The Unreasonable Effectiveness of Artefacts and Documentation: An Exploration of Consensus Using Multi-Agent Simulations in a Two-Team Configuration

Johannes S. Vorster[1][a] and Louise Leenen[1,2][b]
[1]*Department of Computer Science, University of Western Cape, South Africa*
[2]*CAIR, South Africa*

Keywords: Consensus, Consensus Simulation, Stochastic Simulation, Synchronization, Multi-Agent Simulation, Artefacts, Documentation.

Abstract: Documentation and artefact generation is an essential part of business processes. This paper explores the use of artefacts as a means of reaching consensus through the use of Multi-Agent Simulations. In particular we investigate the time to reach consensus with and without the use of artefacts and show the efficiency of artefacts as a means of facilitating consensus, perhaps more importantly, to create efficient consensus processes in the face of difficult organizational communications channels. We found that polyarchies are highly efficient at consensus formation, but are not realistic for larger organizations. For these organisations a small team that facilitate consensus formation is nearly as efficient. The introduction of artefacts significantly improve consensus formation in situations where intra-team communications causes delays in consensus formation.

1 INTRODUCTION

1.1 Organizational Context

Over the past two decades, the evolution of organizational structures and project strategies has been a key discussion, driven by technology companies facing rapid technological advancements, shifting competitive landscapes, and changing customer expectations. As innovation accelerates, businesses must adapt their organizational setups and project delivery methods to remain agile and responsive.

Notable transformations include changes in team composition (Reagans et al., 2016), shifts from hierarchical to lateral organizations (Keupp et al., 2012), and network-like organizational structures (Chang and Harrington, 2000). These changes often lead to faster delivery and reduced resource expenditure, resulting in better investment returns (Will et al., 2019).

Project complexity, marked by difficulties in reaching consensus, is a major cause of delays and failures (Al-Ahmad et al., 2009; Whitney and Daniels, 2013; Kian et al., 2016; Waheeb and Andersen, 2022). During consensus-building, diverse teams must share knowledge, reconcile differences, and collaboratively develop solutions (Eden and Ackermann, 2010; Cheung et al., 2016).

1.2 Consensus Models

In formal settings, particularly for predicting economic outcomes, human consensus is often achieved through Delphi processes, where participants undergo multiple rounds of anonymous feedback to converge on a consensus. Alternative approaches include qualitative studies on social networks (Carter et al., 2015; Jones and Shah, 2016), detailed interviews (Roselló et al., 2010), and computational models (Yan et al., 2017). This paper employs computational models while incorporating aspects of Delphi decision-making.

Reaching consensus in Multi-Agent Systems (MAS) is complex, with research spanning social sciences, economic model simulations, and MAS consensus algorithms. Social sciences have explored crowd and voter behavior for decades, starting with Dunbar (1998) on the 'social brain,' followed by Stocker et al. (2001) on social information exchange, and Leishman et al. (2009) on consensus group for-

[a] https://orcid.org/0000-0001-6452-4186
[b] https://orcid.org/0000-0002-9212-550X

Vorster, J. and Leenen, L.
The Unreasonable Effectiveness of Artefacts and Documentation: An Exploration of Consensus Using Multi-Agent Simulations in a Two-Team Configuration.
DOI: 10.5220/0012785300003758

In *Proceedings of the 14th International Conference on Simulation and Modeling Methodologies, Technologies and Applications (SIMULTECH 2024)*, pages 313-323
ISBN: 978-989-758-708-5; ISSN: 2184-2841

mation, with further insights from Gilbert (2010).

MAS consensus algorithms focus on high-speed applications, as exemplified by Amirkhani and Barshooi (2022). Chang and Harrington (2004) proposed a MAS framework for modeling organizations to address economic questions, and Will et al. (2019) explored the impact of organizational structure on innovation project selection, highlighting polyarchy, hierarchical, and hybrid forms. This paper uses fully connected polyarchies, as discussed in Vorster and Leenen (2023a).

Subversive agent behavior has been studied across various domains, including psychology (McDowell, 2002), business (Manky and Dolores, 2022), politics (Barnes and Prior, 2009), and espionage (Evans and Romerstein, 2012).

This work aligns with MAS in Computational Economics (Tesfatsion and Judd, 2006) and Consensus formation (Vorster and Leenen, 2023a).

1.3 Social Settings

In social networks, the topology and geospatial distribution of the population, along with their interconnectedness, are crucial for shaping opinion dynamics, such as voter views. Accurate spatial distribution models are essential for modeling community behavior (Amblard and Quattrociocchi, 2013).

The earliest computer-based consensus model was developed by Johnson and Feinberg (1977), where crowd members seek support from subsets to sway the crowd towards a specific action. Consensus is achieved by reducing opinion variability.

Subsequent studies used complex adaptive systems to examine virtual team behavior, highlighting the importance of team cognition, trust, cohesion, and conflict (Curşeu, 2006). Social group formation occurs in two phases: initial interaction among closely related individuals and broader community interaction (Leishman et al., 2008).

The impact of network topologies on consensus has been explored using various graphs, including Erdős-Rényi random graphs (Erdős et al., 1960; Amblard et al., 2015), Watts-Strogatz small-world networks (Watts and Strogatz, 1998a), and Barabási-Albert scale-free networks (Barabási et al., 2000; Leskovec and Mcauley, 2012), providing insights into consensus dynamics in social and political contexts.

Pro-social behavior propagates through social interactions (Christakis and Fowler, 2013; Keizer et al., 2013; Tsvetkova and Macy, 2014). Dunbar's 'Social Brain' hypothesis suggests group size in apes and humans is a function of language use rather than grooming (Dunbar, 1998). Stocker et al. (2001) tested this hypothesis by simulating individual influence and idea communication, showing critical connectivity levels are needed for consensus.

Michalski et al. (2022) explored the impact of social network connectivity on consensus with two options. Agents' probabilistic beliefs evolved through interactions, using various network topologies (complete, cycle, wheel, Erdős-Rényi, Watts-Strogatz, Barabási-Albert). They found that complete networks led to the fastest consensus, followed by wheel, scale-free, random, cycle, and small-world topologies.

