A Quasi-Experiment to Expose Attention-Decision-Learning Cycles in Engineering Project Teams

Lorena Pelegrin Alvarez
MIT Global Teamwork Lab
pelegrin@mit.edu

Bryan R. Moser, PhD
MIT Global Teamwork Lab
bry@mit.edu

Vivek Sakhrani, PhD
MIT Global Teamwork Lab
sakhrani@mit.edu

Abstract

Engineering project outcomes are driven by a dynamic mix of the social physics of teams, the unique complexities of the engineering challenge at hand, and stakeholder pressures in context. Various related research has demonstrated formal experiments for tightly controlled problems in small teams, including work in organizational psychology, computational organization theory, design thinking, and coordination science. We realize there is room for testing these foundational concepts in quasi-controlled environments with distributed teams challenged by problem, solution, and organization complexity common today. This paper presents a quasi-experiment to study how engineers proceed through attention, decision, and learning cycles in the design of a System of Systems. The experiment utilized an ensemble of an agent-based model, a decision-support interface, and a variety of sensors to record behavior and activity. Four pilots for a maritime industry challenge were conducted with experienced industry experts, followed by a primary experiment for data collection. Though this work is preliminary, the experimental approach detects (for this case) how designers focused on different variables (attention), manipulated variables to accomplish desired outcomes (decisions), and explored the system performance trade space variously over time to reveal false assumptions and uncover better decisions (learning). Lessons learned from this quasi-experiment are guiding this research team to prepare scalable and reproducible engineering teamwork experiments that include sensors of events over time in the problem, solution, and socials spaces of engineering projects.

1. Background

This research is rooted in observations of engineering as a social activity across a team of teams as they explore fundamental and often counterintuitive tradeoffs. [1] In a stable environment, with teams, markets, and technologies well understood, engineering can be characterized as a drive to efficiency through an analytic process, improved with decreasing uncertainty over time. Tacit capabilities and mental models for successful teams remain aligned with internal and external realities. However, for modern, disruptive, and strategic industrial needs, engineering is much the opposite, proceeding by innovation under uncertainty.

Innovation is a collective capability, involving individual behaviors and group dynamics. Design is a crucial component of innovation. One example of how social behavior affects design decisions is the manner in which designers engage with each other to frame and re-frame the design problem itself, which subsequently influences solutions [2], [3].

A team's mental models can limit their capacity for awareness during complex work. Organization processes, including those for engineering teams, have been proposed to assess, refresh, expand and make explicit the mental models of coordinating teams. [4]–[9]

While a century ago organizations were conveyed as structured, centrally controlled entities; that view has given way in recent decades to a more natural representation. Kozlowski and others have articulated teamwork in an organization as distributed and dynamic, and thus our models and research on organizations must be multi-level. As such, a mix over time of elements, relationships, topology, externalities, and dynamics leads to organization as system with performance as an emergent outcome. [10], [11]

Supported by advances in artificial intelligence, operations research, and computational organization theory, models of teams and organizations are simulated, allowing virtual experiments to validate, question, and expand existing ethnographic and management frameworks. [12]–[15] While there have being increasing variants and interesting explorations, relatively few of these models for engineering teamwork have been validated with reproducibility at real world scale.

URI: https://hdl.handle.net/10125/59475
ISBN: 978-0-9981331-2-6

Recently, with the advent of low-cost pervasive sensors and digital twin models, a research opportunity has risen to supplement existing frameworks and formal experiments with a broader empirical basis. [16]–[18] The motivation of this research is to build quasi-experiments for real-world engineering project teams supported by digital models and sensors as instrumented teamwork.

1.1. Research Framing

We characterize design challenges as being composed of the problem space, solution space, and social space. A broad motive is to systematically observe how the social space influences or perhaps even governs how teams navigate the links between problems and solutions. In other words, we wish to study how the social physics of engineering teams influences design process and outcomes [19].

These perspectives build on the information processing view of innovation and learning in project organizations [13], [20]. With change, old information loses relevance, uncertainties arise, and new information is generated. The position and value of information across the network of the organization evolves. Information new to actors in a given situation may be a surprise [21]. The relevance of new information is not proportional to volume, but a value given topological significance of the knowledge. [22] Small facts may yield big insights with systemic effect. How the organization frames the problem and potential solutions may encourage exploration and recognition of these surprises. [3]

An alternative framing for design behavior is Berglund & Leifer's [23] Triple-Loop model (process, product, and context) variables, building on Argyris's (1977) Double Loop model of learning [24].

