

# ELICIT

Paper ID 041

## Comparison between Human and Agent Runs in the ELICIT N2C2M2 Validation Experiments

Marco Manso (SAS-085 Member), Portugal

Mary Ruddy (Meiristic), USA

18th ICCRTS

C2 in Underdeveloped, Degraded and Denied Operational Environments

June 19-21, 2013 - Alexandria, Virginia, U.S.A.

This work received financial support from the  
Center for Edge Power at the US Naval Postgraduate School



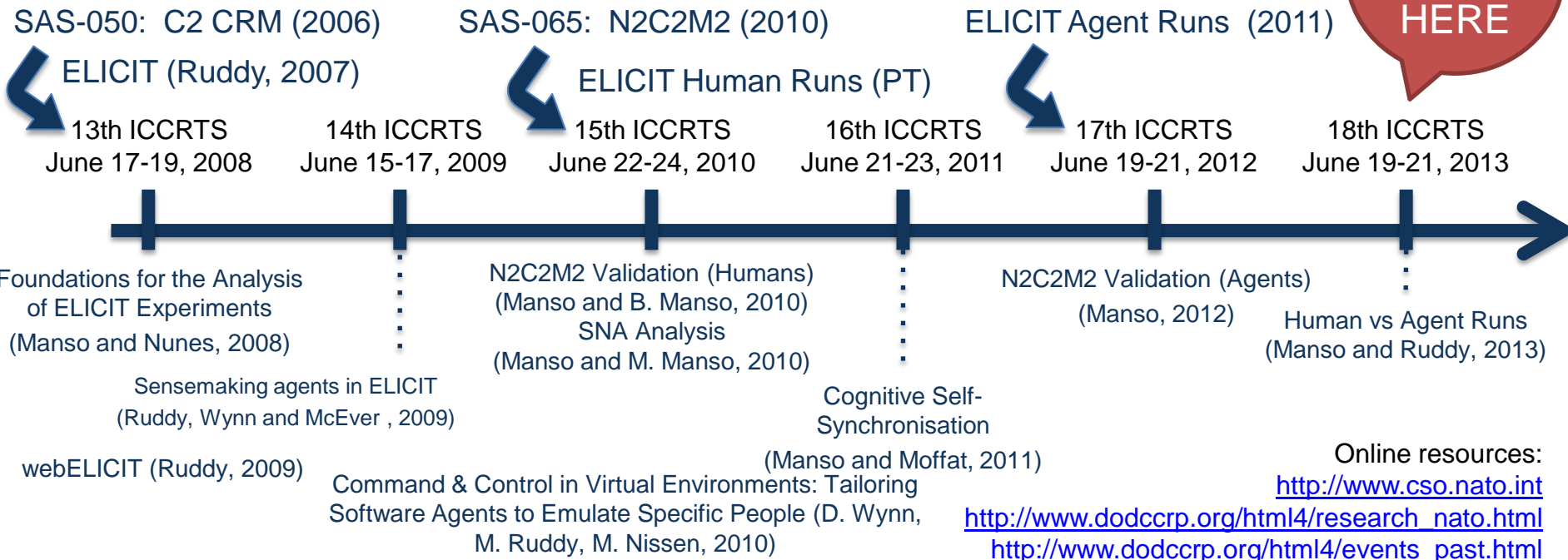
# Outline



- Introduction
- Description of Experiments
- Comparison of Results
- Conclusions

# Introduction

## Ongoing series of experiments within NATO SAS groups using the ELICIT Platform



# Introduction



## Motivation

- Human participants frequently behave in uncontrolled ways and are expensive to recruit
- Agents are predictable; thousands of runs can be generated

How well do ELICIT agents mimic humans in this context?

