



Multi-Robot Coordination for Warehouse Automation Using Artificial Intelligence

Author 1: Basu Valmiki Arthur Cotton

Author 2: Dr.B. Girimurugan

**Business School, Koneru Lakshmaiah Education Foundation,
Vaddeswaram, A.P. – 522302**

Abstract

Online retail and smart manufacturing keep growing, and warehouses are expected to ship more, faster, with fewer errors. AI could help autonomous robots communicate, decide and adapt when the floor changes under them. Earlier research on autonomous mobile robots, Industry 5.0 and machine learning has mostly stayed with single robots. Groups of robots working together get far less attention. We wanted to know how practitioners see AI-based multi-robot coordination, so we put 15 statements to them on a five-point Likert scale. The dataset has 150 records, 149 of them analysable: 23 original questionnaire responses (22 usable) and 127 labelled synthetic sample data. Our methods were descriptive statistics, Cronbach's alpha, one-sample t-tests, Pearson correlation, chi-square tests, one-way ANOVA with Tukey post hoc tests and multiple linear regression.

Respondents rated AI-enabled coordination ($M = 2.76$, $SD = 0.61$) and coordination challenges ($M = 2.79$, $SD = 0.80$) slightly below the neutral score of 3 in the combined data. AI-enabled coordination did not significantly predict perceived warehouse performance ($R^2 = 0.008$, $p = .286$), so our main hypothesis was not supported. Only one group difference appeared: IT/AI professionals reported more coordination challenges than warehouse employees ($F(4, 144) = 3.71$, $p = .007$, $\eta^2 = .09$). The 22 usable original responses, taken alone, told a much stronger story ($R^2 = 0.59$). Still, that is a very small group, and several of its members gave the same rating to every statement. Because most records were synthetic, we treat all of this as exploratory.

Keywords: multi-robot coordination; warehouse automation; artificial intelligence; autonomous mobile robots; Likert survey; regression analysis

1. Introduction

Warehouses today have to be efficient and flexible at the same time, largely because of online shopping and complicated supply chains. If a site depends on manual labour or one robot, a sudden jump in demand or an unexpected fault can quickly cause trouble. So operators are

looking at AI, Autonomous Mobile Robots (AMRs), the Internet of Things (IoT) and cyber-physical systems.

Automation has changed how goods are stored, moved and managed. Automated Storage and Retrieval Systems (AS/RS), Warehouse Management Systems (WMS), IoT devices, cloud computing and AI all offer gains in speed, accuracy and productivity. Several robots working at once is a different, and much harder, problem than a single one, however. Many current approaches still polish each robot separately. They say little about how a group should cooperate once conditions start to shift.

AI gives robots the ability to learn from data, notice what is happening around them and coordinate with each other. Machine learning and intelligent decision-making can handle task allocation, route planning and resource management, and they speed up reactions to surprises. A robot doesn't have to follow a fixed script any more. It can change what it does as the warehouse changes.

That is why coordination among robots matters. Robots that cooperate finish work sooner, stay out of each other's way, and tend to be more accurate and flexible. We look at how AI could support this and at what makes adoption difficult for warehouses.

2. Literature Review

Guohua Wu (2026) observes that Robotic Mobile Fulfillment Systems (RMFS) have become common in e-commerce warehouses. The robots bring storage pods to human pickers, who take out the items they need, and picking can speed up. Even so, results depend on how comfortable, tired and safe workers feel near the robots. Earlier order-picking research mostly assumed workers walking to the products. Grosse et al. split human factors into five areas: perceptual, mental, physical, psychosocial and work environment. Physical problems get the most attention (fatigue, repeated bending, lifting, risk of injury) because they hit both worker health and output.

Xiang He (2022) asks how Industry 5.0 technologies such as robotics, cyber-physical systems and intelligent decision-making are changing warehouse management. Warehouses that used manual labour or individual robots had trouble reacting quickly to changes in demand. Researchers have started to favour collaborative systems, in which several intelligent agents share tasks and work together more efficiently.

Miguel Rodríguez García (2026), modern warehouses are intelligent systems and not merely storage space. AS/RS, AMRs, WMS, IoT, cloud computing and AI can raise accuracy, productivity and efficiency. More reliance on technology also means new risks. Researchers and managers therefore have to set possible system failures and other technology-related disruptions against the benefits of automation.

Earlier studies suggest automation can lower operating costs, shorten delivery lead times and improve customer service. Autonomous robots are a good fit for repetitive storage and retrieval work, since they do it consistently. A lot of systems, though, are designed around how each robot performs alone. When a group of robots meets a new request or an unexpected change, that design can leave the warehouse less flexible than it needs to be.

Matt Danford (2020) writes about AI in manufacturing robots. Different tasks, he points out, need different kinds and levels of intelligence, and manufacturing demands ac

curacy and consistency. Micropsi Industries' MIRAI system, which uses imitation learning, is one example. A human demonstrates the task several times while cameras and sensors record it. The robot then learns what the task is meant to achieve and doesn't just copy one fixed path. It can cope when objects are moved, and it may need less manual programming.

Yossi Sheffi (2026) notes that most earlier research on warehouse automation concerned efficiency, inventory visibility, real-time decisions and logistics performance. Other reviews follow how warehouse systems have evolved, such as Tubis and Rohman (2023), who looked at IoT, RFID, autonomous vehicles and augmented reality. These studies show what digital technology can do. They say less about the risks that appear when operations become heavily automated.

