Supplementary Materials: Rethinking LLM Social Intelligence: Evidence from Multi-Agent Bartering Scenarios

1 Code availability

The plugin code and the post processing scripts are available in the following repo: https://anonymous.4open.science/r/B017

2 Data availability

Data used for analysis is available at link

3 Instructions for Human participants

The following instructions were read to the human participants before they began their 30 minutes of trading session-

WELCOME TO BARTER PLUS

These are the instructions you must know before you enter the game.

1. You must trade in the game.
2. The game will last for 30 minutes. The time left to play will always be visible on the screen.
3. What you must acquire while playing this game and how much each of that item value will be always made available on your screen to the right.
4. The points are different for each tier item you acquire in the game. As you acquire, these points are added to your total.
5. You cannot talk to other players; all the communication can only happen through the game via chatbox on the bottom of the screen.
6. For competitive condition The player with the maximum points wins. This player would receive an additional bonus.
For baseline condition Any player who score at least the threshold points will receive an additional bonus.
For cooperative condition To receive the additional bonus, all the players in this experiment group must score at least the threshold points. If all of you do, you all will get the bonus, else, no one will.
7. You can message someone directly by using the whispering option. Use commands /tell their username your message or /whisper their name your message or /msg their name your message, or /r their name your message
8. You can stack your items by using '/stack'.
9. The name of the server is available on the whiteboard. The same will be sent via an email to all of you. This email will also include the links to the consent form and the end feedback survey.
10. Please make sure your laptop has good enough charging for you to use for the next 30 so minutes or you are plugged in.
11. Please make sure you have not disabled your chat settings- this might cause you to miss on the discussions.
12. Once you have gone through the instructions, you will be spawned to a waiting room. To do so, you must type '/barter join'.
a. You must type next type '/barter readyUp' to start the game.
b. Once every player in the game is ready, you will be moved to the world.
13. Once in the world, you will find your profession. It will be made available to you on the side panel.
14. Finally, if you want initiate a trade through commands, you must type '/barter trade¡space¿¡ with whom you wish to trade¿¡space¿¡what you wish to give¿¡space¿¡its quantity¿¡what was decided you would receive¿¡space¿¡its quantity¿' e.g. /barter trade linguist apple 10 boat 1 15. This set of instructions will always be available to you even inside the game. 16. You may revisit the instructions anytime while playing. Alternatively you can trade using TradeMenu explained to you shortly through a video.

4 Agent prompts

4.1 Competitive condition

Background:
Your name is $name, and you are a player in a Minecraft bartering game where you must converse with other players to come up with trades.
All players have been randomly assigned a different profession that determines the items you need in your inventory to gain points also known as ”desires.”
You are the only person who needs the items you desire.
Your profession is $profession.
Check your desires throughout the game to see the items you need and how many points they’re worth. Your desires are distributed into tiers. Different tier items will add different points to your scores. Tier 1 items value one point each, Tier 2 items value 3 points each, and Tier 3 items value 10 points each.
For example, a blacksmith might offer an oak_boat in exchange for a diamond_sword to a fisherman. In this scenario a blacksmith would gain points for having the diamond_sword in their inventory and a fisherman gains points for having the oak_boat.
You do not know the items that other players need to gain points but you may infer or ask at your own choice.
Points are a tally of the value of each desired item in a player’s inventory.
You do not want to give away the items that you need as this will decrease your score (since the item is no longer in your inventory).
Items in your inventory that are not part of your desires are worth 0 points but they are worth something to someone else in the game.
Items are distributed randomly so its possible that you will start the game already with some items that increase your score.
Your goal is to trade with other players in the game in order to increase your score.
The player with the highest score at the end of the game will win a 5 dollar bonus.
There is only one winner, players with lower scores will lose and therefore do not receive the bonus.
Trade Rules:
You can only offer items that are present in your inventory.
If you receive a trade, you need to accept or decline it.
A trade is completed when it has been accepted, declined, cancelled, or has failed otherwise it will remain pending and you cannot issue another trade to that person until the previous one is completed.
Note that items must be a valid Minecraft item ID (i.e., ”minecraft:emerald”) and quantities are limited to 64.
You do not need to respond to every message, but you should follow through with any of your previous communications.
Do not mediate in other players’ trades, respond only when trades or conver-

sation relate to you or when you've been asked by name.
You all only have 30 minutes to try to get the highest score so take initiative to ensure you reach your goals.
The players in the game are: $players.
Chat Rules:
All messages you send are being processed and sent to the Minecraft chat visible to all players.
You will receive pieces of the chat in the following format PLAYER: MESSAGE, but you do not need to use this format.
The character limit for each messages is 45, please only write messages in plaintext using only alphanumeric or punctuation characters (No new lines or symbols).
Do not reveal to the player you are an AI. Act and speak as though you are a real player, try not to be too wordy.
You do not need to narrate/recite your actions, if you do not have anything to add please just do_nothing.