A social space is viewed through the position and interaction of teams of both individuals and other teams (Team of Teams – ToT). Some researchers have sought to measure the quality of the social space through constructs such as collective intelligence (or C factor) [25] or the Interaction Dynamics Notation (IDN) [26]. These efforts are among the first to develop a cognitive-behavioral model of engineering design team performance. Of particular interest in the literature on innovation is the role of influence cycles [18] in team interactions and the impact on what teams focus on (attention), how they arrive at choices (decisions), and how they improve on past choices (learning) [27].

In this ongoing experimental study, we focus on the idea that the design process is path-dependent. We seek to study how the latent social behaviors associated with path-dependent exploration influence design choices and outcomes in complex system engineering [1], [28].

Path-dependence in a design trade space is explained as follows. Engineering teams begin at a legacy position in the tradespace, determined by the prevailing solution to the problem, tacit knowledge, and influenced by externally determined specifications. Teams then move through the space, attempting to successively improve on the previous positions, i.e. do better than where they have been. They eventually converge to final choices either by arriving at a pareto location, or satisfying requirements under resource constraints. The phenomenon of the design walk – the path moving through the design space -- can be observed and is akin to project shaping [29]. However, many underlying latent behaviors such as attention, decision, and learning that govern the exploration process have until recently been difficult to observe and influence.

A **Platform for Quasi Experiments**: We realize there is room for testing these foundational concepts in quasi-controlled environments with distributed teams challenged by problem, solution, and organization complexity common today. We refer to this level of complexity as the meso-scale, in contrast to micro-scale experiments with individuals and small teams and macro-scale experiments relying on population scale data. This paper presents a quasi-experiment to study how engineers proceed through attention, decision, and learning cycles in the design of a System of Systems.

We pose the following research questions to link the unobservable / latent behaviors in the social space to the observable events in the problem and solution spaces:

a) Attention allocation – how do individual and team behaviors influence the particular design variants that teams focus on as they proceed? What are the social signals and factors to which the designers are attuned in relation to this focus set of variants?

b) Decision – how do teams evaluate and process their design moves by either progressing or regressing through the space? What interactions result in agreement, or a choice?

c) Learning – how do teams recognize, process and engage over the trade-offs that result from design decisions, and how do they alter them to improve upon previous choices?

Our methodology and approach are accordingly structured to observe how individuals and teams behave at the individual and collective levels [25], [30], [31], so that patterns of attention, decision, and learning that influence path-dependent design tradespace exploration are revealed.

2. Methodology & Approach

2.1. Experimental Platform

Improvements in computation and sensing have now made it possible to study the activities of engineering using a platform that combines models and experiments. Such a platform enables (i) the deployment of complex computation to the edge, i.e. where individuals and teams can use distributed devices to handle complex computation in near real-time, (ii) enable visualization of complex design tradespaces to support trade-off evaluation and decision-making, and (ii) instrument individuals and their environment to observe individual and team interaction and behavior.

These technologies promise to minimize the cognitive burden on individuals trying to process complex information, which is a major concern in the engineering of complex systems [32]. An objective in the deployment of these technologies is to free up cognitive and emotional capacity for individuals to engage in meaningful exchange of insights as they explore the tradespace. Recent studies have demonstrated that this experimental approach requires model development as both boundary object for engagement and as support for instrumentation and observation [33], [34].

2.2 Design Challenge Setting

We formulated an engineering challenge for teams of teams (ToT) as part of a commercial maritime cluster of companies and a national laboratory. The cluster consists of stakeholder representatives (**Figure 1**) from the Japanese shipping industry -- cargo suppliers and buyers, ship owners and operators, infrastructure assets owners including ports and bunkering facilities, regulatory system principals, and the ship building sector (designers and builders).

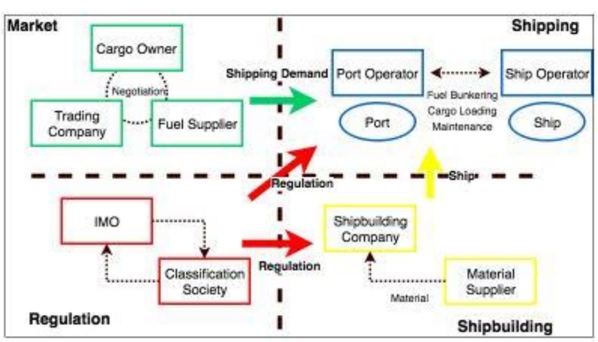

Figure 1. Maritime shipping stakeholder landscape

The challenge for these stakeholders (the team of teams) is to re-design the integrated marine system to comply with revised emissions reduction regulations enacted by the IMO MARPOL (International Convention for the Prevention of Pollution from Ships).

The decisions necessary to invest and change the industry are dependent across actors and therefore require coordination. This revision of the regulation mainly sets limits for Sulphur Oxides (SOx) and Nitrous Oxides (NOx) emissions from ships' exhausts, and will go in effect in 2020. Thus the new regulations create performance targets; participants must study how design variants trade-off other performance dimensions to meet the newly specified performance targets.