Nathan Hupp (2024) describes the growing role of Computational Fluid Dynamics (CFD) in engineering, where designs can be tested and refined through simulation before any physical model is built. The article Certifying CFD reports that the U.S. Environmental Protection Agency (EPA) has started accepting CFD simulations for certain product certifications. Virtual testing can clearly replace costly physical experiments such as wind tunnel tests.

Gokcen Bas (2020) discusses Smart Truck, which used NASA's FUN3D software and computing resources from Oak Ridge National Laboratory to study the aerodynamics of tractor-trailer attachments. In their simulations, aerodynamic drag fell by 6.6% and fuel efficiency rose by an estimated 4–5%. Simulation, the case suggests, can support product design while saving time and money.

Peter Zelinski (2019) studies how AI and automation are changing work processes and workforce planning. AI-based Human-Robot Collaboration (AI-HRC) matters particularly in intralogistics, where warehouses and distribution centres rely more and more on intelligent systems. Breaking work into individual tasks, the research suggests, helps organisations plan their workforce. When you look at tasks rather than job titles, it is easier to see what can be automated and what still requires human skills.

A.Saruhasan (2026) presents robotics and AI as key technologies for raising productivity, reducing manual effort and helping machines with complex tasks. Research indicates that AI allows robots to learn and respond to changing conditions instead of following fixed instructions. LeCun, Bengio, and Hinton (2015) explain how deep learning helps machines recognise patterns and interpret images. Kollar, Tellex, and Roy (2018) studied how robots can link language to physical actions, which makes human-robot communication feel more natural.

3. Research Gap

Researchers have written a good deal on warehouse automation, AI, AMRs and Industry 5.0. Most of that work deals with individual robots, specific technologies or overall efficiency. Few studies ask what happens when many AI-enabled robots must make decisions and coordinate at once in a busy warehouse, and that is the gap we try to fill.

4. Objectives of the Study

- To analyse the role of Artificial Intelligence in improving multi-robot coordination in warehouse automation systems.
- To examine the challenges of coordinating multiple autonomous robots that operate at the same time in dynamic warehouse environments.
- To evaluate how AI-based decision-making, machine learning and intelligent algorithms improve robot collaboration, task allocation and operational efficiency.
- To investigate how multi-robot coordination affects warehouse performance, including productivity, accuracy, flexibility and response to real-time operational changes.
- To identify the limitations of existing warehouse automation approaches that focus mainly on optimising individual robots rather than collaborative multi-robot systems.

5. Research Questions

The following research questions guide the study:

RQ1. How does Artificial Intelligence improve coordination among multiple autonomous robots in warehouse automation systems?

RQ2. What are the major challenges involved in managing and coordinating multiple AI-enabled robots in dynamic warehouse environments?

RQ3. Does perceived AI-enabled coordination predict perceived improvements in warehouse performance (productivity, accuracy and flexibility)?

RQ4. Do perceptions of AI-enabled coordination, coordination challenges and the limitations of existing systems differ across respondent groups (for example, occupation, organisation type or familiarity with AI)?

6. Hypotheses of the Study

We expected that AI-based decision-making, machine learning and intelligent algorithms would help multiple robots work together better in warehouses. Better coordination might let robots share tasks, use resources well, avoid delays and react to change. We also wanted to know whether systems coordinating several robots can overcome the limits of approaches centred on individual robot performance.

To test this general expectation empirically, four statistical hypotheses were set up ($\alpha = .05$):

- H1: Mean agreement with statements on AI-enabled coordination (AIC) is significantly different from the neutral midpoint of 3 (one-sample t-test). H0: $\mu = 3$.
- H2: Perceived AI-enabled coordination significantly and positively predicts perceived warehouse performance impact (linear regression). H0: $\beta = 0$.
- H3: Perceptions (AIC, CHL, PERF, LIM) differ significantly across occupational, organisational, educational, age, experience and familiarity groups (one-way ANOVA). H0: all group means are equal.

- H4: Respondent characteristics (such as occupation, familiarity with AI, and current robot adoption) are significantly associated with one another (chi-square test of independence). H0: the variables are independent.

7. Research Methodology

7.1 Research design

This is a quantitative, cross-sectional survey. Each participant filled in a structured questionnaire once. It asked about their views on AI-based multi-robot coordination, the challenges of coordination, the possible effects on warehouse performance and the limitations of current automation systems.

7.2 Instrument

The questionnaire opened with eight background questions: age, gender, education, occupation, work experience, organisation type, familiarity with AI and robotics, and current use of warehouse robots. Next came fifteen statements rated from 1 (Strongly Disagree) to 5 (Strongly Agree), coded according to the codebook in Appendix A. Following the study objectives, we grouped the statements into four sets:

- AI-Enabled Coordination (AIC): Q1, Q2, Q3, Q7, Q8, Q9 – the role of AI, machine learning and intelligent algorithms in coordination, task allocation and collaboration (Objectives 1 and 3).
- Coordination Challenges (CHL): Q4, Q5, Q6 – communication, dynamic conditions and safety (Objective 2).
- Warehouse Performance Impact (PERF): Q10, Q11, Q12 – productivity, accuracy and flexibility (Objective 4).
- Limitations of Existing Systems (LIM): Q13, Q14, Q15 – the focus on individual robots and the need for improved coordination (Objective 5).

A respondent's construct score is just the mean of the items they answered for that construct.