1.4 Subversive Agents

Research shows that constructive task conflict can enhance team decision-making and performance, especially in complex tasks involving uncertainty or subjective factors (Bradley et al., 2015; Enyinda et al., 2022; Kirschner et al., 2008; Badke-Schaub et al., 2010; Paletz et al., 2017). Effective communication, collaboration, and social skills are essential for leveraging task conflict benefits (Wu et al., 2017; Hirvonen, 2019). Properly managed discourse, with issues explicitly articulated and addressed, is crucial (Holmes and Marra, 2004).

Xie et al. (2011) demonstrated that a small fraction p of committed agents advocating an opposing opinion can shift the majority opinion in a population. Using a 2-option Erdős-Rényi random graph model, they showed that when the committed fraction exceeds a critical threshold of about 10% ($p_c \approx 10\%$), the time T_c for the population to adopt the committed opinion significantly decreases.

Iacopini et al. (2022) identified three behavioral regimes in group-based consensus processes, introducing a susceptibility variable β. Small increases in β initially enhance the committed minority's influence (regime 1). At a critical β value, the entire population adopts the minority's views (regime 2). Further increases isolate the committed minority, preserving the original group's beliefs (regime 3).

Vorster and Leenen (2023b) used stochastic MAS simulations to study subversive agents' impact on consensus in project teams. These agents aim to delay project delivery by influencing views and decisions. They found that expanding options or polarizing the group significantly delays consensus. Coordination among subversive agents deeply influences outcomes, and even a small minority can substantially prolong consensus times (Vorster and Leenen, 2024).

1.5 Organizational Structure

Chang and Harrington (2004) described a scheme for modeling organizations using multi-agent systems (MAS) to address economic questions. Will et al. (2019) later examined the role of organizational structure in innovation project selection, highlighting three forms: polyarchy, hierarchical, and hybrids. In a polyarchy, team members are fully connected and can communicate across hierarchical boundaries.

Network topology significantly influences consensus time. Michalski et al. (2022) found that polyarchies yield the quickest consensus compared to random (Erdős et al., 1960), small-world (Watts and Strogatz, 1998b), and scale-free networks (Barabási et al., 2000).

Will et al. (2019) also studied the economic impact of organizational structure on risky project selection, arguing that decision-making is influenced by organizational structure, not just team evaluative abilities. They used a mathematical model to analyze project selection dynamics, showing that hybrid structures can have error management side effects. Their work builds on earlier economic models by Sah and Stiglitz (1984, 1988). Similarly, Sáenz-Royo and Lozano-Rojo (2023) extended Chang and Harrington (2004)'s work using simulated structures to investigate innovation project selection.

Motivation for This Paper

Our earlier statistical analysis and MAS simulations of consensus processes has shown the importance of MAS in the study of consensus formation within organizations (Vorster and Leenen, 2023a). That study showed that within Polyarchies (Figure 1) the introduction of artefacts (documentation) can increase the formation of consensus by 30%. The formation of consensus in social settings has been studied using various 'random' networks such as Erdős-Rényi random graphs, Watts-Strogatz small-world networks, and Barabási-Albert scale-free networks, as mentioned earlier. However, these networks do not reflect the reality of organisational structures.

Indeed, a Polyarchy is the consensus equivalent of the spherical cow in physics. To study all organisational hierarchies is also not viable, and thus we select and study a specific organizational configuration consisting of a group that provides requirement specifications and another group that provides the implementation (or realization) of the requirements.

In this paper we investigate the formation of consensus across two teams where the one team is responsible for providing the requirements and the other team needs to provide an implementation. We are interested in the consensus processes across these teams, given that both the teams form Polyarchies on their own, see Figures 4, 5 and 6. The two teams can communicate with each other through a small intermediary team consisting of members of both teams (usually the business analyst and solutions architect). In this paper we present results from simulating these organizational configurations and control the interconnectedness of these two groups, as well as the impact of artefacts (documentation) on the ability to reach consensus quickly on a large set of issues.

As can be expected, the bottleneck between the two groups are pivotal to how fast consensus can be reached across the two groups and there is no surprise in finding that the configuration depicted in Figure 5 with one person representing the requirements team talking to one person representing the implementation team is the worst case scenario. How much will this worst case improve if the requirements team generate artefacts that accurately and completely depict the requirements? And similarly for the implementation team.

Since the authors have extensive experience working for, working with, and consulting to such large organizations, they have often seen that a thorough documentation process is seen as a waste of time. We would like to offer this research to at least theoretically prove the importance of thorough artefact generation.

2 METHODOLOGY

In this section we discuss the essential aspects of the simulator focussing only on relevant topics to the below experimental setup.

2.1 Teams and Topics

The members of the teams are agents in the MAS. To be generic for later simulations with multiple teams, we will denote the specification team as team a and the implementation team as team b. There are ${}_aN$ agents in the specification team, and ${}_bN$ agents in the implementation team. ${}_aN \ll {}_bN$. In general we will use a prefix small a and b to denote the teams in variables. Each agent keeps track of a number of topics. The specification team has to consider and agree on ${}_a\mathcal{B}^{\max}$ topic and the implementation team ${}_b\mathcal{B}^{\max}$. With ${}_a\mathcal{B}^{\max} \ll {}_b\mathcal{B}^{\max}$ because the implementation details and topics to agree on are, as a rule, much larger than the number of requirements topics. The first ${}_a\mathcal{B}^{\max}$ topic are the same for both teams. That is,

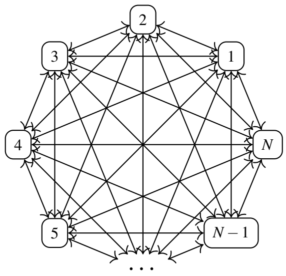

Figure 1: Polyarchy

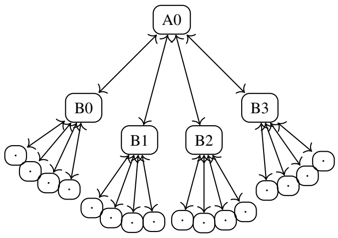

Figure 2: Hierarchical.