Teams are asked to modify a reference crude oil shipping system involving a tankers' fleet composed of Very Large Crude Carriers (VLCCs), currently fueled with Heavy Fuel Oil (HFO) and transporting crude oil from a supply port in the Persian Gulf to a delivery port in Japan.

The challenge addresses the expected progressive transition from the currently predominant use of HFO to Liquefied Natural Gas (LNG), and considers the LNG infrastructure needed to support this transition. The design goal is to reduce SOx and NOx emissions, while fulfilling shipping contracts, at the lowest possible cost (**Table 1**). Individuals representing various stakeholders consider, enumerate, and evaluate feasible system architectures in a tradespace simulated in a computer model. Designers seek the Pareto frontier of non-dominated architectures, and choose a subset of preferred architectures.

Table 1. System performance metrics in the design exercise

Objective	Description	Metrics	Units
Emissions Reduction	Ability of the system to reduce emissions, in respect to baseline emissions of HFO combustion.	NOx Emissions	Ton/ Ton Cargo * km
		SOx Emissions	Ton/ Ton Cargo * km
		CO2 Emissions	Ton/ Ton Cargo * km
Schedule	Ability of the system to meet the contract schedule for crude oil shipping, with minimal disruption.	Waiting Time	%
		Cargo Moved	Cargo Ton
CAPEX	The cost of installation of new LNG bunkering facilities, and retrofitting crude oil tankers.	Initial Cost	MUSD
OPEX	Cost of operating the retrofitted crude oil tankers fleet, incl. voyage.	Fuel Cost Efficiency	USD/ Cargo Ton * km

2.3 Quasi-Experiment Setup

A quasi-experiment based on the maritime transition design challenge was established (**Table 2**). In the exercise, designers can modify designs to achieve one or more variants by playing with a limited set of architectural decisions. These decisions and options for each decision are represented in a morphological matrix. For the crude oil shipping system these include propellant fuel for the fleet of ships, engine and overall propulsion system architecture, fuel tank and vessel layout, fuel bunkering and refueling implications, berth flexibility options and levels (**Table 3**).

Table 2. Experiment variables

Dependent: Emissions (NOx, SOx, CO_2), Waiting Time, Cargo Moved, Initial Cost, Fuel Cost Efficiency
Independent: Number of architectures enumerated, Attention allocation to the problem space, Attention allocation to the solution space, Number of path dependent sequences, Number of surprises, Timestamps
Controlled: Time limit, System of systems model, Demographics (No randomization, quasi-experiment)

Table 3. Matrix of design variables and levels

Decision	Alternatives			
# Ships HFO	0	5	15	20
# Ships LSFO				
# Ships LNG				
# Ships HFO/LNG				
LNG Bunkering Location	Persian Gulf	Singa-pore.	Japan	-
LNG Bunkering Method	Truck to Ship	Ship to Ship	Shore to Ship	-
# LNG Bunkering Facilities by Location	0	1	3	-

The design challenge variables were selected based on the research questions to be explored, the typical tasks of design teams, and the associated teamwork phenomena mapped in **Table 4**.

The quasi-experiment was developed through a series of four pilot experiments with experienced industry professionals at sites in the USA and Japan. The pilot experimentation phase also served for prototyping the computer simulator that implements the system of systems (SoS) model and the interactive visualization software user interface.

Table 4. Experiment variables mapped to teamwork tasks and phenomena

Teamwork Task	Teamwork Phenomena	Experiment Variables
Collective interpretation of design goals and definition of strategy	Attention Allocation	Attention allocation on elements of the problem space
	Decision	System performance metrics priorities
Collective enumeration of architectures	Attention Allocation	Attention allocation on elements of the solution space
	Decision	Number of architectures enumerated
Collective consideration and evaluation of architectures	Attention Allocation	Attention allocation on Elements of the problem space
	Learning	Number of path dependent sequences
		Number of surprises
Collective selection of best architecture	Attention Allocation	Attention allocation on elements of the problem space
	Decision	System performance metrics from selected architecture

2.3 Tradespace Simulation

MOSES is an agent-based simulator developed by Wanaka for the evaluation of architectural decisions in ship transportation systems. [35] The simulator was developed based on specific technical and physical realities of the ship, shipping, and port for the case in the experiment. The simulation was then improved through the pilot experimentation phase specifically for this exercise.

There are five types of agents and five types of demands (Table 5). Agents have four functions: (i) observe, (ii) select, (iii) checkNextEvent, and (iv) update. For each iteration of the simulation, agents observe their status and task list according to the demand, select their next task, and estimate the next event.