7.3 Sample and data provenance

The dataset contains 150 records. According to the 'Data Status' field, 23 are 'Original response' and 127 are 'Synthetic sample data'. We ran the main analysis on all of them, then analysed the original responses separately to see whether anything changed (Section 8.8). The combined dataset should not be mistaken for 150 independently surveyed people. All 150 records list a male respondent, so gender differences could not be examined. We neither used nor reported the names in the raw data.

7.4 Data screening and missing data

Of the answers to the fifteen attitude questions, 22 were missing, and all of those were in original responses. One respondent left all fifteen statements blank and also skipped the AI-familiarity and robot-use questions, so we excluded that record from the analyses of the survey statements.

That left 149 usable cases. The other seven missing answers were handled with the available data: pairwise deletion for item-level analyses, and construct scores calculated from the answered items whenever at least 12 of the 15 statements were present. Analyses involving AI familiarity or robot use therefore used $n = 149$.

7.5 Statistical tools and procedures

We ran the analysis in Python, mainly with SciPy and statsmodels. The accompanying Excel workbook contains formulas for the descriptive statistics, correlations and simple regression, and Appendix B gives equivalent SPSS syntax for anyone who wants to repeat the analyses. The methods were as follows:

- Descriptive statistics: mean, median, standard deviation, minimum, maximum and percentage agreement for every item and construct.
- Reliability: Cronbach's alpha for each construct.
- One-sample t-tests of construct means against the neutral midpoint of 3 (with Wilcoxon signed-rank tests as a non-parametric check).
- Pearson correlation among constructs.
- Chi-square tests of independence with Cramér's V as the effect size; the percentage of cells with an expected count below 5 is reported because small expected counts weaken the chi-square approximation.
- One-way ANOVA with eta-squared (η^2) effect sizes, Levene's test for homogeneity of variance, Kruskal–Wallis tests as a robustness check, and Tukey HSD post hoc comparisons.
- Linear regression (OLS): a simple model (PERF on AIC), a multiple model adding CHL and LIM, and an extended model adding familiarity, experience and education as controls. Assumptions were checked with variance inflation factors, the Durbin–Watson statistic and a Shapiro–Wilk test of residuals.

8. Data Analysis and Results

8.1 Respondent profile

Table 1. Respondent profile (full dataset, N = 150)

Variable	Category	n	%	n (original only)
Age group	Below 25 years	56	37.3	21
	25–35 years	50	33.3	1
	36–45 years	44	29.3	1
Gender	Male	150	100.0	23
	Female	0	0.0	0
Education	Diploma	44	29.3	2
	Undergraduate	52	34.7	18

Variable	Category	n	%	n (original only)
	Postgraduate	21	14.0	2
	Doctorate	33	22.0	1
Occupation	Warehouse Employee	42	28.0	6
	Warehouse Manager/Supervisor	22	14.7	4
	IT/AI Professional	17	11.3	1
	Researcher/Academic	29	19.3	5
	Other	40	26.7	7
Experience	Less than 1 year	63	42.0	17
	1–5 years	39	26.0	3
	6–10 years	48	32.0	3
Organisation type	Manufacturing Industry	34	22.7	6
	Logistics and Supply Chain	27	18.0	4
	Robotics and Automation Company	33	22.0	3
	Research/Education Institution	30	20.0	5
	Other	26	17.3	5
AI familiarity	Beginner	36	24.0	11
	Basic Knowledge	24	16.0	3
	Intermediate	25	16.7	5
	Advanced	39	26.0	2
	Expert	25	16.7	1
Uses warehouse robots	Yes	34	22.7	10
	No	38	25.3	3
	Planning to implement	39	26.0	6
	Not sure	38	25.3	3

Note. Familiarity and robot-use percentages are based on N = 150, although one respondent did not answer these items (149 valid).

The sample covered a range of occupations and organisation types. Age groups were fairly even (37.3% under 25, 33.3% aged 25–35, 29.3% aged 36–45). Warehouse employees were the largest occupational group at 28.0%, followed by 'Other' roles at 26.7% and researchers or academics at 19.3%. Manufacturing made up 22.7% of records, robotics and automation companies 22.0% and logistics 18.0%. Many respondents were short on experience: 42.0% had under one year, and none had more than ten. Postgraduate or doctoral qualifications were held

by 36.0%. When asked about robot use, 22.7% said they used robots, 25.3% said they did not, 26.0% planned to bring them in and 25.3% weren't sure.

8.2 Descriptive Statistics: Item level

Table 2. Item-level descriptive statistics (full dataset)

Item	n	Mean	Median	SD	% Disagree	% Neutral	% Agree
Q1. AI improves coordination	149	3.01	3.0	1.50	40.3	18.1	41.6
Q2. AI enables communication/collaboration	149	2.98	3.0	1.45	39.6	18.8	41.6
Q3. AI reduces conflicts and delays	149	2.58	3.0	1.07	45.0	30.9	24.2
Q4. Communication is a major challenge	147	2.30	2.0	1.02	55.8	30.6	13.6
Q5. Dynamic conditions create difficulty	148	3.12	3.0	1.32	31.1	25.7	43.2
Q6. Safety/collision avoidance is a challenge	149	2.92	3.0	1.50	41.6	18.1	40.3
Q7. AI improves task allocation	149	2.43	2.0	1.14	55.7	18.8	25.5
Q8. ML helps robots adapt over time	147	2.77	3.0	1.39	46.3	21.1	32.7
Q9. Intelligent algorithms aid collaboration	148	2.81	3.0	1.38	43.2	23.6	33.1
Q10. Coordination raises productivity	149	2.95	3.0	1.35	39.6	21.5	38.9
Q11. Coordination improves accuracy	148	2.95	3.0	1.32	34.5	29.1	36.5
Q12. Multi-robot systems improve flexibility	149	2.97	3.0	1.42	42.3	17.4	40.3
Q13. Existing systems focus on individual robots	149	3.17	3.0	1.41	29.5	23.5	47.0
Q14. Automation needs better AI coordination	149	3.13	3.0	1.39	32.2	26.2	41.6
Q15. Lack of collaboration reduces efficiency	149	2.82	3.0	1.40	46.3	16.8	36.9