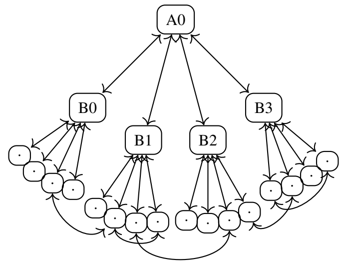

Figure 3: Hybrid.

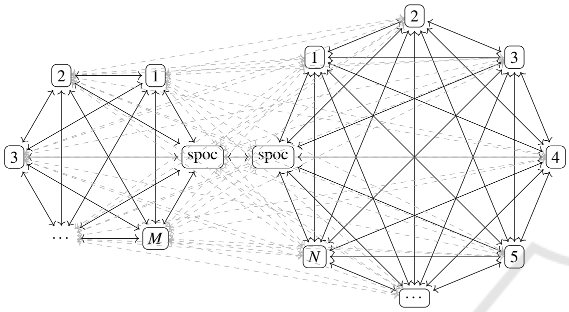

Figure 4: Polyarchy Binary configuration.

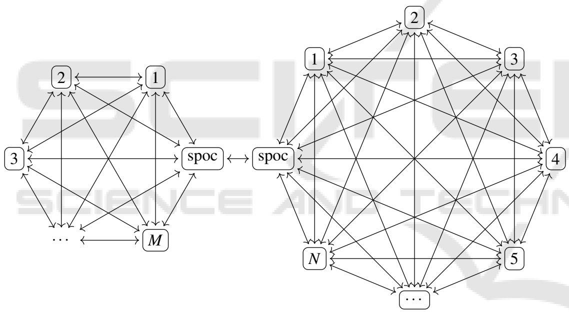

Figure 5: Loose Binary configuration.

the implementation team must agree and reach consensus with the specification team on the specification.

2.2 Artefacts

The requirement specification artefact encodes the topics sequentially one-to-one with the topics agents keep track of. The specification contains ${}_a\mathcal{C}^{\max}$ topics and both the specification team and the implementation team will need to reach consensus on these specifications. It is so arranged and agreed that if an artefact contain fewer topics than what agents discuss, that is, if ${}_a\mathcal{C}^{\max} < {}_a\mathcal{B}^{\max}$, then the first ${}_a\mathcal{C}^{\max}$ topics of the *agents* coincide with the artefact's topics. In this arrangement it is possible to model incomplete artefacts because artefacts can contain fewer topics than what is needed to reach consensus.

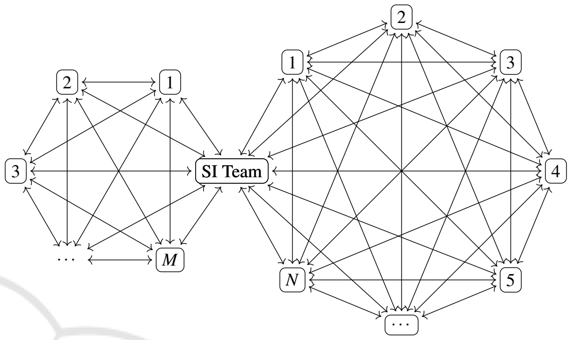

Figure 6: Small Interaction Team Binary configuration.

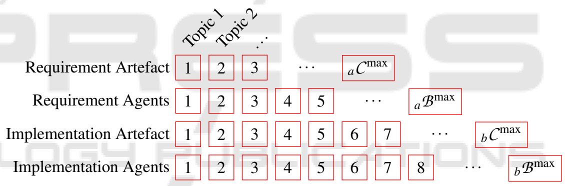

Figure 7: Topics across agents and artefacts.

Figure 7 provide a visual representation of the situation in terms of topics, with

$$ {}_a\mathcal{C}^{\max} \leq {}_a\mathcal{B}^{\max} < {}_b\mathcal{C}^{\max} \leq {}_a\mathcal{B}^{\max}. $$

Topic i is the same across all agents and artefacts (in this specific simulation), but need not be in general. Agents will set up meetings using their connectivity graph, and will try and resolve topics until consensus has been reached between all agents with each other, and agents with the artefacts, on all topics.

2.3 Time, Duration, and Meetings

The simulation takes into account calendar time and meeting duration. A limited number of topics can be discussed per meeting. Agents will set up time to meet (following earliest available time rules) with each other if there are disagreements in their views. Each meeting lasts a fixed time (30 minutes and 16 meetings per day). Within that meeting time-slot, the number of topics that can be discussed and resolved are determined stochastically (at least one and maximum ten). Three outcomes are possible for such a discussed topic: (a) both agree on a new position in the middle of their earlier positions (compromise consensus), (b) the first agent convinces the second of its view, and (c) visa-versa.

Agents can also choose to interact with an artefact within each 30 minute time-slot. In that case the interaction is similar to that of another agent, in that the agent will select a random number (at least one, maximum ten) of topics where it is in disagreement with the artefact. The three outcomes for each topic of interaction are (a) the reader is partially convinced by the artefact (its view changes) but it also updates the artefact to the new view, (b) the reader is completely convinced and fully internalizes the view expressed in the artefact, and (c) the reader feels the artefact is in error and corrects it by modifying it with the agent's current view on the topic.

Agents can only set up meetings with other agents within their connectivity network. This connectivity network is modelled as a directed graph with agents as nodes and edges represent the ability to set up a meeting and have a discussion. Agent i denoted by node ν_i can meet with agent j denoted by node ν_j if there is an edge within the graph from ν_i to ν_j. Agent i, node i and ν_i are all representations of the same thing.

The simulation initializes each agent with a random view on each topic. The simulation allows the agents to set up meetings and discuss topics, or interact with an artefact. One activity is allowed per time-slot. The simulation counts time in measures of time-slots. So that in the below graphs a time measure of 100 would mean that it is the 100$^{\text{th}}$ time-slot from the start of the simulation. Simulations with the same configuration are repeated numerous time (sometimes up to 20000) to ensure statistically significant and smooth results.