Table 5. Types of agents and demands in MOSES

Agent	Demand	To
Cargo owner	Shipping demand	Ship operator
Ship operator	Transportation demand Berthing demand	Ship Terminal operator
Terminal operator	Bunkering demand Loading demand	Port facility Port facility
Ship	N/A	N/A
Port facility	N/A	N/A

Further to the computer simulator MOSES introduced here, the experiment developed is instrumented and aided by several other software packages, the most significant of them being the so-called Maritime Decision Support System (DSS).

The Maritime DSS software is a package of open source software developed by Winder. It enables teams engaged in the design problem to generate and evaluate architectures with enhanced visualization features. It provides a UI that augments a team's understanding of the problem and the solution spaces, as well as the underlying system model of the simulator and its assumptions. During the experiment, socio-metric data is collected passively while the teams engage in solving the design problem.

Figure 2 provides a snapshot of the UI. The left-hand toolbar allows users to configure and simulate a unique shipping fleet. The right-hand toolbar allows users to explore seven dimensions of KPIs (the system performance objectives or -ilities). The KPI plot (lower right) allows users to track and revisit the performance of multiple scenarios over time in a tradespace format.

The Maritime DSS has a module called "Team Space IO" for analysis of "fingerprints" data collected (**Figure 3**). By cross-referencing fingerprint data with other time-stamped observations about intention, behavior, and/ or strategy, we can verify or preclude research hypotheses.

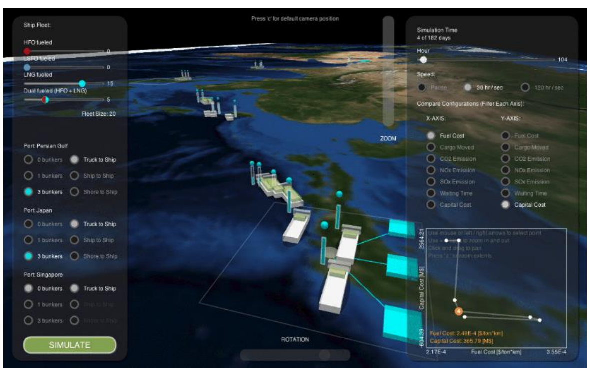

Figure 2. Maritime DSS UI

We focus on two broad categories of data:

● Attention: We measure attention by knowing what subset of information (i.e. KPIs) users are viewing over time. We can also view which inputs are changed.

● Performance: Performance is evaluated relative to other teams and relative to a simulated tradespace.

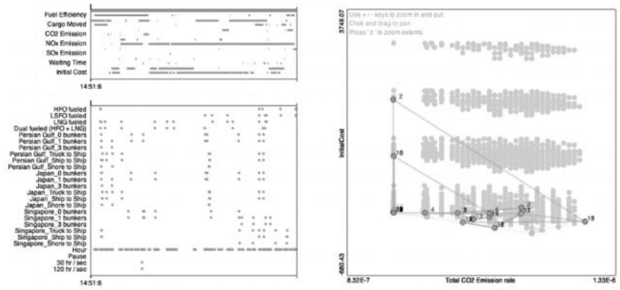

Figure 3. Sensor output interface for Maritime DSS system

3. Experiment Results

3.1. Attention Allocation by Teams

Attention allocation on the variables of the problem space can be sensed by recording how much time a tradespace variable was set or remained set to a certain system performance objective.

The major objectives subject of attention in this experiment have been Fuel Cost, Cargo Moved and Initial Cost. The following **Figure 4** provides the distribution of attention into the different trade variables in the very moment that a team performed a simulation or a recall.

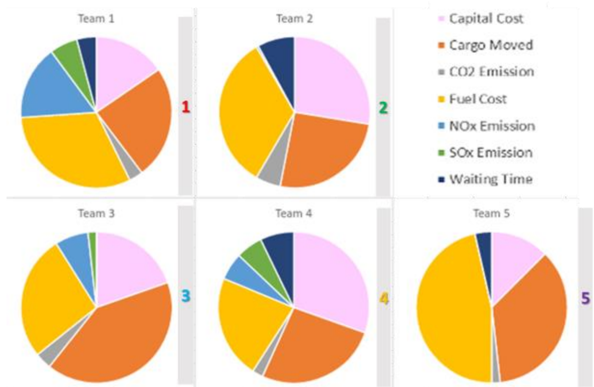

Figure 4. Tradespace variables on which teams focused

The results documented in **Table 6** show that there is a variance in how teams allocated attention to variables.