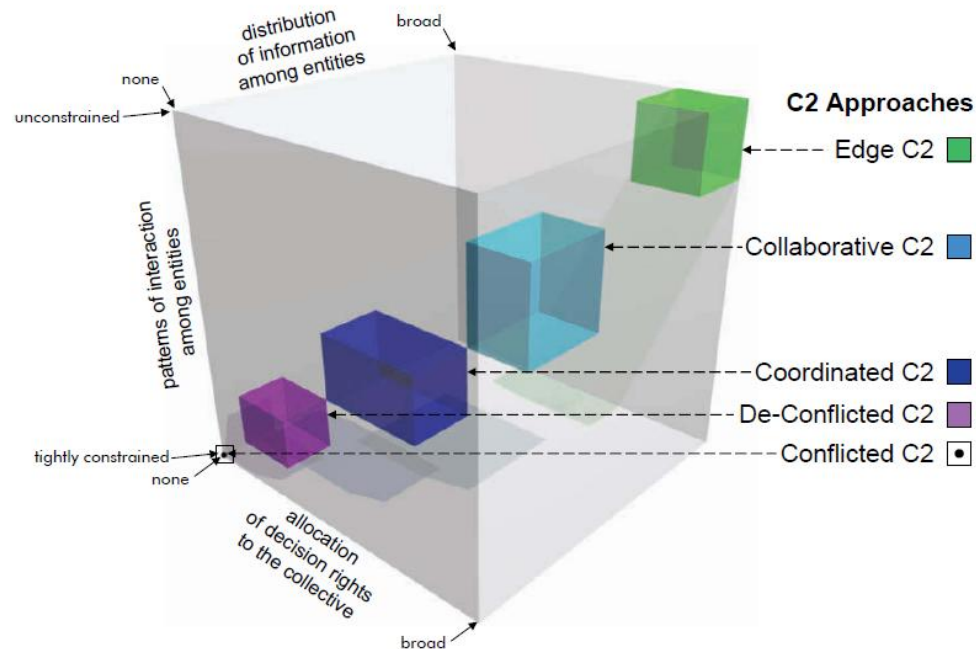
# Experiments



- Background Theory
- Design
- Measurements

# Experiments: background

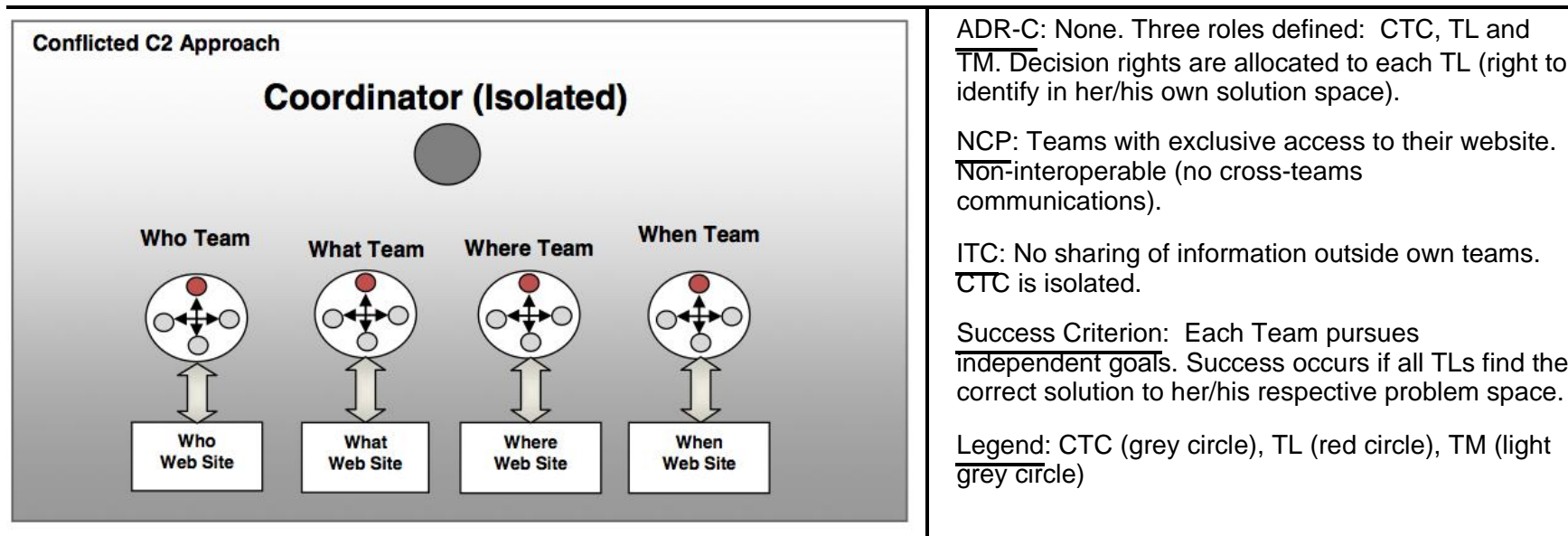
**Foundations:** NCW Tenets, NCW Value Chain (SAS-065, 2010), C2 Domains (Alberts and Hayes, 2006), C2 CRM (SAS-050, 2006), C2 Approach Space (SAS-050, 2006), N2C2M2 (SAS-065, 2010)



Source: NATO SAS-065 (2010)

# Design of Experiments

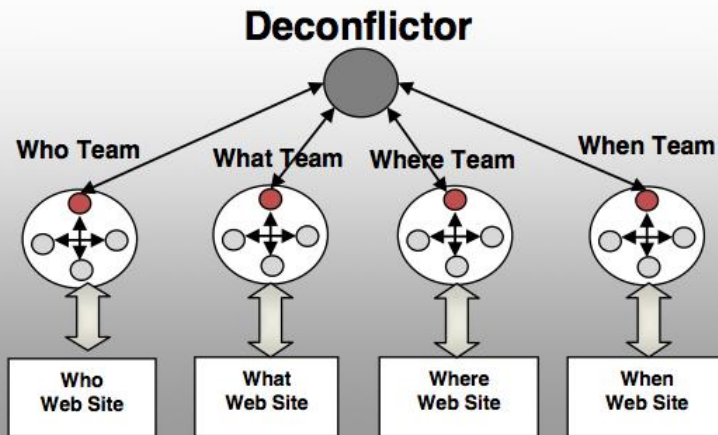
## Instantiation of Conflicted C2



# Design of Experiments

## Instantiation of De-Conflicted C2

### De-conflicted C2 Approach



**ADR-C:** Established constraints. Three roles defined: Deconf, TL and TM. Decision rights are allocated to each TL (right to identify in her/his own solution space).

**NCP:** Minimum connectivity allowed. Stove-pipe: between TLs and Deconf. Teams have exclusive access to their websites.