2.4 Measuring Consensus

In agent-based modelling, consensus between an agent and the rest of its connected group on a specific topic k is often expressed as the sum of differences between the views b^k on topic k of agent i and all other agents j, (see e.g. Wei et al. (2021)),

$$u_i^k = -\sum_{j=1}^{N} \delta_{ij}(b_i^k - b_j^k), \quad i \in \{1, 2, \ldots, N\}.$$

Where δ_{ij} is a cost factor and we use $\delta_{ij} = 1$ if i can schedule meetings with j and $\delta_{ij} = 0$ otherwise. We however need to use absolute differences and sum over all agents, all artefacts and all topics, for a discussion see Vorster and Leenen (2023a).

If I_ν is the set of all agents and I_A is the set of all artefacts, then we define a measure of consensus for agents with each other $(i, j \in I_\nu)$ and agents with artefacts $(i \in I_\nu, p \in I_A)$ on a specific topic k as

$$u_{ij}^k = \delta_{ij}|b_i^k - b_j^k| \text{ and } u_{ip}^k = \delta_{ip}|b_i^k - c_p^k|$$

which leads to an overall measure of consensus for an agent i with agents and artefacts it has contact with as

$$u_i^k = \sum_{j \in I_\nu} \delta_{ij}|b_i^k - b_j^k| + \sum_{p \in I_A} \delta_{ip}|b_i^k - c_p^k|. \quad (1)$$

That is, the level of consensus that an agent i has relative to the rest of the group on a topic k, is the sum of absolute differences between that view b_i^k and the views on the same topic for all other agents, b_j^k that it is connected to, $(\delta_{ij} > 0)$, as well as the same measure for that topic in all artefacts, c_p^k, to which it has access.

We can now define an overall level of consensus over all topics, agents, and artefacts as

$$u = \sum_{i \in I_\nu} \sum_{j \in I_\nu} \sum_{k=1}^{k_{ij}^{max}} \delta_{ij}|b_i^k - b_j^k| + \sum_{i \in I_\nu} \sum_{p \in I_A} \sum_{k=1}^{k_{ip}^{max}} \delta_{ip}|b_i^k - c_p^k|. \quad (2)$$

Some care needs to be taken with k since it runs from 1 to some maximum number of topics which are dependent on the agents and artefacts being evaluated, see Figure 7. For example, if agent i is from team a and agent j from team b, then $k_{ij}^{max} = min({}_a\mathcal{B}^{\max}, {}_b\mathcal{B}^{\max})$. In reality (in the software code) each agent keeps a list of its views, as does artefacts, so that a direct sum of absolute differences only runs to the minimum length of the two applicable lists.

2.5 Time and Effort to Reach Consensus

Each agent keeps track of what they do in each time-step and they record this in their dairy so that d_i^t is the entry for agent i's diary at time t. It records the agent number it met with, or 'z' if it did nothing at time t. The effort, $e^{\max}$, to reach consensus is the sum of all actions taken by all agents, that is

$$e^{\max} = \sum_{t=1}^{t^{max}} \sum_{i=1}^{N} busy(d_i^t), \; busy(d_i^t) = \begin{cases} 0 & d_i^t = \text{'z'} \\ 1 & \text{otherwise,} \end{cases} \quad (3)$$

where $t^{\max}$ is the total time it took to reach consensus. Since each agent will always take some action if an action is available, the simulation terminates at time t^{max} when $d_i^t = \text{'z'} \; \forall i \in I_\nu$. That is, when no agent takes any actions any more, the simulation stops. Both $e^{\max}$ and $t^{\max}$ will be determined by simulation and will differ on each stochastic simulation. For each configuration a large number of simulations were run to determine $e^{\max}$ and $t^{\max}$ accurately.

2.6 Meeting Efficiency

The last topic and variable of interest here are the efficiency of meetings. A meeting will discuss a random number of topics from one to nine, with an average of five in our simulations. As we will see shortly, this average of five topics per meeting is only realistic at the start of the project, and as topics gets resolved, it starts to happen that there are no longer five topics of discord left between agents. When this happens, the meeting efficiency starts to drop. If $\bar{z}(t)$ is the observed average number of topics discussed in meetings at time t and $\bar{z}^{max}$ is the maximum expected number of topics, then we define a measure of meeting efficiency at time t as

$$e(t) = \bar{z}(t)/\bar{z}^{max} \tag{4}$$

Although it is possible to measure the effectiveness of individual agents, here we are interested in the effectiveness of meetings by the team. In the below graphs and results we plot the meeting efficiency of team a, the requirements team, independently from the meeting efficiency of the implementation team b.

2.7 Mathematical Model Summary

The following are important concepts for the remainder of the paper:

- u – measure of the overall consensus in the group and is the pairwise sum over all differences in views over all members and artefacts;
- u_i^k – measure of the difference on a spesific topic (k) between all group memers and a specific member i, and is the pairwise sum of differences in view between i and all other members and artefacts;
- t^{max} – the time to reach consensus for a specific scenario (group-size, artefacts, problem-size), which is averaged over many runs;
- e^{max} – the effort to reach consensus (which is the sum of all meetings) to reach consensus.

3 EXPERIMENTAL RESULTS

In this sections the basics of the simulation and important measurements are discussed together with results that can be used as a baseline for later results.