Table 6. Attention Allocation on Variables of the Problem Space for different Teams

	Team 1	Team 2	Team 3	Team 4	Team 5
Attention Allocation Problem Space	Fuel Cost Eff.: 31.23% Cargo Moved: 24.25% NOx Emissions: 15.95% Initial Cost: 15.45%	Fuel Cost Eff.: 33.14% Initial Cost: 27.56% Cargo Moved: 25.51%	Cargo Moved: 41.07% Fuel Cost Eff.: 26.79% Initial Cost: 19.64%	Initial Cost: 30.43% Cargo Moved: 26.36% Fuel Cost Eff.: 22.15%	Fuel Cost Eff.: 46.43% Cargo Moved: 35.71% Initial Cost: 12.50%

Figure 5 illustrates the distribution of attention into the possible trade-offs between objectives by different teams. Some teams evaluated architectures from more viewpoints (checked more types of trade-offs) than the other teams.

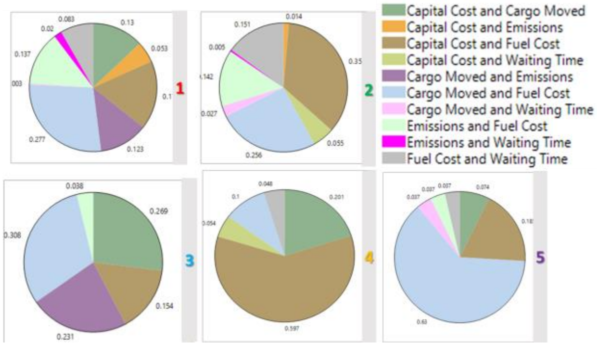

Figure 5: Joint variables on which teams focused

Figure 6 provides an aggregated view of the density of categories of changes through time by different teams.

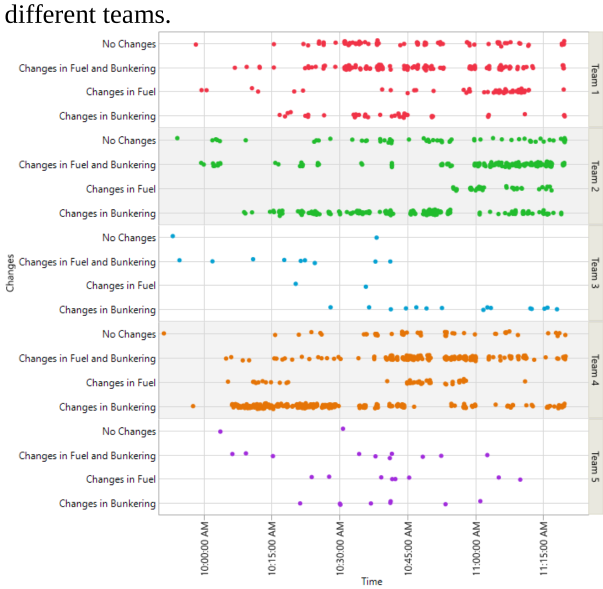

Figure 6. Architectural Changes by Teams over Time

A few teams seem to have concentrated more on changing bunkering infrastructure options rather than on ship portfolio options, while the other teams seem to have had a more balanced distribution of focus on design inputs. This can also be observed in the following **Table 7**, which shows the dominant areas of attention.

Table 7. Attention Allocation on Variables of the Solution Space for different Teams

	Team 1	Team 2	Team 3	Team 4	Team 5
Attention Allocation Solution Space	Fuel Type: 52.04% LNG Bunkering: 47.96%	Fuel Type: 36.06% LNG Bunkering: 63.94%	Fuel Type: 31.43% LNG Bunkering: 68.57%	Fuel Type: 43.81% LNG Bunkering: 56.18%	Fuel Type: 50% LNG Bunkering: 50%

3.2. Decisions by Teams

The design goal of the exercise was to reduce SOx emissions and NOx emissions, while fulfilling shipping contracts, at the lowest possible cost. **Figure 7** plots the time series of the outcomes for a sample of different teams. Team 1, 2 and 4 have a very large number of data points, while Team 3 and 5 have much less.