Note. Scale: 1 = Strongly Disagree, 5 = Strongly Agree. % Disagree = ratings of 1–2; % Agree = ratings of 4–5.

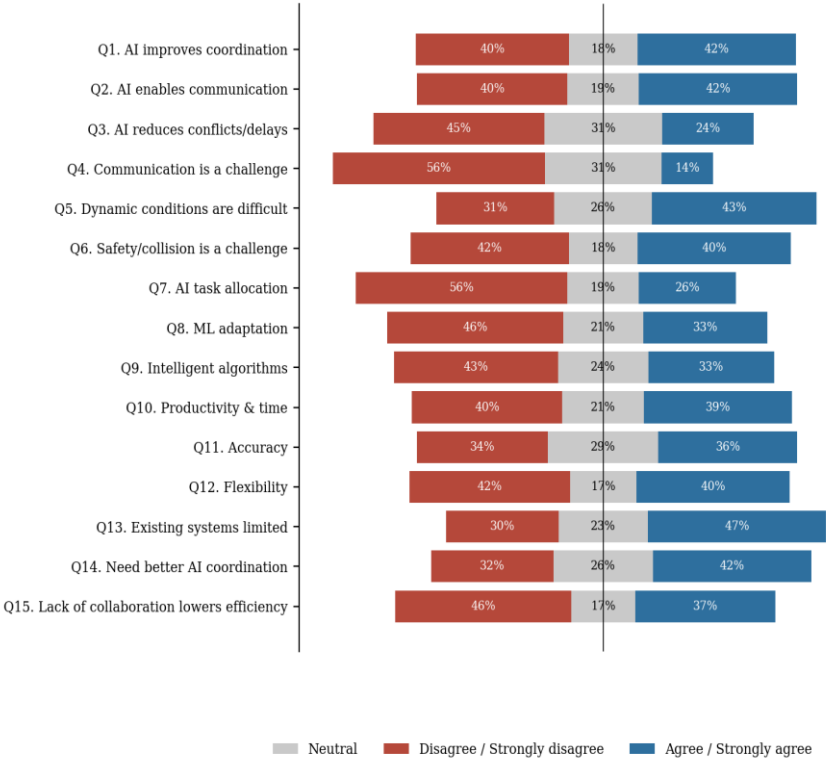


Figure 1. Distribution of responses by item (full dataset). Neutral responses are centred; percentages shown within bars.

Item means ranged from 2.30 (Q4) to 3.17 (Q13), so most answers hovered around the neutral score of 3. Standard deviations were between 1.02 and 1.50, which means opinions varied a lot. Most items had a median of 3, with Q4 and Q7 at 2. Q13 drew the most agreement (47.0%; existing systems focus on individual robots), followed by Q5 (43.2%; changing warehouse conditions). Q4, on communication challenges, drew the least (13.6% agree, 55.8% disagree), and so did Q7, on AI-based task allocation (25.5% agree, 55.7% disagree). Respondents were, on balance, not clearly positive about AI-enabled coordination.

8.3 Construct-level statistics, reliability and tests against the neutral midpoint

Table 3. Construct descriptive statistics, reliability and one-sample t-tests against 3 (full dataset)

Construct	n	Mean	Median	SD	α	t-test	p	Cohen's d
AI-Enabled Coordination (6 items)	149	2.76	2.83	0.61	0.25	t (148) = -4.78	< .001	-0.39
Coordination Challenges (3 items)	149	2.79	2.67	0.80	0.20	t (148) = -3.26	.001	-0.27
Warehouse Performance Impact (3 items)	149	2.96	3.00	0.79	0.01	t (148) = -0.69	.491	-0.06
Limitations of Existing Systems (3 items)	149	3.04	3.00	0.89	0.25	t (148) = 0.58	.560	0.05

Construct	n	Mean	Median	SD	α	t-test	p	Cohen's d
Overall (15 items)	149	2.86	2.87	0.47	0.45	t (148) = -3.57	< .001	-0.29

Note. Wilcoxon signed-rank tests gave the same conclusions (AIC $p < .001$; CHL $p = .003$; PERF $p = .597$; LIM $p = .600$).

The mean for AI-enabled coordination (AIC) was 2.76 (median = 2.83, SD = 0.61), significantly below the neutral value of 3 ($t(148) = -4.78, p < .001, d = -0.39$), a small-to-moderate effect. H1 is supported, although not in the direction we expected, because the overall rating was slightly negative. Coordination challenges were also below neutral ($M = 2.79, p = .001$). Neither perceived performance impact ($M = 2.96$) nor limitations of existing systems ($M = 3.04$) differed significantly from neutral.