3.1 Experimental Configuration

The primary variables that are changed are the group size, from 2 to 20, and the presense of artefacts. Variables that can change but are kept constant are the number of topics for discussion in meetings and in artefacts. It was shown earlier (Vorster and Leenen, 2023a) that the time and effort to reach consensus are linear with the number of topics, and thus for these experiments we kept the topics constant at 50 since it does not play a role in the results we are discussing in this paper. Number of topics discussed per meeting is randomized from 1 to 10. The pseudo-Python code for meetings between agents i and j is:

```
random.shuffle(topics)
issuesToDiscuss=randint(1,11)
for k in topics:
  if agent[i].view[k]==agent[j].view[k]:
    continue
  rnd = randint(0,3)
  if (rnd==0):
    val = int((agent[i].view[k]
        + agent[j].view[k]))/2.0)
    agent[i].view[k]=agent[j].view[k]=val
  if (rnd==1):
    agent[j].view[k] = agent[i].view[k]
  if (rnd==2):
    agent[i].view[k] = agent[j].view[k]
  issuesToDiscuss-=1#
  if issuesToDiscuss<=0: return
```

3.2 Polyarchies

As a first experiment, and as a way to construct a baseline for comparison, we start with two teams; team a with seven team members (the requirements specification team); and team b, the implementation team, with 14 members. In the next section we generalize team size. The two teams are allowed to set up meetings with members in the opposite team as they see fit. That is, the organization acts like a polyarchy, see Figure 4.

Stochastic simulations were repeated 20000 times. The averages for all the relevant variables were computed. Figure 8 shows the results using the consensus measure (top) and the same measure but using $\log u$ (middle). The consensus measure decreases exponentially (linear on log-scale) over time until consensus is reached. A mean time of 279 ($\sigma = 10.7, n = 20000$) was recorded.

The introduction of artefacts, a requirements specification artefact and an implementation specification

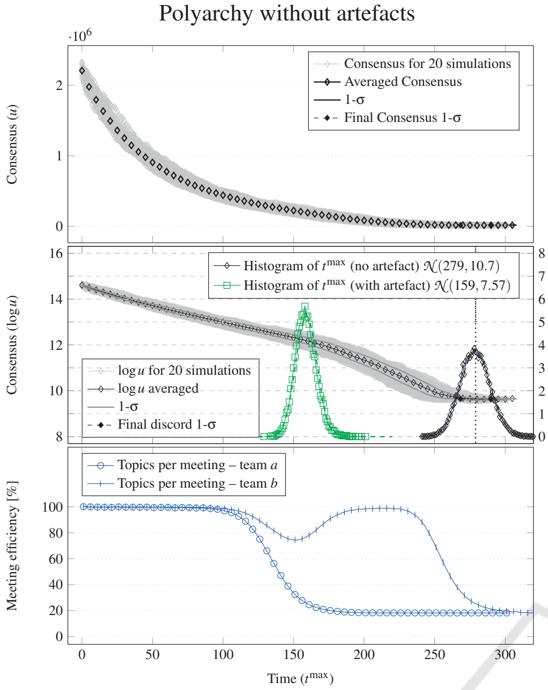

Figure 8: (Top) Various simulations of the 7-group showing the consensus measure over time. (Middle) The same data as in top graph, but now using $\log_e$(consensus). Histogram of the time it takes to reach consensus over many such runs ($\mu = 279$, $\sigma = 10.7$, n=20000) and Normal and Lognormal fits to the histogram data. The central (green) histogram is the same as that of Figure 9 transposed here for comparison of how much artefacts shifts the time to reach consensus. (Bottom) Meeting effectiveness graphs for the two groups.

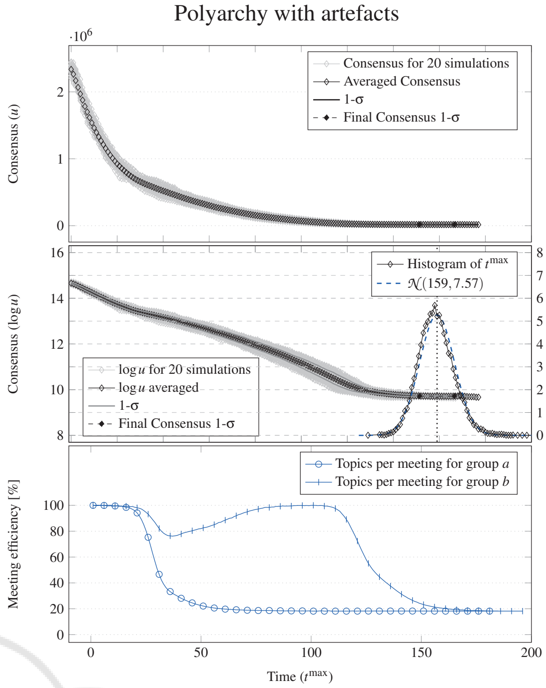

Figure 9: (Top) Various simulations of the 7-group showing the consensus measure over time. (Middle) The same data as in top graph, but now using $\log_e$(consensus). Histogram of the time it takes to reach consensus over many such runs ($\mu = 159$, $\sigma = 10.7$, n=20000) and Normal and Lognormal fits to the histogram data. (Bottom) Meeting effectiveness for the two groups.

artefact, led to significant improvements in time to reach consensus, see and compare Figure 9, with a mean time to reach consensus now shifting to 159 ($\sigma = 7.57, n = 20000$). This is an improvement of 43.0%, or, conversely, not using an artefact will increase the time to reach consensus by 75.5%. The middle plot of Figure 8 shows both the histogram for time to reach consensus with and without an artefact for a visual reference of how significant this improvement is.

The productivity of meetings in the two teams are also interesting and can be divided into four phases, Figures 8 and 9 (bottom). During phase 1, both teams are engaged in trying to reach consensus on the requirements topics and meetings are highly effective, since there are significant numbers of topics to discuss. In phase 2, team a's meeting efficiency steadily reduces as the meetings become more inefficient since there are no longer as many topics available to make every meeting effective, for both teams. During phase 3 the team dynamics shifts, with team a's meetings being very ineffective, mostly dealing with one-topic discussions while team b's meetings remain highly efficient as they are trying to reach consensus on the implementation specification topics, taking specific issues back to team a as they surface. Lastly, in phase 4, the topics for discussion for the implementation also reduce in number and meeting efficiency decreases, until close to when consensus is reached, when all meetings, in all teams are highly inefficient and only one-topic meetings occur.

These four phases are present in both cases with and without artefacts. When artefacts are present, the phases are accelerated, see Table 1. The largest phase contraction happens in phase 1, with a significant reduction in time due to the presence of an artefact. Per-

Table 1: Length of phases in the meeting productivity measure.