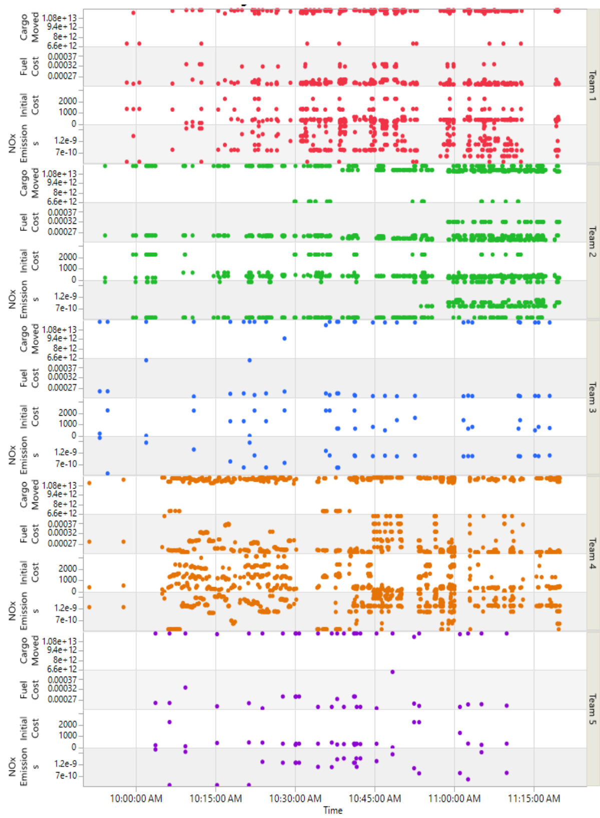

Figure 7. Performance variable changes over time by each of 5 teams; team 1 at top (red), team 5 at bottom (purple)

Teams interpreted the design goals differently. The Post-survey collected from each team identifies the design strategy followed for this experiment:

- Team 1 (red): "Maximize Cargo Moved and minimize Fuel Cost, while keeping Emissions as low as possible".
- Team 2 (green): "Decrease Fuel Cost, with a compromise on Emissions and Cargo Moved".
- Team 3 (blue): "Minimize Emissions and long-term operation cost".
- Team 4 (yellow): "Maximize revenue first, and achieve lower SOX/NOX as a secondary goal. To this end, we identified that the main variables contributing to revenue were Cargo moved, Fuel cost, and capital cost. Our model is (roughly) Revenue = CM - FC – CC".

● Team 5 (purple): “Maximize Cargo Moved, and minimize Fuel Cost and Capital Cost, while keeping Emissions reduction reasonable”.

In terms of design strategies, the teams approached the design walk differently:

● Team 1 considered/ enumerated new architectures, evaluated them and then compared them with (recalled) previous architectures in cycles, converging to their selected architecture.

● Team 2’s design walk shows that they simulated many times and did not use the recall function. It is not clear how many of the simulations are in effect actually a recall without further analysis.

● Team 3’s design walk shows they simulated fewer architectures and recalled a few times.

● Team 4 attempted to generate the simulated tradespace, analyzed it (through recalls), defined design goals, and selected an architecture.

● Team 5’ design walk shows they simulated fewer architectures and recalled only a handful of times.

3.2. Learning Cycles by Teams

A method used in this quasi-experiment to detect and analyze learning cycles is illustrated with a detailed review of Team 1. The nature of the surprises encountered, the associated outcomes, and subsequent changes are visualized.

A learning cycle is defined as a process by which a team considers, evaluates, and reflects about design choices. Learning cycles are characterized by the encounter of surprises followed by a set of changes in a sequence as the team explores in response to the new information in the surprise.

Identification of learning cycles of teams is performed by detecting when and why design teams encounter surprises while carrying out a design task, and reviewing the design changes they execute. The changes are driven by design goals. Mental models of the team evolve with the new information obtained and the associated reflection and reframing that happens after encountering a surprise. When a team fixes design variables after having encountered a surprise (i.e. after having learned something new about system dynamics) then continues exploring changes to other design variables, we speak about a path-dependent sequence of design solutions. The new design solutions tested may be dependent on insights derived from previous design solutions. This procedure has been applied to four different teams in the frame of two design workshops.

Table 9. Definitions for identifying learning cycles

Variable	Description
Surprise	Indicates the point in time in their design walk when a team encountered a surprise.
Architecture	The system configuration, a set of design decisions, that was considered and evaluated when the team encountered the surprise.
Recorded Reasons for Surprise	The reasons the team judged the outcomes different than expected. A surprise is encountered when system performance is better or worse than expected.
Potential Learning and likely decision in course of action	What the teams possibly learned about the specific system dynamics and what they probably decided for the next moves in their design process.
Subsequent Changes until next Surprise	The type of design changes that the team explored until the time they encountered another surprise in their design process.

Figure 8 provides the time series of outcomes and design changes, whereby the architectures enumerated (i.e. the time that an architecture considered was first simulated -marked with asterisks), and recalled (i.e. the times that an architecture previously enumerated was recalled -marked with red dots) are highlighted.

Learning cycles are plotted as black lines connecting the asterisks. Phases of the design walk are identified in the figure as follows: (1) is an exploratory phase, whereby mental models were likely initially confirmed or challenged. In this initial phase we can see fewer recalls, as there are fewer architectures to be recalled, and the first systemic surprises/ learnings emerge. In this phase we also see more testing of both Ship Portfolio and Bunkering options; (2) comprises an analysis phase, where there is higher density of recalls (i.e. comparison between architectures) and where path-dependent sequences are established. In this phase we also see more testing of bunkering options; (3) appears to be a deliberation phase, including a 4min pause before 11 am; (4) indicates a fine-tuning phase, with some last new architectures and very high density of recalls from the array of LNG and Dual Fuel hybrid fleet options, i.e. testing of Ship Portfolio details with a fixed bunkering decision; ending in (5) a final decision phase characterized by a 7min pause and some recalls.