Reliability was weak. Cronbach's alpha was 0.25 for AIC, 0.20 for CHL, 0.01 for PERF and 0.25 for LIM, all far below the usual .70 benchmark. The questions within each group, in other words, did not consistently capture one idea in the combined dataset. Construct scores should be read with care, and the item-level results in Table 2 may say more. The original responses were more reliable (AIC = 0.88, CHL = 0.80, PERF = 0.61, LIM = 0.79), which Section 8.8 comes back to.

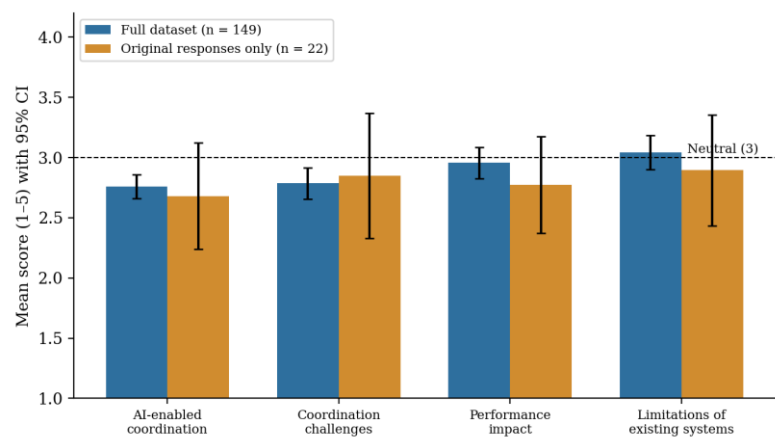


Figure 2. Construct means with 95% confidence intervals: full dataset versus original responses only.

8.4 Correlation analysis

Table 4. Pearson correlations among constructs (full dataset, n = 149)

	AIC	CHL	PERF	LIM
AIC	1	.35**	.09	.24**
CHL	.35**	1	.10	.20*
PERF	.09	.10	1	.15
LIM	.24**	.20*	.15	1

Note. * $p < .05$, ** $p < .01$ (two-tailed). AIC = AI-enabled coordination; CHL = coordination challenges; PERF = performance impact; LIM = limitations of existing systems.

AIC correlated weakly and positively with CHL ($r = .35$, $p < .001$) and LIM ($r = .24$, $p = .003$), and CHL with LIM ($r = .20$, $p = .017$). People who saw benefits in AI coordination also tended to notice a few more coordination problems and flaws in current systems. Perceived performance impact (PERF) was unrelated to every other construct (all $|r| \leq 0.15$).

8.5 Chi-square tests of independence

Table 5. Chi-square tests of independence (full dataset)

Variable pair	χ^2	df	p	n	Cramér's V	Cells with expected < 5
Robot use × AI familiarity	8.15	12	.773	149	0.14	0%
Robot use × Occupation	23.78	12	.022*	149	0.23	20%
Robot use × Organisation type	11.62	12	.477	149	0.16	0%
Robot use × Education	10.71	9	.296	149	0.15	6%
Robot use × Experience	4.65	6	.589	149	0.12	0%
AI familiarity × Occupation	30.93	16	.014*	149	0.23	44%
AI familiarity × Age group	17.92	8	.022*	149	0.25	0%
AI familiarity × Experience	11.87	8	.157	149	0.20	0%
Occupation × Organisation type	20.31	16	.206	150	0.18	40%
Education × Occupation	9.01	12	.702	150	0.14	30%
Age group × Experience	4.72	4	.317	150	0.13	0%
AIC level (agree/neutral/disagree) × Occupation	8.88	8	.352	149	0.17	33%
AIC level (agree/neutral/disagree) × AI familiarity	16.50	8	.036*	149	0.24	33%
AIC level (agree/neutral/disagree) × Organisation type	7.04	8	.532	149	0.15	33%
AIC level (agree/neutral/disagree) × Robot use	4.91	6	.556	149	0.13	33%

Note. * $p < .05$. Gender was constant (all male) and could not be tested. AIC level: mean AIC ≤ 2.5 = disagree; 2.5–3.5 = neutral; > 3.5 = agree.

Four of the fifteen chi-square tests were significant at the .05 level: robot use and occupation, AI familiarity and occupation, AI familiarity and age group, and AIC level and AI familiarity. Effect sizes were small to moderate (Cramér's $V = .23$ – $.25$). Robot use was linked to occupation ($\chi^2(12) = 23.78$, $p = .022$), AI familiarity to age group ($\chi^2(8) = 17.92$, $p = .022$) and occupation

($\chi^2(16) = 30.93$, $p = .014$), and AIC agreement level to AI familiarity ($\chi^2(8) = 16.50$, $p = .036$). Robot use had no significant link with AI familiarity, organisation type, education or experience, so H4 is only partly supported. In some tables as many as 44% of cells had low expected counts, which makes the results shakier, and none survived a Bonferroni correction for fifteen tests ($\alpha = .0033$).