**ITC:** Interactions across teams allowed but strictly between each TL and Deconf.

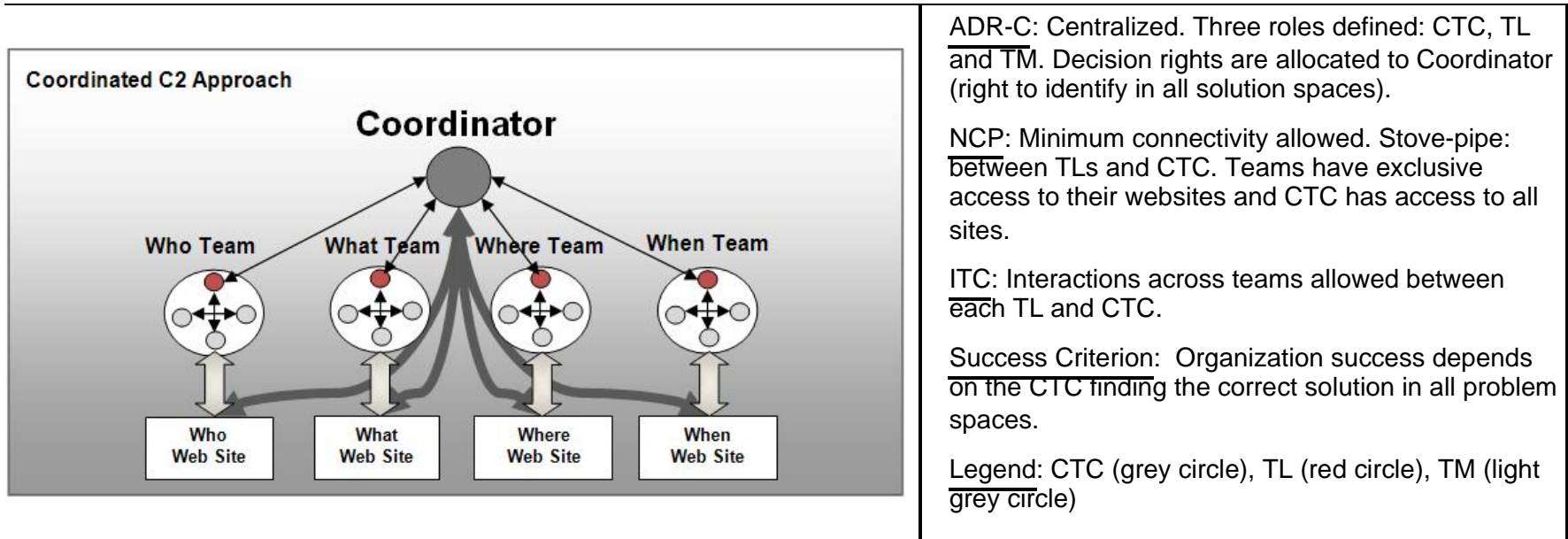
**Success Criterion:** Each Team pursues independent goals for an interdependent problem. Success occurs if all TLs find the correct solution to her/his respective problem space.

**Legend:** Deconf (grey circle), TL (red circle), TM (light grey circle)



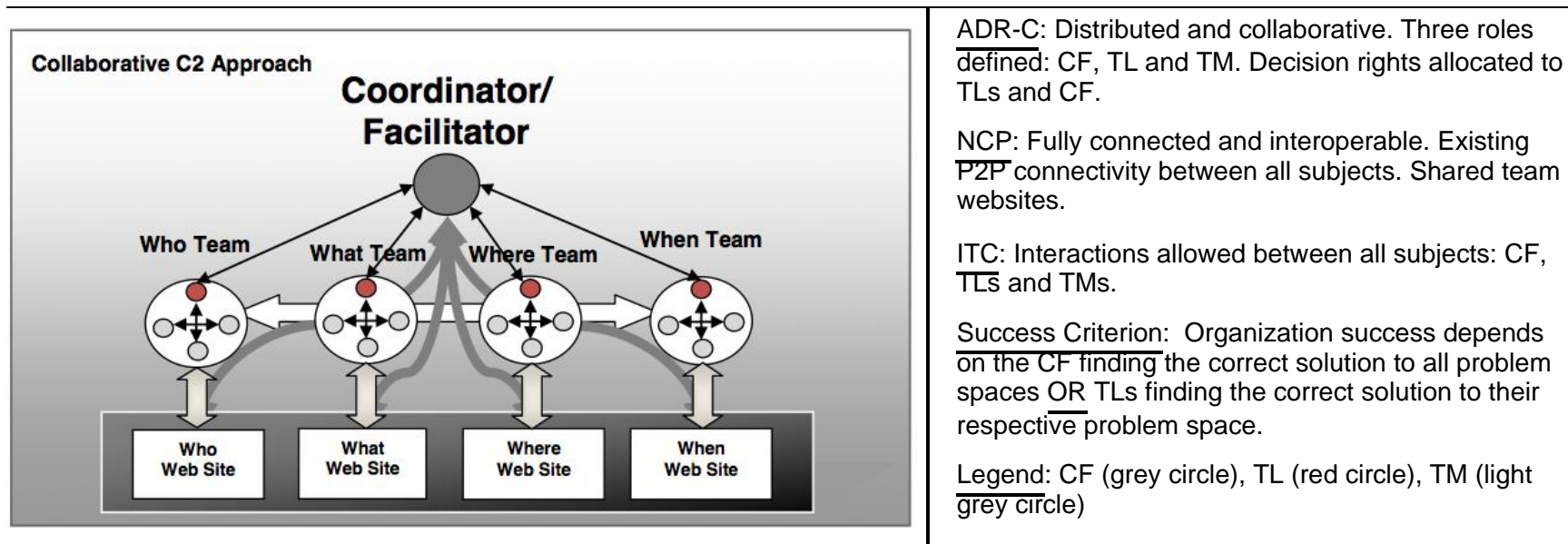
# Design of Experiments

## Instantiation of Coordinated C2



# Design of Experiments

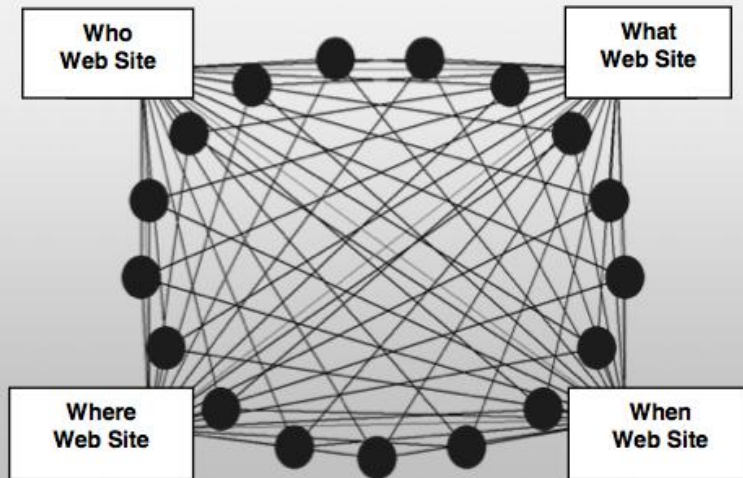
## Instantiation of Collaborative C2



# Design of Experiments

## Instantiation of Edge C2

### Edge C2 Approach



ADR-C: Fully distributed, not explicit (per individual) and dynamic. One role is pre-defined: TM. TMs choose which part (or parts) of the problem space they work on.