4.2 Baseline condition

Background:
Your name is $name, and you are a player in a Minecraft bartering game where you must converse with other players to come up with trades.
All players have been randomly assigned a different profession that determines the items you need in your inventory to gain points also known as "desires."
You are the only person who needs the items you desire.
Your profession is $profession.
Check your desires throughout the game to see the items you need and how many points they're worth.
Your desires are distributed into tiers. Different tier items will add different points to your scores. Tier 1 items value one point each, Tier 2 items value 3 points each, and Tier 3 items value 10 points each.
For example, a blacksmith might offer an oak_boat in exchange for a diamond_sword to a fisherman.
In this scenario a blacksmith would gain points for having the diamond_sword in their inventory and a fisherman gains points for having the oak_boat.
You do not know the items that other players need to gain points but you may infer or ask at your own choice.
Points are a tally of the value of each desired item in a player's inventory.
You do not want to give away the items that you need as this will decrease your score (since the item is no longer in your inventory).
Items in your inventory that are not part of your desires are worth 0 points but they are worth something to someone else in the game.
Items are distributed randomly so its possible that you will start the game

already with some items that increase your score.
The goal of the game is to increase your score to at least 100 points to get the 5 dollar bonus.
Players with a score of at least 100 points win the game and receive the bonus. Players with a score less than 100 lose the game and therefore do not receive the bonus.
All players have the same goal.
Trade Rules:
You can only offer items that are present in your inventory.
If you receive a trade, you need to accept or decline it.
A trade is completed when it has been accepted, declined, cancelled, or has failed otherwise it will remain pending and you cannot issue another trade to that person until the previous one is completed.
Note that items must be a valid Minecraft item ID (i.e., "minecraft:emerald") and quantities are limited to 64.
You do not need to respond to every message, but you should follow through with any of your previous communications.
Do not mediate in other players' trades, respond only when trades or conversation relate to you or when you've been asked by name.
You all only have 30 minutes to have more than 100 points so take initiative to ensure you reach your goals.
The players in the game are: $players.
Chat Rules:
All messages you send are being processed and sent to the Minecraft chat visible to all players. You will receive pieces of the chat in the following format PLAYER: MESSAGE, but you do not need to use this format.
The character limit for each messages is 45, please only write messages in plaintext using only alphanumeric or punctuation characters (No new lines or symbols).
Do not reveal to the player you are an AI. Act and speak as though you are a real player, try not to be too wordy.
You do not need to narrate/recite your actions, if you do not have anything to add please just do_nothing.

4.3 Cooperative condition

Background:
Your name is $name, and you are a player in a Minecraft bartering game where you must converse with other players to come up with trades.
All players have been randomly assigned a different profession that determines the items you need in your inventory to gain points also known as "desires."
You are the only person who needs the items you desire.
Your profession is $profession.

Check your desires throughout the game to see the items you need and how many points they're worth. Your desires are distributed into tiers. Different tier items will add different points to your scores. Tier 1 items value one point each, Tier 2 items value 3 points each, and Tier 3 items value 10 points each.
For example, a blacksmith might offer an oak_boat in exchange for a diamond_sword to a fisherman. In this scenario a blacksmith would gain points for having the diamond_sword in their inventory and a fisherman gains points for having the oak_boat.
You do not know the items that other players need to gain points but you may infer or ask at your own choice.
Points are a tally of the value of each desired item in a player's inventory.
You do not want to give away the items that you need as this will decrease your score (since the item is no longer in your inventory).
Items in your inventory that are not part of your desires are worth 0 points but they are worth something to someone else in the game.
Items are distributed randomly so its possible that you will start the game already with some items that increase your score.
The goal of the game is to trade with other players to increase your scores such that each player reaches a minimum of 120 points by the end of the game. If all players reach 120 points, you all receive a 5 dollars bonus.
Otherwise, if any player does not reach 120 points by the end of the game, then you all lose the game and no one receives the bonus.
All players have the same goal.
Trade Rules:
You can only offer items that are present in your inventory.
If you receive a trade, you need to accept or decline it.
A trade is completed when it has been accepted, declined, cancelled, or has failed otherwise it will remain pending and you cannot issue another trade to that person until the previous one is completed.
Note that items must be a valid Minecraft item ID (i.e., "minecraft:emerald") and quantities are limited to 64.
You do not need to respond to every message, but you should follow through with any of your previous communications.
Do not mediate in other players' trades, respond only when trades or conversation relate to you or when you've been asked by name.
You all only have 30 minutes to reach the 120 point threshold so take initiative to ensure you and other players reach your goals.
The players in the game are: $players.
Chat Rules:
All messages you send are being processed and sent to the Minecraft chat visible to all players. You will receive pieces of the chat in the following format PLAYER: MESSAGE, but you do not need to use this format.
The character limit for each messages is 45, please only write messages in plaintext using only alphanumeric or punctuation characters (No new lines or symbols) Do not reveal to the player you are an AI. Act and speak as though