Table 8. Features of architecture selected by Team 1

Team No.-arch.	# Ships fueled with				Persian Gulf		Japan		Singapore	
	HFO	LSFO	LNG	Dual-fuel	# LNG Bunkers	Bunkering Method	# LNG Bunkers	Bunkering Method	# LNG Bunkers	Bunkering Method
Team 1	-	-	10	10	-	-	-	-	1	Ship to Ship

A first path-dependent sequence of Ship Portfolio (incl. 10xDual-fueled ships) seems to have been triggered by Surprise 5 (the first time the selected architecture was considered). This surprise likely caused a reframing about the positive effects of designing for a half fleet of LNG ships. This is identified in **Table 10**. The elements of surprise are highlighted with yellow circles (i.e. Better performance than anticipated in all variables), and the associated path-dependent sequence with a yellow rectangle.

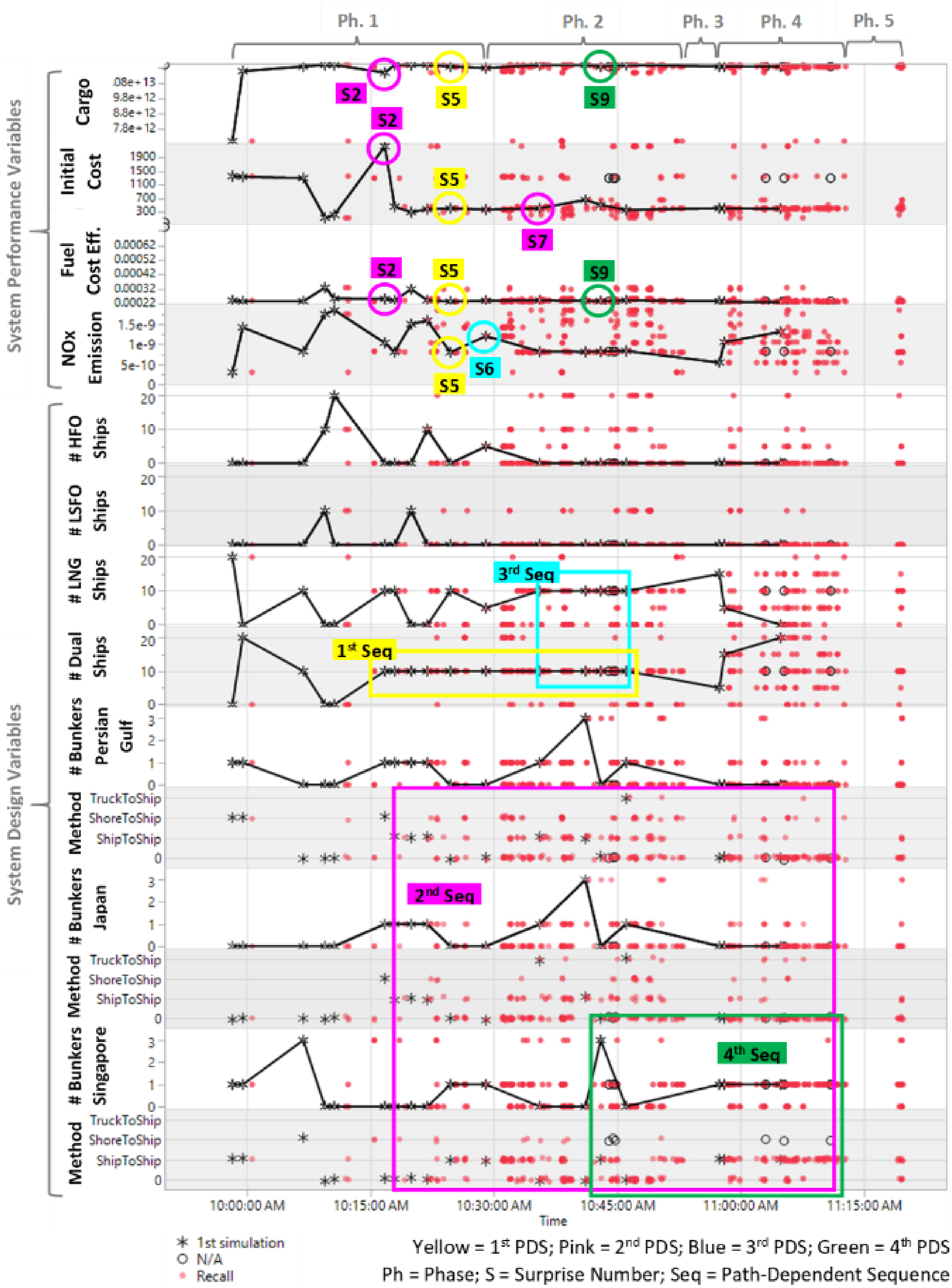

Figure 8. Analysis of design walk showing phases, surprises, and sequences across design & performance variables