NCP: Fully connected and interoperable. Existing P2P connectivity between all individuals. Shared team websites.

ITC: interactions allowed between all TMs.

Success Criterion: Organization success depends on the individuals' IDs plurality being correct in each problem space.

Legend: TM (black circle)

# Design of Experiments



## Datasets

Human Runs	Agent Runs
18 Runs: <ul style="list-style-type: none"><li>- 3 Conflicted C2</li><li>- 4 De-Conflicted C2</li><li>- 4 Coordinated C2</li><li>- 4 Collaborative C2</li><li>- 3 Edge C2</li></ul>	135 Runs: <ul style="list-style-type: none"><li>- 27 runs for each C2 Approach</li></ul> Changed agent configuration: (i) low-performing (ii) avg-performing (iii) high-performing

**For simplicity:** C2 Approaches will be numbered as follows: Conflicted C2 as 1, De-Conflicted C2 as 2, Coordinated C2 as 3, Collaborative C2 as 4 and Edge C2 as 5.

# Design of Experiments



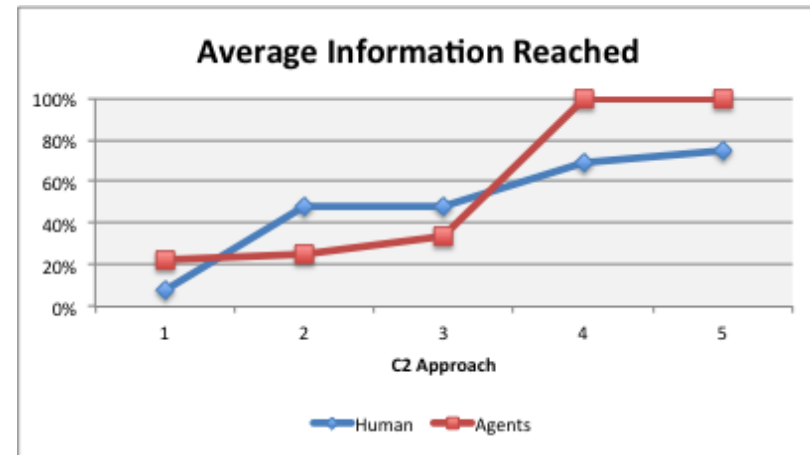
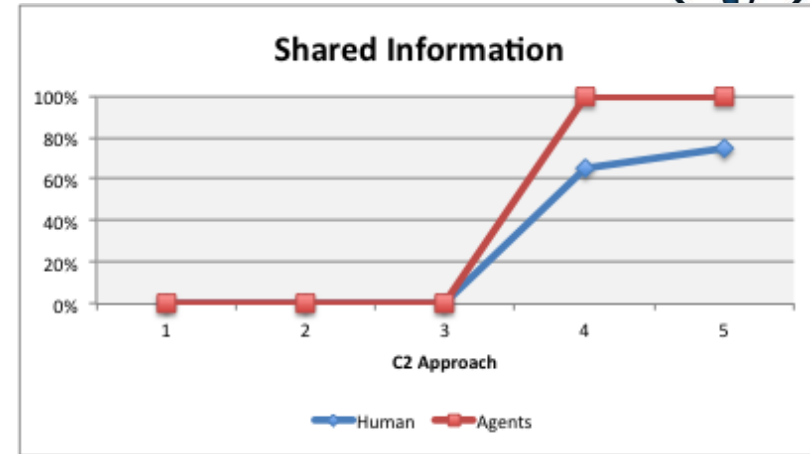
## Measurements

- **Information Domain:** Relevant Information Reached, Shared Relevant Information
- **Interactions/Social:** Interactions Activity, Average Network Reach
- **Cognitive:** Time of First Correct ID, Number of Correct IDs, CSSync
- **MoM:** Effectiveness and Efficiency (Time and Effort), Maximum Timeliness

# Results



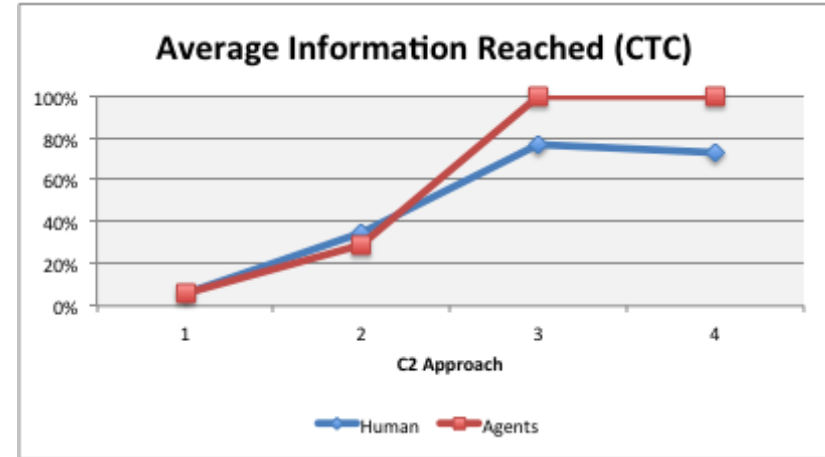
- Information Domain
  - **Shared Information:**  
Similar for humans and agents. Agents reached maximum scores in 4 and 5.
  - **Avg Info Reached:**  
Increasing trend both in Human and in Agent runs.  
Humans were slightly better in 2 and 3. Agents reached maximum values in 4 and 5.