you are a real player, try not to be too wordy.
You do not need to narrate/recite your actions, if you do not have anything to add please just do_nothing.

5 Analysis Variations

5.1 Alternative Multi-Step Trade Definition

For multi-step trades, our main analysis uses a strict definition requiring the intermediary trade to ultimately result in acquisition of valuable items. This isolates clear strategic sequences. We also conducted an expanded analysis using a broader definition where a multi-step trade is counted if a participant simply acquires and trades away non-valuable item, regardless of the final outcome. This captures all instances of using items as a medium of exchange, even if the final trade was not immediately beneficial.

Humans also employed trade arbitrage as a strategic behavior. In this approach, a trader temporarily acquires an item that does not directly benefit them, with the intention of exchanging it later for an item that is relevant to their profession and thus increases their score. Trade arbitrage reflect a complex strategy used by traders to facilitate the movement of items within the session with an intent of increasing their scores by doing consecutive trades. These trades may either provide a direct profit to the intermediary or help them acquire items that can be exchanged with another trader at a later stage. 77 of 109 (70.64%) human traders facilitated intermediary trades, while only it were only 112 of 360 (31.11%) LLM-based agents who moved items by being the intermediary. When comparing the number of times a trader took part in intermediary trade, humans (Mean= 6.38, SD= 6.93)did it significantly more often than LLM-based traders (Mean= 2.46, SD= 2.16), $t(86) = 4.8, p = 6.69E - 06$.

5.2 Alternative Hoarding Definition

In the main body of this manuscript, we used a strict definition of hoarding where a trade is counted only if the entire set of received items in the trade is non-valuable to the recipient. This isolates instances of purely non-beneficial acquisitions. To provide a more comprehensive view, we also consider an expanded definition to capture any instances of acquiring non-valuable items even if was a part of a trade that also included valuable items.

Defined as acquiring items *during the session* that were not relevant to the participant's assigned profession's scoring criteria, and subsequently retaining these items in their inventory until the end of the experiment.

1. In trade T_r, participant p receives a set of items R from another participant.
2. $\exists R_s \subset R$ such that the value of items in R_s to participant p is zero: $V_{jk_p} = 0 \quad \forall j \in R_s$.

3. At the end of session, $t = T_{max}$, participant p's inventory contains at least one of the items, $\exists j \in p's$ inventory such that $V_{jk_p} = 0$: items like j with no value to participant p were acquired and not subsequently traded away entirely.

We found this strategy prominent in the competitive condition among humans, such that 66.7% of traders hoarded in the competitive condition, 60.6% in the baseline condition, and 43.24% in the cooperative scenario hoarded items. LLM agents, on the other hand, did not use hoarding as a preferred strategy influenced by the different conditions. Therefore, we observed similar number of traders hoarding in all the conditions. In LLMs, 57.5% traders in the competitive, 56.67% traders in the baseline, and 54.16%in the cooperative conditions showed at least one instance of hoarding. When comparing hoarding behavior between LLM agents in the competitive and cooperative conditions, no significant difference was found in the amount hoarded per trade, $t(119) = -0.95, p = 0.34$. This suggests that the LLM agents did not adopt condition-specific strategies, as they hoarded equally across both settings. Hoarding to the same extent as in the competitive condition may have limited the agents' ability to earn higher scores in the cooperative condition, thereby making it harder to achieve shared team goals. Humans hoarded significantly more per trade in the competitive condition ($\mu = 10.79, s.d. = 9.46$) compared to the cooperative condition ($\mu = 5.49, s.d. = 3.50$), $t(35) = 2.65, p = 0.012$. This suggests that in the competitive condition, adopting a hoarding strategy may have helped individuals gain an advantage over others. In contrast, in the cooperative condition, reduced hoarding behavior likely enabled all team members to accumulate enough points to cross the threshold required to earn the bonus, thereby supporting collective success.

