A4 – Selected question and response related to the project outcomes

Indicator code	Question: In summary, how satisfied are you with the following aspects?	Average score*	SD
RES	Results achieved by the OG	3.94	1.07

Scores: 1 = very dissatisfied; 5 = very satisfied.
Source: Authors' elaboration based on data from the Italian NRN

A.1 Evaluation of the reflective measurement models

The evaluation of reflective measurement models involves four main steps: assessment of indicator reliability, internal consistency reliability, convergent validity, and discriminant validity. The seminr package, version 2.3.7, in R 4.5.2, was used for the evaluation of reflective measurement models.

A.1.1 Indicator reliability

Indicator reliability is assessed by examining the variance explained by each indicator's construct. Loadings ≥ 0.708 indicate that the construct explains more than 50% of the indicator variance. Indicators with loadings between 0.40 and 0.708 are retained if their removal decreases internal consistency or convergent validity below acceptable thresholds. Indicators with loadings < 0.40 are removed to preserve model integrity.

A.1.2 Internal consistency reliability

Internal consistency reliability is evaluated using Cronbach's α, the exact reliability coefficient (ρ_A) (Dijkstra, 2014; Dijkstra & Henseler, 2015), and composite reliability (ρ_C) (Jöreskog, 1971). Cronbach's α and ρ_C represent the lower and upper bounds of reliability, respectively, while ρ_A falls between them. Values between 0.70 and 0.90 indicate satisfactory reliability, whereas values above 0.90 suggest redundancy and reduced construct validity.

A.1.3 Convergent validity

Convergent validity is assessed via the Average Variance Extracted (AVE), which measures the proportion of variance explained by the indicators of a construct. An AVE ≥ 0.50 indicates that the construct explains at least 50% of the variance in its indicators.

A.1.4 Discriminant validity

Discriminant validity is examined using the heterotrait–monotrait ratio (HTMT), which compares the correlations between different constructs with those within the same construct (Henseler et al., 2015). A threshold of 0.90 (or 0.85 if the constructs are conceptually distinct) is applied. One-sided bootstrap confidence intervals (CI), using the percentile method and 10,000 samples, are computed. Discriminant validity is supported when the upper bound of the 95% CI is below 0.90 (or 0.85).

Table A5 – Internal consistency reliability and convergent validity of constructs in the measurement model

Iteration / constructs	Cronbach's α	Composite reliability (ρ_C)	Exact reliability (ρ_A)	AVE
Iteration 1				
Information	0.822	0.871	0.843	0.534
Interaction	0.887	0.912	0.889	0.596
Participation	0.764	0.838	0.812	0.481
Iteration 2				
Information	0.822	0.871	0.843	0.534
Interaction	0.887	0.912	0.889	0.596
Participation	0.809	0.867	0.815	0.568

Note: The metrics concerning the outcomes construct are always 1, as they are measured with a single item, with $\rho_A \equiv 1$.

Table A6 – Indicator loadings by construct

Construct	Code	Name	Loading
Information	INF1	Frequency of receiving information through documents reserved for OG members	0.583
	INF2	Frequency of receiving information through direct contact with other OG members	0.672
	INF3	Frequency of receiving information through documents also distributed externally to the OG	0.820
	INF4	Frequency of receiving information by attending public events (e.g., seminars, media interviews)	0.814
	INF5	Frequency of receiving information by consulting information available online (e.g., website, blog, social network)	0.807
	INF6	Frequency of receiving information through tools specifically created for communication among OG partners	0.652
Interaction	INT1	Frequency of interactions with the project leader	0.702
	INT2	Frequency of interactions with the innovators	0.801
	INT3	Frequency of interactions with the partners providing technical and informational support	0.779
	INT4	Frequency of interactions with the partners testing/adopting the proposed innovation	0.815
	INT5	Frequency of interactions with the partners facilitating/disseminating the innovation	0.785
	INT6	Frequency of interactions with the farms targeted by the innovation	0.762
	INT7	Frequency of interactions with other OG participants	0.754
Participation	PAR1	Frequency of participation in OG activities through plenary meetings	0.735
	PAR2	Frequency of participation in OG activities through subgroup meetings	0.774
	PAR3	Frequency of participation in OG activities through remote work (email and social networks)	0.808
	PAR4	Frequency of participation in OG activities through remote work with collaborative tools (e.g., online platforms, shared folders or documents)	0.710
	PAR5	Frequency of participation in OG activities through field visits (farms, laboratories, etc.)	0.721
	PAR6	Frequency of participation in OG activities through other methods	0.262*

* Deleted because it is below 0.4.