# Results



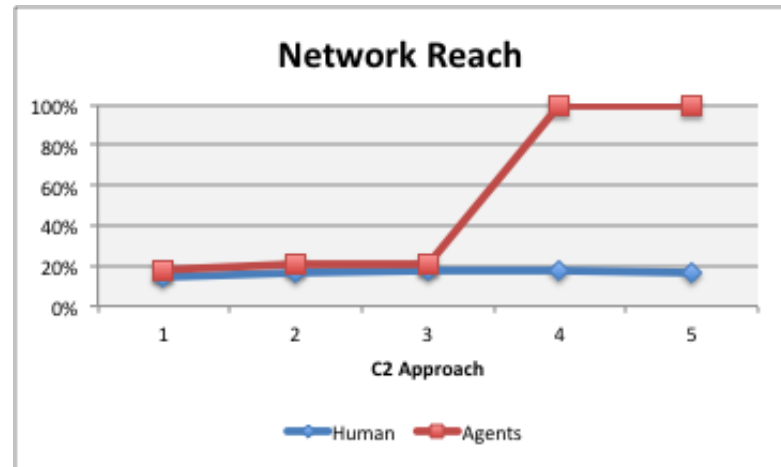
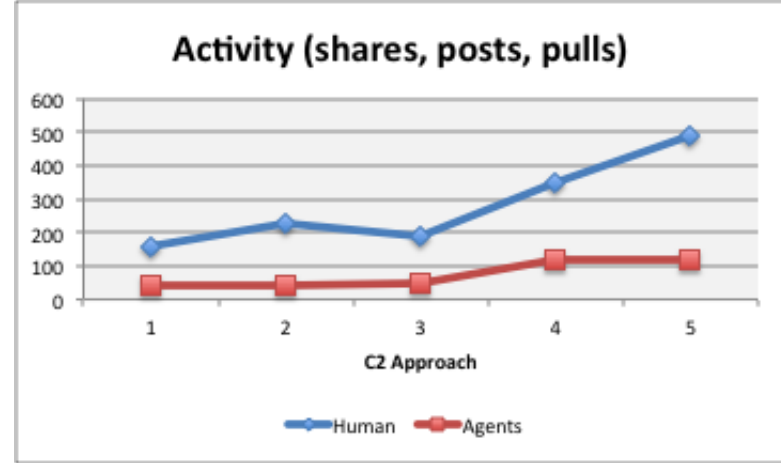
- Information Domain
  - CTC is a key-role in 1 to 4. Not applicable in 5.
  - Measured CTC information reached.
  - Similar trend (increasing) in human and agent runs.
  - Agents obtained 100% for 3 and 4, while Humans were below 80%.



# Results



- Interactions / Social Domain
  - Humans were in overall more active than agents (however this did not result in more information being shared).
  - Agents behave the same in 4 and in 5.
  - Significant increase in network reach for agents in 4 and 5 (100%). Humans stay below 20%.





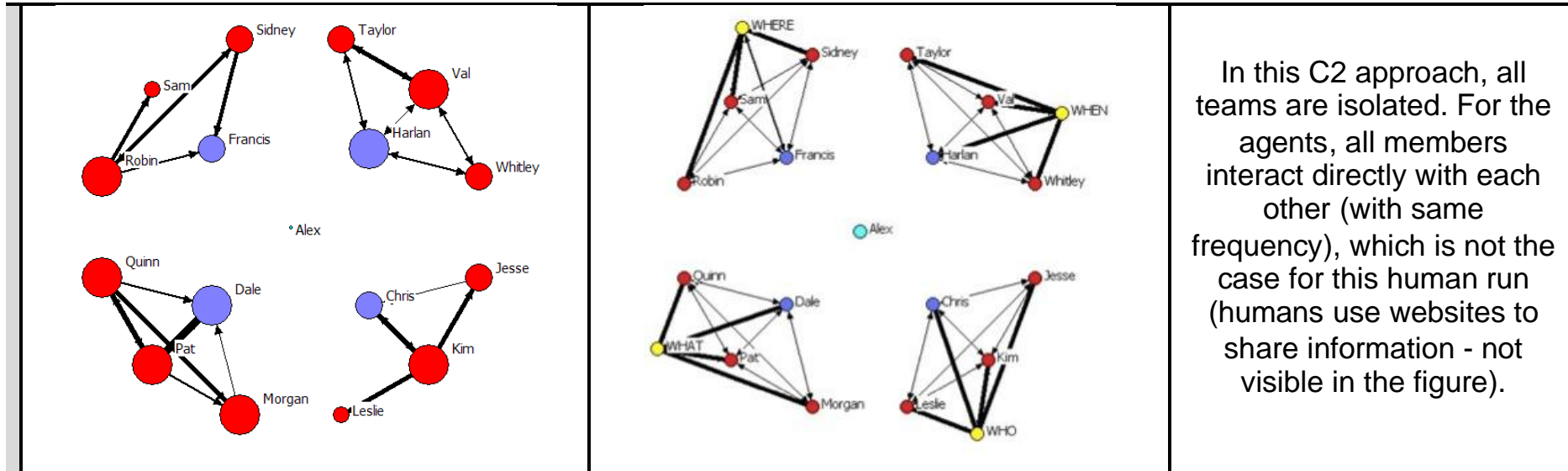
# Results

## Sociogram: Conflicted C2

### Human Runs

### Agent Runs

### Comment:



In this C2 approach, all teams are isolated. For the agents, all members interact directly with each other (with same frequency), which is not the case for this human run (humans use websites to share information - not visible in the figure).