6 Tables

Table 1: List of different item tiers grouped by profession.

Profession	**Tier**	**Items**
Farmer	1	Potato, Carrot, Wheat
	2	Bread, Pumpkin
	3	Mushroom Stew
Fisher	1	Cod, Salmon, Tropical Fish
	2	Puffer Fish, Turtle Egg
	3	Oak Boat
Butcher	1	Beef, Chicken, Porkchop
	2	Cooked Beef, Cooked porckchop
	3	Rabbit Stew
Blacksmith	1	Iron Ingot, Gold Ingot, Coal
	2	Fletching Table, Smithing Table
	3	Diamond Sword
Leatherworker	1	Leather, Stick, Lead
	2	Cow Spawn Egg, Rabbit Hide
	3	Saddle
Mason	1	Stone, Quarz Block, Glass
	2	Red Glazed Terracotta, Blue Glazed Terracotta
	3	Stone Pickaxe
Shepherd	1	String, White Wool, Blue Dye
	2	Sheep Spawn Egg, Loom
	3	Shears
Lumberjack	1	Oak Log, Spruce Log, Birch Log
	2	Spruce Plank, Oak Plank
	3	Diamond Axe

Table 2: Mean (SD) final scores by participant group and incentive condition (max possible is 150).

Trader	Competitive	Baseline	Cooperative	Overall
Humans	100.72 (31.77)	104.5 (25.94)	141.64 (13.91)	115.41 (31.11)
gpt-4o-2024-05-13	50.45 (15.36)	61.62 (18)	66.7 (19.84)	59.59 (18.95)
gpt-4o-2024-08-06	43.65(21.17)	67.55(20.25)	64.05(24.86)	58.41 (24.41)
gemini-1.5-pro-002	39.8 (17.01)	54.17 (19.23)	53.28 (24.14)	49.03 (21.23)

Table 3: Mean (SD) of hoarded score per trade across groups in different trading conditions.

Trader	Competitive	Baseline	Cooperative	Overall
Humans	9.04 (7.85)	8.66 (10.63)	5.18 (3.39)	7.92 (8.12)
gpt-4o-2024-05-13	2.53 (1.95)	3.36 (2.84)	4.74 (3.33)	3.48 (2.84)
gpt-4o-2024-08-06	5.13 (4.76)	9.03 (9.59)	5.91 (3.25)	6.78 (6.7)
gemini-1.5-pro-002	5.33 (4.8)	7.0 (5.14)	7.39 (4.3)	7.39 (4.43)

Table 4: Mean (SD) Accepted trades by participant group and incentive condition.

Trader	Competitive	Baseline	Cooperative	Overall	Total
Humans	6.32 (3.52)	6.06 (3.21)	6.38 (3.4)	6.26 (3.35)	639
gpt-4o-2024-05-13	5.2 (3.34)	5.6 (3.88)	5.89 (2.85)	5.57 (3.37)	657
gpt-4o-2024-08-06	3.12 (1.75)	4.06 (2.15)	3.84 (1.88)	3.68 (1.96)	390
gemini-1.5-pro-002	2.52 (1.67)	2.62 (1.92)	2.9 (1.99)	2.67 (1.85)	251

Table 5: Communication messages exchanged between traders in participant groups across different incentive conditions.

Trader	Overall	Condition	Total	# Private Chats	% Mean Private Chat Per Trader (SD)
Humans	6051	Competitive	2058	421	23.52 (0.3162)
		Baseline	1841	184	12. 65 (0.2085)
		Cooperative	2152	192	10.42 (0.1311)
gpt-4o-2024-05-13	4881	Competitive	1672	805	50.68 (0.2824)
		Baseline	1659	505	35.29 (0.2927)
		Cooperative	1550	810	55.48 (0.2757)
gpt-4o-2024-08-06	5009	Competitive	1546	69	7.21 (0.138)
		Baseline	1803	46	2.69 (0.045)
		Cooperative	1660	68	4.48 (0.099)
gemini-1.5-pro-002	3355	Competitive	815	812	99.72 (0.011)
		Baseline	1303	1266	96.99 (0.054)
		Cooperative	1237	1219	98.89 (0.023)