8.6 One-way ANOVA

Table 6. One-way ANOVA of construct scores by respondent characteristics (full dataset)

Outcome	Grouping factor	F	p	η^2	Kruskal–Wallis p	Levene p
AIC	Occupation	F (4, 144) = 2.28	.064	.060	.093	.085
	Organisation type	F (4, 144) = 1.30	.273	.035	.311	.929
	AI familiarity	F (4, 144) = 0.63	.641	.017	.217	.240
	Experience	F (2, 146) = 0.33	.722	.004	.524	.516
	Education	F (3, 145) = 1.33	.268	.027	.139	.495
	Age group	F (2, 146) = 0.09	.918	.001	.866	.020
	Robot use	F (3, 145) = 0.64	.587	.013	.535	.214
CHL	Occupation	F (4, 144) = 3.71	.007*	.093	.006	.322
	Organisation type	F (4, 144) = 1.25	.291	.034	.319	.945
	AI familiarity	F (4, 144) = 0.85	.495	.023	.389	.407
	Experience	F (2, 146) = 1.07	.344	.015	.437	.894
	Education	F (3, 145) = 0.11	.954	.002	.936	.032
	Age group	F (2, 146) = 0.97	.381	.013	.263	.167
	Robot use	F (3, 145) = 1.25	.293	.025	.277	.069
PERF	Occupation	F (4, 144) = 0.71	.583	.019	.631	.635

Outcome	Grouping factor	F	p	η^2	Kruskal–Wallis p	Levene p
	Organisation type	F (4, 144) = 1.15	.336	.031	.542	.106
	AI familiarity	F (4, 144) = 0.19	.943	.005	.894	.880
	Experience	F (2, 146) = 0.56	.573	.008	.655	.920
	Education	F (3, 145) = 1.61	.190	.032	.266	.660
	Age group	F (2, 146) = 0.36	.695	.005	.978	.662
	Robot use	F (3, 145) = 2.00	.116	.040	.059	.433
LIM	Occupation	F (4, 144) = 1.94	.106	.051	.107	.005
	Organisation type	F (4, 144) = 0.18	.950	.005	.929	.787
	AI familiarity	F (4, 144) = 1.99	.099	.052	.102	.610
	Experience	F (2, 146) = 1.87	.158	.025	.179	.329
	Education	F(3, 145) = 0.54	.657	.011	.565	.976
	Age group	F(2, 146) = 0.12	.884	.002	.939	.101
	Robot use	F(3, 145) = 1.87	.137	.037	.125	.142

Note. * $p < .05$. Levene $p < .05$ indicates unequal variances (AIC $\times$ age; CHL $\times$ education; LIM $\times$ occupation); Kruskal–Wallis results agree with ANOVA in every case.

Table 7. Coordination challenges (CHL) by occupation: group means

Occupation	Mean	SD	n
Warehouse Employee	2.61	0.85	42
Warehouse Manager/Supervisor	2.64	0.62	22
IT/AI Professional	3.33	0.73	17
Researcher/Academic	3.01	0.67	29
Other	2.65	0.86	39

Note. Tukey HSD: IT/AI professionals scored higher than warehouse employees (mean difference = 0.72, $p = .013$) and 'Other' roles (0.68, $p = .024$); the difference from managers/supervisors was borderline (0.69, $p = .052$). All other pairs were not significant.

Just one of the 28 ANOVA tests came out significant. Coordination challenges differed by occupation ($F(4, 144) = 3.71, p = .007, \eta^2 = .093$), a medium effect, and the Kruskal–Wallis test agreed ($p = .006$). IT/AI professionals reported the most challenges ($M = 3.33$), compared with warehouse employees ($M = 2.61$) and 'Other' roles ($M = 2.65$). Perhaps technical professionals simply know more about communication, safety and planning problems. AIC ratings across occupations got close ($F(4, 144) = 2.28, p = .064$) without reaching significance, and nothing else differed between groups. H3 is partly supported, then. But the occupation result would not hold up after correcting for 28 comparisons ($\alpha = .0018$), so we see it as a possible pattern, not a firm finding.

8.7 Linear regression

Table 8. Regression models predicting warehouse performance impact (PERF), full dataset

Model	Predictor	B	SE	t	p	95% CI
Model 1	Constant	2.641	0.300	8.79	< .001	[2.05, 3.23]
	AI-enabled coordination (AIC)	0.114	0.106	1.07	.286	[-0.10, 0.32]
Model 2	Constant	2.318	0.356	6.51	< .001	[1.61, 3.02]
	AI-enabled coordination (AIC)	0.047	0.115	0.41	.686	[-0.18, 0.27]
	Coordination challenges (CHL)	0.060	0.087	0.69	.491	[-0.11, 0.23]
	Limitations of existing systems (LIM)	0.112	0.076	1.48	.140	[-0.04, 0.26]
Model 3	Constant	2.204	0.461	4.78	< .001	[1.29, 3.11]
	AI-enabled coordination (AIC)	0.062	0.119	0.52	.602	[-0.17, 0.30]
	Coordination challenges (CHL)	0.052	0.089	0.58	.560	[-0.12, 0.23]
	Limitations of existing systems (LIM)	0.101	0.079	1.28	.203	[-0.05, 0.26]
	AI familiarity (1–5)	-0.005	0.047	-0.10	.919	[-0.10, 0.09]
	Experience (1–3)	0.064	0.079	0.82	.416	[-0.09, 0.22]
	Education (1–4)	0.009	0.060	0.15	.878	[-0.11, 0.13]

Note. Model 1: $R^2 = 0.008$, $F(1, 147) = 1.15$, $p = .286$. Model 2: $R^2 = 0.028$, adj. $R^2 = 0.008$, $F(3, 145) = 1.37$, $p = .253$. Model 3: $R^2 = 0.032$, adj. $R^2 = -0.009$, $F(6, 142) = 0.79$, $p = .581$.