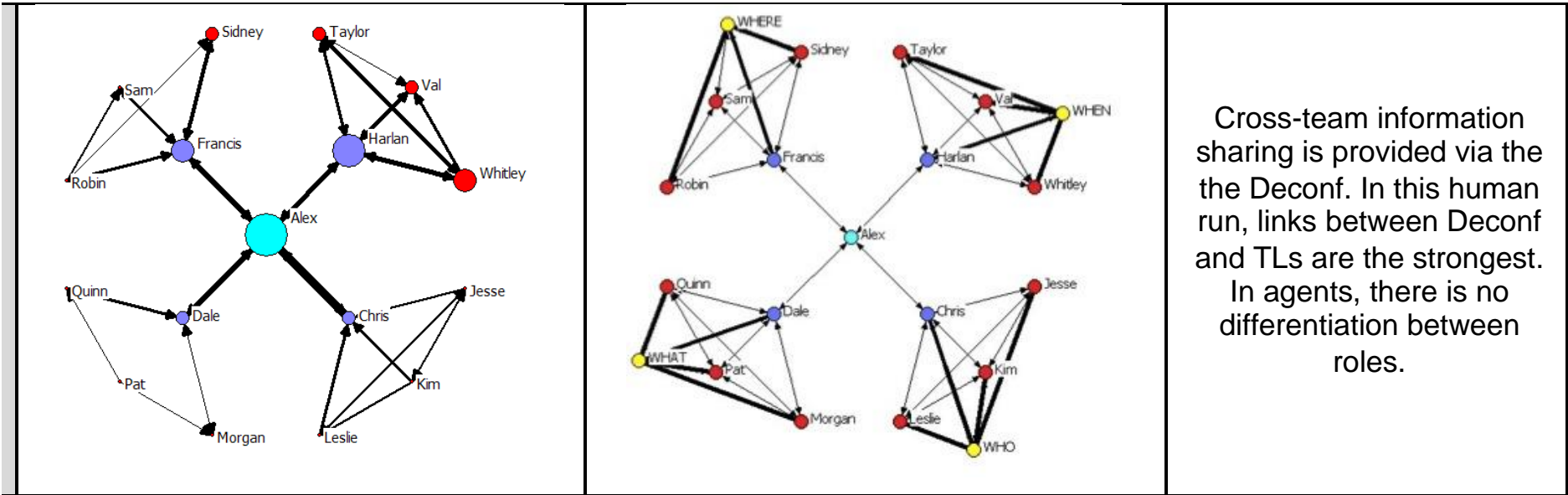
# Results

## Sociogram: De-Conflicted C2

### Human Runs

### Agent Runs

### Comment:

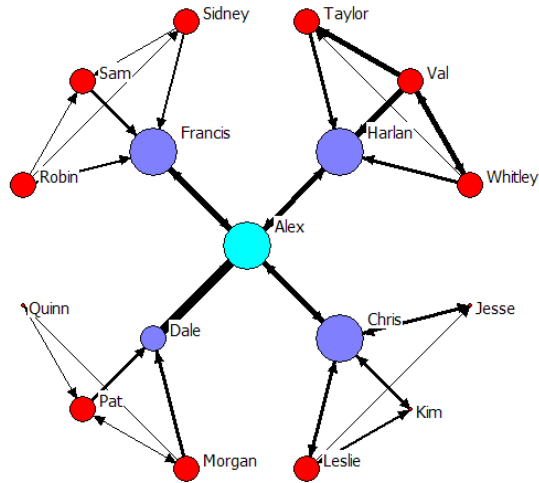


Cross-team information sharing is provided via the Deconf. In this human run, links between Deconf and TLs are the strongest. In agents, there is no differentiation between roles.

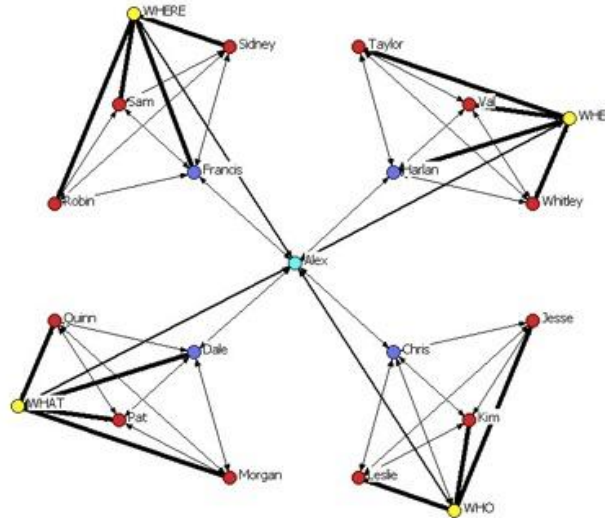
# Results

## Sociogram: Coordinated C2

### Human Runs



### Agent Runs



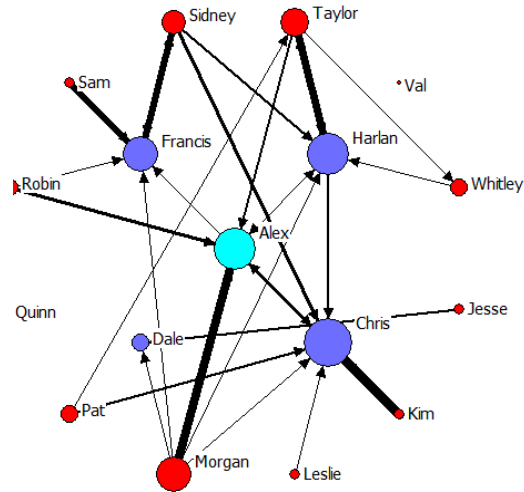
### Comment:

The CTC role is created with access to all websites. In this human run, links between the CTC and TLs are the strongest. In agents, there is no differentiation between roles.

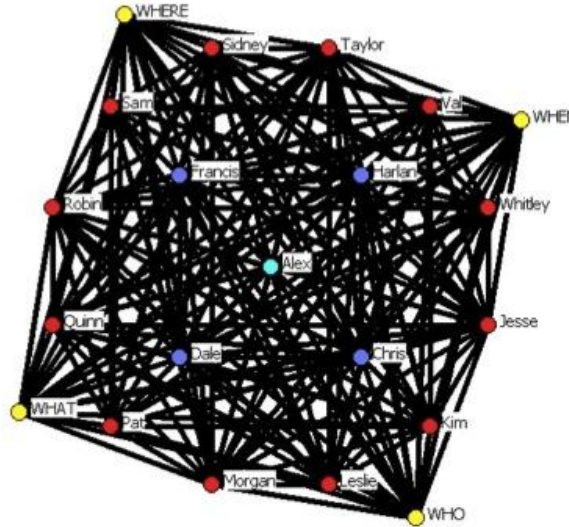
# Results

## Sociogram: Collaborative C2

### Human Runs



### Agent Runs



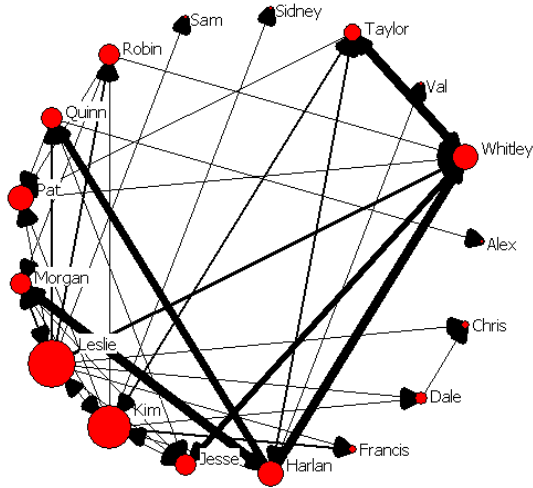
### Comment:

A fully connected network is provided in a three-level organisational structure. Humans differentiated their interactions based on node role, while agents were fully connected (any member reached all members).

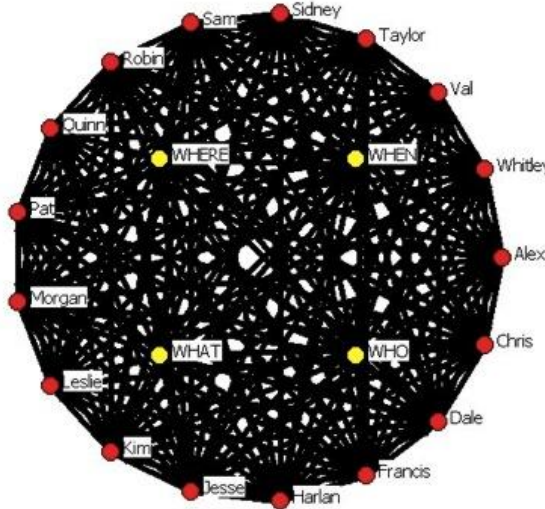
# Results

## Sociogram: Edge C2

### Human Runs



### Agent Runs



### Comment:

A fully connected network is provided in a flat organisation (no pre-assigned roles). Humans interacted with each other (although websites was their preferred method to post and pull information) while agents were fully connected (any member reached all members).

# Results



- Cognitive Domain

- Correct IDs:

Low values for 1, 2 and 3. Increase for 4 and 5.

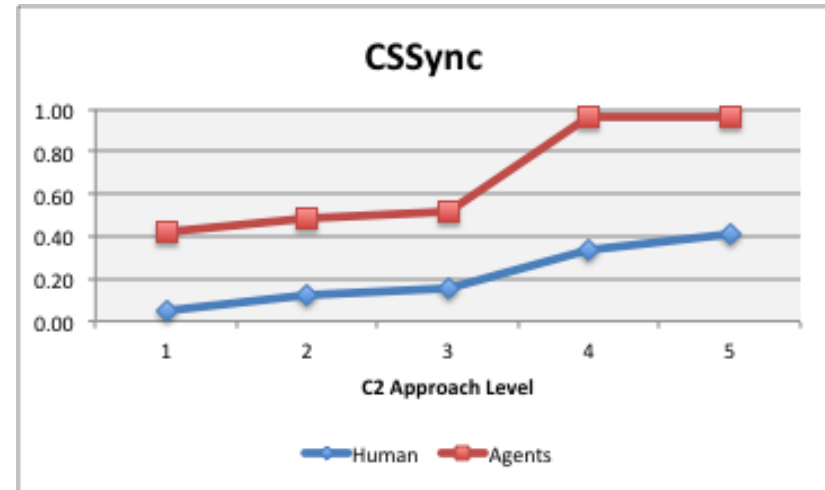
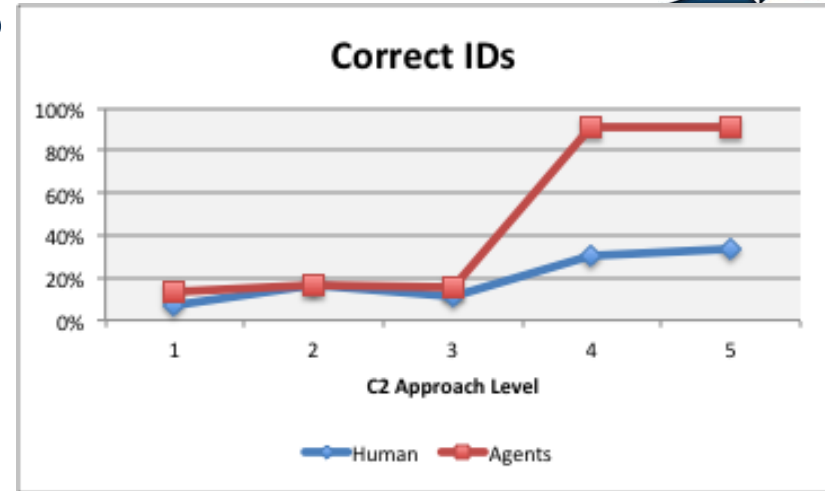
Agents reached ~90% in 4 and 5.

Humans stayed below 40%

- CSSync (Manoso and Moffat, 2011):

Increasing trend from 1 to 5.

Agents always obtained higher values than humans.