Table A7 – Bootstrapped results and confidence intervals for the heterotrait–monotrait ratio

	Original estimate	Bootstrap Mean	Bootstrap SD	T Stat.	5% CI	95% CI
Information → Interaction	0.489	0.488	0.051	9.529	0.403	0.570
Information → Participation	0.582	0.581	0.050	11.639	0.496	0.659
Information → Outcomes	0.363	0.363	0.048	7.623	0.284	0.442
Interaction → Participation	0.752	0.751	0.033	22.817	0.695	0.803
Interaction → Outcomes	0.488	0.488	0.049	10.029	0.405	0.565
Participation → Outcomes	0.536	0.536	0.048	11.142	0.454	0.614

A.2 Evaluation of the structural model

The structural model evaluation follows three main steps: assessment of collinearity, path coefficients and effect sizes, and explanatory and predictive power of the model. The seminr package, version 2.3.7, in R 4.5.2, was used for structural model evaluation.

A.2.1 Collinearity assessment

Collinearity among predictor constructs is tested via the Variance Inflation Factor (VIF). Values above 5 indicate potential multicollinearity issues.

A.2.2 Path coefficients and effect sizes

Path coefficients are estimated**,** and their significance is tested through bootstrapping (95% CI). Coefficients are significant when zero is not within the interval. Effect sizes (f^2) are calculated to measure the impact of removing each predictor on the endogenous construct's R^2. Values < 0.02, 0.02–0.15, 0.15–0.35, and > 0.35 represent no, small, medium, and large effects, respectively (Cohen, 1988). Consistency between f^2 values and path coefficient magnitudes confirms the robustness of the results.

A.2.3 Explanatory and predictive power

The explanatory power of the model is assessed using R^2 and adjusted R^2 for endogenous constructs. Values ≥ 0.75 are considered substantial, 0.50–0.75 moderate, and 0.25–0.50 weak. Predictive power is evaluated using the PLSpredict procedure (Shmueli et al., 2016), a case-wise, out-of-sample prediction method implemented in PLS-SEM software (seminr package in R). The procedure relies on k-fold cross-validation, where the model is estimated on the training folds and then used to generate predictions for the holdout fold. The minimum sample size per fold is determined using the inverse square root method (Kock & Hadaya, 2018). Prediction errors are assessed using either the Root Mean Square Error (RMSE) or the Mean Absolute Error (MAE). If the error distribution is symmetric, RMSE is used. Lower RMSE/MAE values in PLS-SEM compared with naïve linear regression benchmarks indicate high predictive power.

Table A8 – Variance inflation factor values for the structural model

Construct	Code	Name	Values
Information	INF1	Frequency of receiving information through documents reserved for OG members	1.482
	INF2	Frequency of receiving information through direct contact with other OG members	1.601
	INF3	Frequency of receiving information through documents also distributed externally to the OG	2.053
	INF4	Frequency of receiving information by attending public events (e.g., seminars, media interviews)	2.274
	INF5	Frequency of receiving information by consulting information available online (e.g., website, blog, social network)	2.053
	INF6	Frequency of receiving information through tools specifically created for communication among OG partners	1.388
Interaction	INT1	Frequency of interactions with the project leader	1.660
	INT2	Frequency of interactions with the innovators	2.405
	INT3	Frequency of interactions with the partners providing technical and informational support	2.126
	INT4	Frequency of interactions with the partners testing/adopting the proposed innovation	2.262
	INT5	Frequency of interactions with the partners facilitating/disseminating the innovation	1.893
	INT6	Frequency of interactions with the farms targeted by the innovation	1.985
	INT7	Frequency of interactions with other OG participants	1.860
Participation	PAR1	Frequency of participation in OG activities through plenary meetings	1.650
	PAR2	Frequency of participation in OG activities through subgroup meetings	1.778
	PAR3	Frequency of participation in OG activities through remote work (email and social networks)	1.998
	PAR4	Frequency of participation in OG activities through remote work with collaborative tools (e.g., online platforms, shared folders or documents)	1.677
	PAR5	Frequency of participation in OG activities through field visits (farms, laboratories, etc.)	1.492

A.3 Multi-group analysis and measurement invariance

Multi-group analysis (MGA) is employed to examine whether structural relationships differ across professional roles. As a prerequisite for meaningful group comparisons, measurement invariance is first assessed using the Measurement Invariance of Composite Models (MICOM) procedure, followed by the evaluation of multi-group differences. All analyses are conducted using the cSEM package, version 0.6.1, in R 4.5.2.