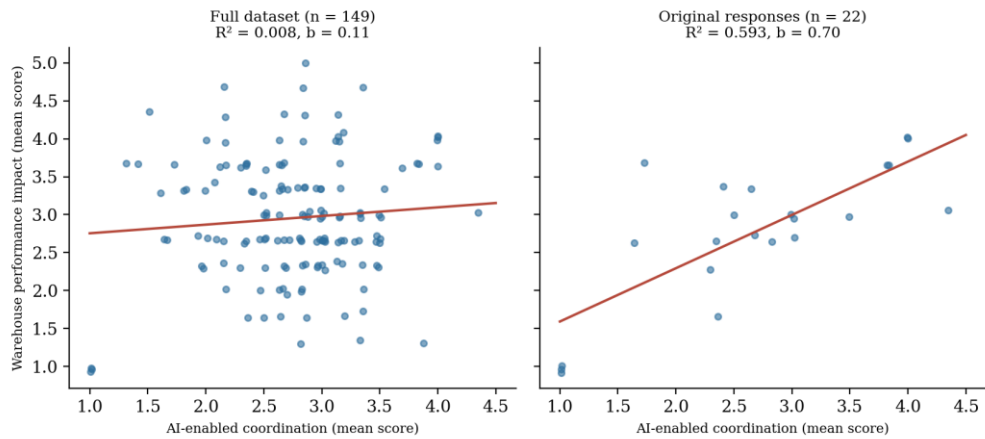


Figure 3. Relationship between AI-enabled coordination and performance impact with fitted regression line (full dataset and original responses only). Points are jittered slightly for visibility.

AIC explained less than 1% of the variation in perceived warehouse performance in Model 1 ($R^2 = 0.008$), and its coefficient was small and not significant ($B = 0.114$, $SE = 0.106$, $p = .286$). We added coordination challenges and system limitations in Model 2 and background characteristics in Model 3. Neither model was significant ($R^2 = 0.028$ and 0.032 ; both F-tests $p > .25$), and no individual predictor was either. Model checks showed no major problems. Variance inflation factors ranged from 1.08 to 1.18, the Durbin–Watson statistic was 1.99 and the Shapiro–Wilk test of residuals gave $p = 0.528$. In the combined dataset, H2 was not supported.

8.8 Sensitivity analysis: original responses only

Table 9. Sensitivity of key results to data source

Statistic	Full dataset	Original responses	Original, excluding uniform responders
Respondents analysed	149	22	16
AIC mean (SD)	2.76 (0.61)	2.68 (0.99)	2.87 (0.80)
PERF mean (SD)	2.96 (0.79)	2.77 (0.90)	3.00 (0.60)
Cronbach α (AIC / CHL / PERF / LIM)	0.25 / 0.20 / 0.01 / 0.25	0.88 / 0.80 / 0.61 / 0.79	—
r (AIC, PERF)	.09	.77	.44
Regression: PERF on AIC, B (p)	0.11 (.286)	0.70 (< .001)	0.32 (.092)
R ²	0.008	0.593	0.190
Multiple regression (Model 2) R ²	0.028	0.682	—
Maximum VIF (Model 2)	1.2	8.3	—

Note. Uniform responders gave the same rating to all fifteen items (6 of the 22 analysable original respondents; none of the 127 synthetic records did).

Using only the original responses changed the results a great deal. AIC and PERF became strongly and significantly related ($R^2 = 0.59$, $B = 0.70$, $p < .001$), reliability ranged from 0.61 to

0.88, and correlations among constructs were high ($r = 0.69\text{--}0.91$). We'd urge caution for three reasons. One, only 22 original responses could be analysed, so estimates are uncertain; the 95% confidence interval for mean AIC was 2.24–3.12. Two, six respondents gave exactly the same rating to all fifteen questions, which can inflate correlations and reliability. Without them, the AIC–PERF correlation fell to $r = .44$ ($n = 16$, $p = .092$). Three, the multiple regression showed serious multicollinearity (maximum VIF = 8.3).

The original and synthetic records behaved differently too. Average construct scores were similar (Welch t-tests $p > .30$ for all four constructs). But the synthetic records showed very weak relationships among items and had no uniform responders, so they resembled independent random answers. The original answers hung together better. That goes a long way toward explaining the low reliability and almost flat regression result in the combined dataset. It would be wrong to read the weak full-dataset result as proof that AI coordination and warehouse performance are unrelated; it may largely come from mixing genuine and synthetic records.

8.9 Summary of hypothesis tests

Table 10. Summary of hypothesis testing

Hyp.	Statement	Full-dataset result	Decision (full dataset)	Original responses only
H1	Mean AIC differs from neutral (3)	$t(148) = -4.78$, $p < .001$	Supported (mean below neutral)	Not significant ($p = .147$; $n = 22$)
H2	AIC predicts PERF ($\beta > 0$)	$B = 0.11$, $R^2 = 0.008$, $p = .286$	Not supported	$B = 0.70$, $R^2 = 0.59$, $p < .001$ (caution)
H3	Construct scores differ across groups	1 of 28 ANOVAs significant ($\text{CHL} \times \text{occupation}$, $p = .007$)	Partly supported	2 of 28 significant at .05 ($n = 22$; unadjusted)
H4	Respondent characteristics associated	4 of 15 chi-square tests significant ($p = .014\text{--}.036$)	Partly supported	4 of 15 significant (>90% sparse cells)

9. Discussion

We set out to learn how people working with warehousing, robotics and automation view AI-based multi-robot coordination. Four findings stand out.