	Without artefact	With artefact
Phase 1	100	20
Phase 2	80	60
Phase 3	50	30
Phase 4	50	50

haps a surprising result, that warrants further study, is that it seems that the size of phase 4 is unaffected by the artefacts. That is, the phase to resolve the final small outstanding issues does not seem to benefit from the presence of artefacts. This observation will be explored later in the discussion section.

3.3 Inter-Team Communication

In the previous section, the aim was to identify the characteristics of consensus formation for two teams that are highly integrated in terms of their inter-team communications ability. That is, for a polyarchy, any member of the teams can talk to any other member, both within their own and the other team. A Polyarchical interaction network is feasible in smaller organisations, but becomes infeasible as organization grows and formal structures start shaping who talks to who.

Here we investigate consensus formation for two additional communications configurations; namely team *a* and *b* highly disconnected and two semi-connected teams.

In the case of the highly disconnected teams, see Figure 5, the communications channel between the two teams are though a single link form by one member from each team. This is somewhat of a 'round-cow', approximation, it is often the case that teams need to communicate with other teams though a 'spoc', a single point of contact. The aim of such a person is to remove noise from the rest of the team by having external teams work though this spoc.

The third communications model investigated was that of a small two-person interaction team (the SI team), see Figure 6. This is a model often seen within large organisations where the business analyst and the solutions architect form a team what jointly goes to both business meetings for requirements solicitation and technical meetings for implementation design. In this model the two-team are both connected to all team members of both teams.

The primary measure for these scenarios is time to reach consensus under different team sizes to profile the consensus process and to better understand the ratios involved. That is, how do the time to reach consensus differ for the three organizational communications models between team *a* and *b* (polyarchy, SI-team as mediators, and spoc as mediators) and what is the effect of using artefacts to capture topics and consensus discussions.

For each of the three communications models, the team size is varied, with the fixed ratio of 1:2 for the team sizes of team *a* and *b*. The size of team *a* is ranged from 1 to 20, which leads to an overall project team size ranging from 3 to 60 and for each such team size the simulation is executed 200 times to reach statistical stability.

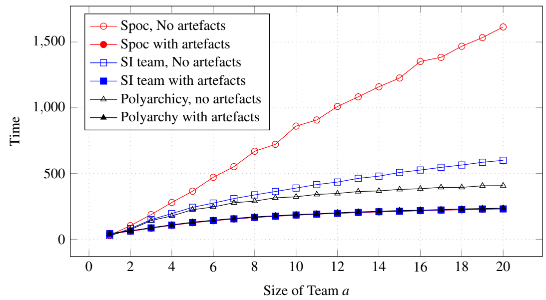

Figure 10: Time to reach consensus under different team configurations. The graphs for 'Spoc with artefacts', 'SI team with artefacts', and 'Polyarchy with artefacts' are so similar they appear on top of each other at this scale and are re-drawn in Figure 11.

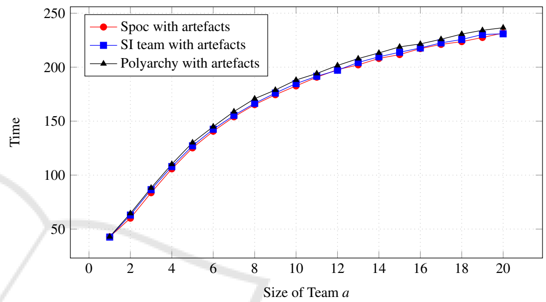

Figure 11: Time to reach consensus under different team configurations with artefacts.

The results are shown in Figure 10. The data depicted in this figure lead to a number of striking conclusions, some obvious, others only after reflection.

A communications strategy that limits teams talking to each other, such as the spoc strategy, leads to very slow consensus formation. This is expected since the communications channel through which information flows between the two teams are highly limited. The spoc-spoc meetings discuss on average $\bar{z}$ topics per meeting (five in our case), see section 2.6.

The SI team is effective in terms of time to reach consensus given how small the team is, compared to the highly effective polyarchy strategy. From this data it is clear why larger organizations opt for this model; it leads to reasonable times, while still keeping the team structure intact. The success here is due to the fact that the SI team can talk to anyone in both teams, breaking the bottleneck observed in the spoc model, and thus allowing information to flow much faster between the teams. It should be obvious that by making the SI team bigger the situation tends more towards a polyarchy and thus the time to reach consensus will approach that of a polyarchy. However, from a practical standpoint it seems that a small SI team derives enough benefit to be a good strategy.

Turning to the effect of artefacts, the results for the three scenarios are so close, see also Figure 11, that we double checked the simulation configuration. The results are surprising to us. The effect of using artefacts is that it eliminates delays in time to reach consensus caused by organizational structure. The agents efficiently use documentation to communicate across team boundaries voiding any delays that could be caused by inter-team structure. This result will be further addressed in the discussion section.

4 DISCUSSION

The results obtained in the previous sections in conjunction with the earlier discussion on meeting efficiency, paints an interesting picture of project delivery which is applicable to larger organizations and projects.

Some caution is appropriate when interpreting the efficiencies obtained by using documentation as shown above. These agents are extremely diligent in following the RTFM (read the documentation) instructions. More so than what we think humans are capable of. That is, the improvements caused by artefact usage are (in our view) an optimistic view of the situation and in real-world situations people will not first read the documentation and then have a meeting. That goes against human nature and experience. Thus, even though these results demonstrate the value of documentation, further modelling is needed to extend the agent's behaviour to reflect the diversity of human behaviour (some people like to read, some to write, some to talk). We plan to publish in a follow-up paper these effects, for example, the effect that only one documenter has on the efficiency of the team, and the time to reach consensus.

Meeting productivity can be used to break the consensus formation process into four phases, see Table 1 and Figures 8 and 9. The data suggests that the most significant contribution to improvements in time to reach consensus is in phase 1. This is in line with research showing that the most significant reason for project delays and failures are due to early misunderstanding, miscommunication and lack of reaching consensus (Al-Ahmad et al., 2009; Whitney and Daniels, 2013; Kian et al., 2016; Waheeb and Andersen, 2022).