A.3.1 Measurement invariance of composite models

Measurement invariance is assessed to ensure that constructs are interpreted consistently across groups (Henseler et al., 2016). The MICOM procedure consists of three sequential steps: (1) configural invariance, which verifies that the basic model specification is identical across groups; (2) compositional invariance, assessed via permutation testing to evaluate the equivalence of composite scores; and (3) equality of composite mean values and variances, which examines potential differences in latent variable distributions across groups. Partial measurement invariance, established through Steps 1 and 2, is considered sufficient to allow meaningful comparisons of structural relationships across groups, whereas Step 3 is required only when comparing construct means.

A.3.2 Multi-group differences

Once partial measurement invariance is established, differences in path coefficients across professional roles are examined. To enhance the robustness of the results and account for the different statistical assumptions underlying alternative MGA approaches, multiple procedures are employed. Specifically, the analysis combines the Sarstedt et al. (2011) omnibus test, the parametric approaches proposed by Keil et al. (2000) and Nitzl (2010), which are based on t-test distributions, and the permutation-based test by Chin & Dibbern (2010). The latter is considered particularly conservative, as it does not rely on distributional assumptions. Differences in path coefficients are regarded as statistically significant only when consistent evidence is obtained across multiple MGA procedures, reflecting a conservative assessment strategy.

Table A9 – MICOM Step 2: Compositional invariance assessment

Comparison	Construct	Composite Correlation (c)	*p*-value	Invariance established
Other vs Research Institutes	Information	0.868	0.660	Yes
Other vs Research Institutes	Interaction	0.794	0.142	Yes
Other vs Research Institutes	Participation	0.895	0.248	Yes
Other vs Research Institutes	Outcomes	1	1	Yes
Other vs Farms	Information	0.791	0.525	Yes
Other vs Farms	Interaction	0.757	0.177	Yes
Other vs Farms	Participation	0.945	0.711	Yes
Other vs Farms	Outcomes	1	1	Yes
Research Institutes vs Farms	Information	0.710	0.261	Yes
Research Institutes vs Farms	Interaction	0.920	0.802	Yes
Research Institutes vs Farms	Participation	0.936	0.599	Yes
Research Institutes vs Farms	Outcomes	1	1	Yes

Note: Results of the permutation-based assessment of the correlations between composite scores across groups. Compositional invariance is established when the *p*-value is greater than 0.05.

Table A10 – MICOM Step 3: Assessment of equality of composite mean values and variances

Comparison	Construct	Parameter	Statistic	*p*-value	Equality
Other vs Research Institutes	Information	Mean	-0.094	0.355	Yes
Other vs Research Institutes	Interaction	Mean	-0.073	0.475	Yes
Other vs Research Institutes	Participation	Mean	-0.206	0.043	No
Other vs Research Institutes	Outcomes	Mean	-0.196	0.059	Yes
Other vs Research Institutes	Information	Variance	-0.041	0.710	Yes
Other vs Research Institutes	Interaction	Variance	0.225	0.194	Yes
Other vs Research Institutes	Participation	Variance	0.245	0.152	Yes
Other vs Research Institutes	Outcomes	Variance	0.234	0.195	Yes
Other vs Farms	Information	Mean	-0.009	0.932	Yes
Other vs Farms	Interaction	Mean	0.026	0.830	Yes
Other vs Farms	Participation	Mean	-0.100	0.386	Yes
Other vs Farms	Outcomes	Mean	-0.067	0.581	Yes
Other vs Farms	Information	Variance	-0.059	0.638	Yes
Other vs Farms	Interaction	Variance	0.126	0.534	Yes
Other vs Farms	Participation	Variance	0.006	0.977	Yes
Other vs Farms	Outcomes	Variance	0.063	0.760	Yes
Research Institutes vs Farms	Information	Mean	0.086	0.440	Yes
Research Institutes vs Farms	Interaction	Mean	0.099	0.378	Yes
Research Institutes vs Farms	Participation	Mean	0.106	0.342	Yes
Research Institutes vs Farms	Outcomes	Mean	0.129	0.257	Yes
Research Institutes vs Farms	Information	Variance	-0.018	0.877	Yes
Research Institutes vs Farms	Interaction	Variance	-0.099	0.606	Yes
Research Institutes vs Farms	Participation	Variance	-0.240	0.202	Yes
Research Institutes vs Farms	Outcomes	Variance	-0.171	0.395	Yes

Note: Evaluation of the differences in latent variable means and the ratios of variances. Full measurement invariance is achieved when both mean and variance differences are non-significant (*p*-value > 0.05).