The first is that views on the benefits of AI coordination were mixed. The mean AIC score was 2.76 out of 5, slightly below neutral, and agreement with individual AI-coordination items ranged from roughly 24% to 42%. That is less positive than earlier work like He (2022) and Saruhasan (2026), who present AI as an important tool for robot collaboration. One explanation could be that respondents haven't yet seen much practical evidence that AI improves cooperation among several robots, as distinct from improving one.

The second is that respondents agreed more with statements about the limits of existing systems than with statements about AI's benefits. About 47% agreed that existing systems focus on individual robots (Q13), and around 42% that current systems need better AI coordination (Q14). Both speak to Objective 5 and to the research gap described earlier. Changing warehouse conditions (Q5) were also widely seen as a challenge, which suggests that adapting to real-time change is still an open concern.

The third is that respondent groups differed little. IT/AI professionals rated coordination challenges higher than warehouse employees did, maybe because they understand the technical problems better. A handful of chi-square tests connected AI familiarity, occupation, robot use and agreement with AI coordination. Comparisons by age, education, experience or organisation type mostly showed nothing clear. None of this held up after correcting for multiple tests, so firm conclusions will need more research.

The fourth is that the combined dataset gave no support to the main hypothesis that higher perceived AI coordination goes with better perceived warehouse performance. AIC explained less than 1% of the variation in PERF. The original responses showed a much stronger relationship, but there were few of them and several came from people who gave one answer throughout. The synthetic records, meanwhile, added many responses with no clear pattern. These data can't reliably show whether AI-based coordination improves multi-robot performance, either way.

10. Conclusion

We used a 15-item questionnaire to examine the role of AI in coordinating multiple warehouse robots. The average rating for AI coordination was slightly below neutral ($M = 2.76$ against 3.00), and respondents somewhat agreed that current systems focus too much on individual robots. AI coordination did not significantly predict perceived warehouse performance in the combined dataset. Occupation was linked to views on coordination challenges and to robot use, and other background characteristics showed little connection. The five objectives were addressed, but the data could not confirm our main hypothesis that AI significantly improves multi-robot coordination.

Warehouse organisations may still find the results useful if they apply them with care. Technical specialists and frontline workers may see coordination challenges differently. Anyone considering AI-based systems should pay attention to robot communication, safety and changing warehouse conditions, and should collect real performance measures (task completion time, collision rates, throughput) rather than rely on opinion alone.

11. Limitations and Future Research

- Data provenance. Only 23 of the 150 records are original responses; the other 127 are flagged as synthetic sample data. Results from the combined dataset are not evidence from 150 real respondents and shouldn't be reported that way. Any submission should state the sample composition and follow the institution's rules on synthetic data.

- Small genuine sample. Six of the 22 analysable original responses are uniform responders, so statistical power is low and estimates are unstable.
- Low reliability. Alpha values in the combined dataset (0.01–0.25) show the constructs were not measured consistently. The researcher defined them, and they were never validated with factor analysis.
- Perception-based measures. Every variable is a self-reported attitude, not a measure of robot performance. The study can't establish actual coordination effects or causation.
- Sample composition. All respondents are male, and most are young and early in their careers. That limits generalisability, and gender could not be analysed.
- Multiple testing and sparse tables. We ran forty-three significance tests without correction, and many chi-square tables had expected counts below 5.
- Ordinal data. Likert items were treated as interval-level in the ANOVA and regression, though the non-parametric checks (Wilcoxon, Kruskal–Wallis) supported the conclusions.

Future studies should gather more genuine responses from a wider range of participants. They should also test, through exploratory and confirmatory factor analysis, whether the questionnaire measures its intended concepts consistently. Reverse-coded questions and attention checks could help catch people who give the same answer throughout. And survey responses ought to be compared with real warehouse measurements or simulation results, to see how multi-robot coordination performs in practice.

References

1. Bas, G. (2020). [Multi Functional Intelligent Measurement Robot, as a Part of a Global Quality Management System for Collaboration Between Small, Medium and Large Scale Enterprises.]
2. Danford, M. (2020). [What Makes Smart Robots Smart?].
3. García, M. R. (2026). [Managing technology-related disruptions and vulnerabilities in highly automated warehouse systems: an integrative review and research agenda].
4. Grosse, E. H., et al. (2023). [Human-Centric Production and Logistics System Design and Management: Transitioning from Industry 4.0 to Industry 5.0].
5. He, X. (2022). [Multi-agent system optimisation in factories of the future: cyber collaborative warehouse study].
6. Hupp, N. (2024). [3D Imaging, AI-Based Software, and Industrial Computing Democratize Application-Specific Automation Systems].
7. Kollar, T., Tellex, S., & Roy, N. (2018). [Learning to Map Words to Actions for Robotic Manipulation].
8. LeCun, Y., Bengio, Y., & Hinton, G. (2015). [Deep learning].
9. Saruhasan, A. (2026). [Robotics and Artificial Intelligence].

10. Sheffi, Y. (2026). [Estimating the task content of work: workforce design for AI-driven human-robot collaboration in intralogistics].
11. Tubis, A., & Rohman, J. (2023). [Intelligent Warehouse in Industry 4.0—Systematic Literature Review].
12. Wu, G. (2026). [Trading off travel distance and fatigue. The effect of storage, order batching, and pod selection in robotic mobile fulfillment systems].
13. Zelinski, P. (2019). [Unguarded Automation].