The results obtained here supports the notion that thorough and early generation of artefacts significantly improves overall project delivery times though the generation of consensus. That is, artefacts are a highly effective way to reach consensus as well as improve the overall success of the project, ... *if* they are read.

5 CONCLUSION AND FUTURE WORK

Earlier research on the causes of project failures, delays, and cost overruns have identified lack of consensus as one of the key contributing factors. The consensus formation process is time-consuming, and often left out of project planning or its effort is underestimated.

This paper investigated the formation of consensus when the project consists of a two-team approach, where the first team generates the requirements and the second team is responsible for the implementation. Here we looked at the formation of consensus on various topics, that is, the process to reach consensus on the requirements and implementation plan.

We showed that in such a team configuration, the more team-members are allowed to communicate inter-team, the faster consensus is reached. This is not always feasible, especially in larger organizations. A small multi-skilled team that form a group and talk to both teams are very efficient at creating consensus.

The introduction of artefacts (documentation) greatly improves the time to reach consensus and eliminates inefficiencies in the inter-team communications structure. Artefacts significantly improve time to reach consensus irrespective the inter-team communications model.

This research suggests further work is needed to understand the efficiency phases that were identified, and in particular, ways to eliminate or improve situations where one-topic meetings dominate. This research suggests that different meeting cultures should be considered earlier rather than later in project life cycle, to avoid one-topic meetings and thus improve delivery time and consensus formation.

Further understanding of the effect of artefacts in consensus formation can benefit from studies into the balance between artefact generation and meetings based on individual preference. For example, what is the benefit of having 10% of team members focus on artefact generation, versus a lower (or higher) number. What is clear is that someone that likes to document topics are worth their weight in gold (as the saying goes).

Lastly, it can be argtued that artefacts have more authority and thus that people would be lesss likely to modify them. This should be modelled, perhaps by setting up meetings with the document author, indroducing a bottleneck that would reduce the effectiveness of artefacts, and should be explored further using models.

REFERENCES

Al-Ahmad, W., Al-Fagih, K., Khanfar, K., Alsamara, K., Abuleil, S., and Abu-Salem, H. (2009). A taxonomy of an it project failure: root causes. *International Management Review*, 5(1):93.

Amblard, F., Bouadjio-Boulic, A., Gutiérrez, C. S., and Gaudou, B. (2015). Which models are used in social simulation to generate social networks? a review of 17 years of publications in jasss. pages 4021–4032. IEEE.

Amblard, F. and Quattrociocchi, W. (2013). Social networks and spatial distribution. *Simulating Social Complexity: A Handbook*, pages 401–430.

Amirkhani, A. and Barshooi, A. H. (2022). Consensus in multi-agent systems: a review. *Artificial Intelligence Review*, 55(5):3897–3935.

Badke-Schaub, P., Goldschmidt, G., and Meijer, M. (2010). How does cognitive conflict in design teams support the development of creative ideas? *Creativity and Innovation Management*, 19(2):119–133.

Barabási, A.-L., Albert, R., and Jeong, H. (2000). Scale-free characteristics of random networks: the topology of the world-wide web. *Physica A: statistical mechanics and its applications*, 281(1-4):69–77.

Barnes, M. and Prior, D. (2009). *Subversive citizens: Power, agency and resistance in public services*. Policy Press, Bristol, UK.

Bradley, B. H., Anderson, H. J., Baur, J. E., and Klotz, A. C. (2015). When conflict helps: Integrating evidence for beneficial conflict in groups and teams under three perspectives. *Group Dynamics: Theory, Research, and Practice*, 19(4):243.

Carter, D. R., DeChurch, L. A., Braun, M. T., and Contractor, N. S. (2015). Social network approaches to leadership: An integrative conceptual review. *Journal of Applied Psychology*, 100(3):597–622.

Chang, M.-H. and Harrington, J. E. (2000). Centralization vs. decentralization in a multi-unit organization: A computational model of a retail chain as a multi-agent adaptive system. *Management Science*, 46(11):1427–1440.

Chang, M.-h. and Harrington, J. E. (2004). Agent-based models of organizations working paper, no. 515.

Cheung, S. Y., Gong, Y., Wang, M., Zhou, L., and Shi, J. (2016). When and how does functional diversity influence team innovation? the mediating role of knowledge sharing and the moderation role of affect-based trust in a team. *Human relations*, 69(7):1507–1531.

Christakis, N. A. and Fowler, J. H. (2013). Social contagion theory: examining dynamic social networks and human behavior. *Statistics in medicine*, 32(4):556–577.

Curşeu, P. L. (2006). Emergent states in virtual teams: A complex adaptive systems perspective:. *Journal of Information Technology*.

Dunbar, R. I. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, 6(5):178–190.

Eden, C. and Ackermann, F. (2010). Decision making in groups: Theory and practice. *Handbook of decision making*, pages 231–272.

Enyinda, C. I., Blankson, C., Cao, G., and Enyinda, I. E. (2022). Why cannot we all just get along? resolving customer-focused team interface conflicts in a b2b firm leveraging ahp-based multi-criteria decision-making. *Journal of Business & Industrial Marketing*.

Erdős, P., Rényi, A., et al. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60.

Evans, M. S. and Romerstein, H. (2012). *Stalin's secret agents: the subversion of Roosevelt's government*. Simon and Schuster, New York, USA.

Gilbert, G. N. (2010). *Computational social science*, volume 21. Sage, Computational Social Science gilbert.

Hirvonen, P. (2019). Positioning, conflict, and dialogue in management teams. *Qualitative Research in Organizations and Management: An International Journal*, 14(4):444–464.

Holmes, J. and Marra, M. (2004). Leadership and managing conflict in meetings. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, 14(4):439–462.

Iacopini, I., Petri, G., Baronchelli, A., and Barrat, A. (2022). Group interactions modulate critical mass dynamics in social convention. *Communications Physics*, 5(1):64.