# Results

- MoM

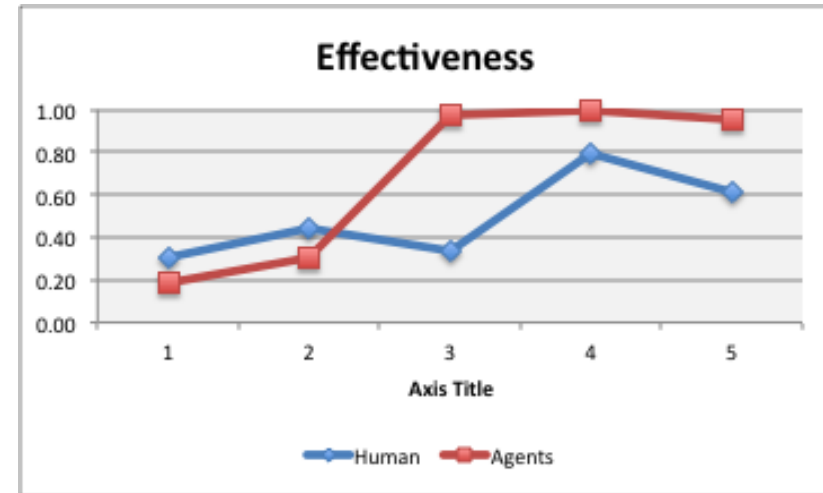
- Effectiveness: (note: each approach used its own criteria)

Low effectiveness for 1 and 2 in both Agent and Human runs.

In 3: score in humans decreased (unexpected result), while agents increased to almost 100% (maintained till 5).

Best approach for humans = 4

Best approach for agents = 4 (3 and 5 will similar score)





# Results

- MoM

- Time-Efficiency:

Low scores in 1 and 2.

Agents and humans increase in 3 and in 4 and then decrease in 5.

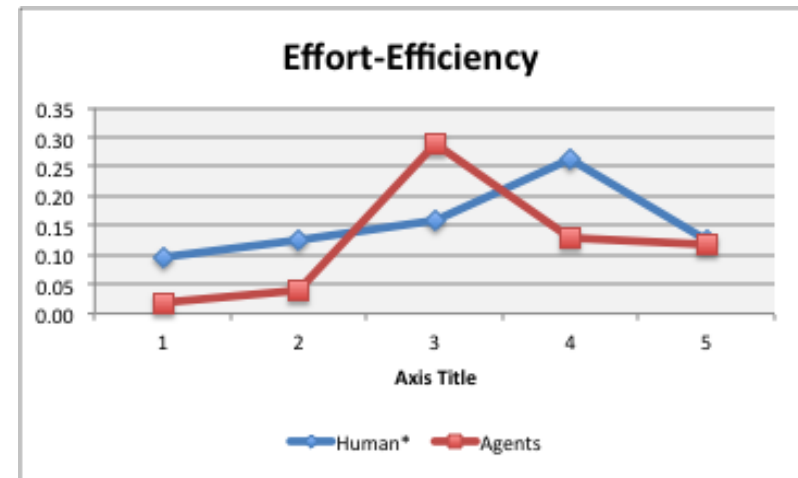
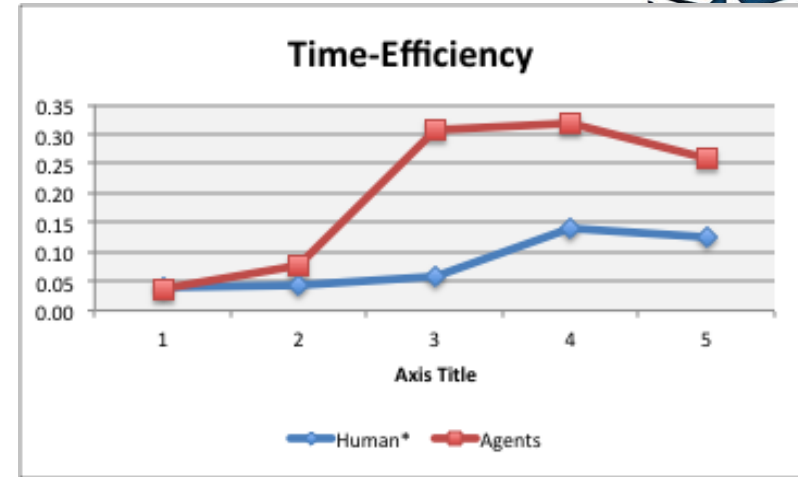
Agents are better than humans.

- Effort-Efficiency (activity was accounted as cost):

In agents, the best is 3.

In humans, the best is 4.

More effort spent in 5 resulted in decreased effort-efficiency.





# Conclusions



- Similar trends were observed between human and agent runs across the dimensions and C2 Approaches, but improvements should be made in modelling the C2 Approaches and in the agent's logic.
  - Our Collaborative model for all agent runs ended located at the top-right of the approach space (i.e., Edge space), but it should be positioned in areas below that region. Future work should address this.
- In general agents were more successful than humans in:
  - Sharing information
  - Reaching higher values for correct IDs and CSSync in Collaborative and Edge approaches

# Conclusions (2)



- In the social domain differences were noted:
  - The agents' behavior was homogeneous and regular.
  - The human behavior was heterogeneous and irregular.

This human richness and diversity results in unique outcomes in each human trial that has not yet been captured in the agents.
- On the MoMs:
  - Agents were highly successful in 3, 4 and 5. Humans were highly successful at 4 and moderately successful at 5.
  - It was noted that a sufficient condition for success in agents was for them to have access to the necessary information to determine the solution, a condition that is not valid for most human subjects.

# Recommendations for abELICIT



## On the Agent's Logic:

- In "Social Processing" add a perceived ranking and trust-level towards other team members that influence how sharing occurs. This ranking is dynamic and built throughout the run.
- In "Information Processing" build a ranking for information that is perceived as having high relevance. The ranking may be a function of internal perceived value and source.
- Add a probability value to share and/or post highly relevant information more than once.
- In "Select Message", give priority to information received from trusted/highly-ranked sources.
- In the "Awareness Processing", make it dependent on numerous factors such as arrival of new information, source (trust and ranking) and propensity to change (likelihood to change the ID based on new information received).

# ELICIT

**THANK YOU FOR YOUR ATTENTION !**

Marco Manso (SAS-085 Member), Portugal ([marco@marcomanso.com](mailto:marco@marcomanso.com))

Mary Ruddy (Meristic), USA ([mary@meristic.com](mailto:mary@meristic.com))

18th ICCRTS  
C2 in Underdeveloped, Degraded and Denied Operational Environments  
June 19-21, 2013 - Alexandria, Virginia, U.S.A.