Johnson, N. R. and Feinberg, W. E. (1977). A computer simulation of the emergence of consensus in crowds. *American Sociological Review*, pages 505–521.

Jones, S. L. and Shah, P. P. (2016). Diagnosing the locus of trust: A temporal perspective for trustor, trustee and dyadic influences on perceived trustworthiness. *Journal of Applied Psychology*, 101:392–414.

Keizer, K., Lindenberg, S., and Steg, L. (2013). The importance of demonstratively restoring order. *PloS one*, 8(6):e65137.

Keupp, M. M., Palmié, M., and Gassmann, O. (2012). The strategic management of innovation: A systematic review and paths for future research. *International journal of management reviews*, 14(4):367–390.

Kian, M. E., Sun, M., and Bosché, F. (2016). A consistency-checking consensus-building method to assess complexity of energy megaprojects. *Procedia-social and behavioral sciences*, 226:43–50.

Kirschner, P. A., Beers, P. J., Boshuizen, H. P., and Gijselaers, W. H. (2008). Coercing shared knowledge in collaborative learning environments. *Computers in human behavior*, 24(2):403–420.

Leishman, T. G., Green, D. G., and Driver, S. (2008). Self-organization in simulated social networks. *null*.

Leishman, T. G., Green, D. G., and Driver, S. (2009). Self-organization in simulated social networks. In

Computer-Mediated Social Networking: First International Conference, ICCMSN 2008, Dunedin, New Zealand, June 11-13, 2008, Revised Selected Papers, pages 150–156. Springer.

Leskovec, J. and Mcauley, J. (2012). Learning to discover social circles in ego networks. *Advances in neural information processing systems*, 25.

Manky, O. and Dolores, J. (2022). Subversive entrepreneurs: Business agency and commodification of peruvian higher education (1992–2012). *Latin American Perspectives*, 49(3):162–180.

McDowell, K. (2002). Roll of thunder, hear my cry: A culturally specific, subversive concept of child agency. *Children's Literature in Education*, 33:213–225.

Michalski, R., Serwata, D., Nurek, M., Szymanski, B. K., Kazienko, P., and Jia, T. (2022). Temporal network epistemology: On reaching consensus in a real-world setting. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 32(6).

Paletz, S. B., Chan, J., and Schunn, C. D. (2017). The dynamics of micro-conflicts and uncertainty in successful and unsuccessful design teams. *Design Studies*, 50:39–69.

Reagans, R., Miron-Spektor, E., and Argote, L. (2016). Knowledge utilization, coordination, and team performance. *Organization Science*, 27(5):1108–1124.

Roselló, L., Prats, F., Agell, N., and Sánchez, M. (2010). Measuring consensus in group decisions by means of qualitative reasoning. *International Journal of Approximate Reasoning*, 51(4):441–452.

Sáenz-Royo, C. and Lozano-Rojo, A. (2023). Authoritarianism versus participation in innovation decisions. *Technovation*, 124:102741.

Sah, R. K. and Stiglitz, J. E. (1984). The architecture of economic systems: Hierarchies and polyarchies. Technical report, National Bureau of Economic Research.

Sah, R. K. and Stiglitz, J. E. (1988). Committees, hierarchies and polyarchies. *The Economic Journal*, 98(391):451–470.

Stocker, R., Green, D. G., and Newth, D. (2001). Consensus and cohesion in simulated social networks. *Journal of Artificial Societies and Social Simulation*, 4(4).

Tesfatsion, L. and Judd, K. L. (2006). *Handbook of computational economics: agent-based computational economics*. Elsevier, Amsterdam, The Netherlands.

Tsvetkova, M. and Macy, M. W. (2014). The social contagion of generosity. *PloS one*, 9(2):e87275.

Vorster, J. and Leenen, L. (2023a). Consensus simulator for organisational structures. In *Proceedings of the 13th International Conference on Simulation and Modeling Methodologies, Technologies and Applications*, pages 15–26.

Vorster, J. and Leenen, L. (2023b). Exploring the effects of subversive agents on consensus-seeking processes using a multi-agent simulator. In *Proceedings of the 13th International Conference on Simulation and Modeling Methodologies, Technologies and Applications*, pages 104–114.

Vorster, J. and Leenen, L. (2024). Covert subversive agents and consensus disruption on large projects. In *Proceedings of the 19th International Conference on Cuber Warfare and Security, Reeding, UK*, pages 421–429.

Waheeb, R. A. and Andersen, B. S. (2022). Causes of problems in post-disaster emergency re-construction projects—iraq as a case study. *Public Works Management & Policy*, 27(1):61–97.

Watts, D. J. and Strogatz, S. H. (1998a). Collective dynamics of small-world networks. *Nature*.

Watts, D. J. and Strogatz, S. H. (1998b). Collective dynamics of 'small-world'networks. *Nature*, 393(6684):440–442.

Wei, Q., Wang, X., Zhong, X., and Wu, N. (2021). Consensus control of leader-following multi-agent systems in directed topology with heterogeneous disturbances. *IEEE/CAA Journal of Automatica Sinica*, 8(2):423–431.

Whitney, K. M. and Daniels, C. B. (2013). The root cause of failure in complex it projects: Complexity itself. *Procedia Computer Science*, 20:325–330.

Will, M. G., Al-Kfairy, M., and Mellor, R. B. (2019). How organizational structure transforms risky innovations into performance–a computer simulation. *Simulation Modelling Practice and Theory*, 94:264–285.

Wu, G., Liu, C., Zhao, X., and Zuo, J. (2017). Investigating the relationship between communication-conflict interaction and project success among construction project teams. *International Journal of Project Management*, 35(8):1466–1482.

Xie, J., Sreenivasan, S., Korniss, G., Zhang, W., Lim, C., and Szymanski, B. K. (2011). Social consensus through the influence of committed minorities. *Physical Review E*, 84(1):011130.

Yan, H.-B., Ma, T., and Huynh, V.-N. (2017). On qualitative multi-attribute group decision making and its consensus measure: A probability based perspective. *Omega*, 70:94–117.