The second path-dependent sequence is triggered by Surprise 2 (i.e. Shore-to-Ship too high an investment), after which most of the design walk considers Ship-to-Ship configurations, with some checks on Truck-to-Ship, on an attempt to verify a significant reduction in Initial Cost (Surprise 7). The related outcomes that resulted different than anticipated are marked with pink circles (i.e. Worse Cargo Moved, Fuel Cost, and Initial Cost than anticipated), and the Ship-to-Ship bunkering choices made thereafter can be seen in the sequences marked with a pink rectangle.

A third path-dependent sequence is triggered by Surprise 6 (i.e. Further investigate 10LNG/10Dual configurations). This surprise likely caused a reframing about the effect of including any number of HFO ships in a fleet, thereby confirming the Ship Portfolio configuration of the selected architecture. Surprise 6 on Worse Emissions than expected has been marked with a blue circle and the associated sequence of Ship Portfolio choices with a blue rectangle.

A fourth path-dependent sequence is triggered by Surprise 9, however the insight was first discovered at Surprise 5 (i.e. Bunkering in Singapore could work too.). This confirms the location of the bunkering point in Singapore. Surprise 9 (Worse Cargo Moved, and Fuel Cost than anticipated) is marked with green circles, and the path-dependent sequence with a green rectangle.

Regarding the number of bunkers, it could be argued that the team had a pre-existing mental model that supported one bunker configurations (most of their architectures considered feature one or no bunkers).

Table 10. Key surprises in Team 1's learning cycle

Sur-pri-se	Previous Arch.	Recorded Reasons for Surprise	Potential Learning and likely decision in course of action	Changes until next Surprise
	Current Arch.			
2	20x HFO	Worse Cargo, Fuel Cost, and Initial Cost than anticipated	Shore-to-Ship too high an investment. Explore other bunkering methods.	Changes in Fuel and Bunke-ring
	10x LNG, 10x Dual 1x in PG, Shore-to-Ship 1x Bunker (JP), Shore-to-Ship			
5	10x LSFO, 10x Dual 1x Bunker (PG), Ship-to-Ship 1x Bunker (JP), Ship-to-Ship	Better than anticipated in all variables	Good candidate. Continue exploring hybrid fleet combinations. Continue exploring bunkering configurations incl. Ship-to-Ship.	Changes in Fuel and Bunke-ring
	10x LNG, 10x Dual 1x LNG (SG), Ship-to-Ship			
6	10x HFO, 10x Dual 1x Bunker (PG), Ship-to-Ship 1x Bunker (JP), Ship-to-Ship	Worse Emissions than antici-pated	Team wrote: "Capital cost not changing huge, NOx and SOx went up, SOx same" Do not further consider fleets incl. HFO. Further investigate 10LNG/10Dual config.	Changes in Bunke-ring
	5x HFO, 5x LNG, 10x Dual 1x Bunker (SG), Ship-to-Ship			
7	5x HFO, 5x LNG, 10x Dual. 1x Bunker (SG), Ship-to-Ship	Worse Initial Cost than antici-pated	Team wrote: "Not huge benefit in capital cost by changing to truck to ship, should keep ship to ship" Continue exploring bunkering configurations incl. Ship-to-Ship.	Changes in Bunke-ring
	10x LNG, 10x Dual. 1x Bunker (PG), Ship-to-Ship 1x Bunker (JP), Truck-to-Ship			
9	10x LNG, 10x Dual. 3x Bunker (PG), Ship-to-Ship 3x Bunker (JP), Truck-to-Ship	Worse Cargo Moved, and Fuel Cost than anticipated	1 bunker in Singapore seems a good option. Continue exploring bunkering. Note: Team wrote: "Didn't change cargo moved, fuel costs higher than 1 bunker in Singapore".	Changes in Bunke-ring
	10x LNG, 10x Dual. 3x Bunker (SG), Ship-to-Ship			

3.4. Team Performance

Teams were ranked based on the performance of their selected architectures. For every trade-off, each architecture is compared to the others in the two objectives of the trade-off. The comparison consists of a simple estimate of the distances between the one architecture subject of analysis and the best performing architecture in both of the trade-off objectives.

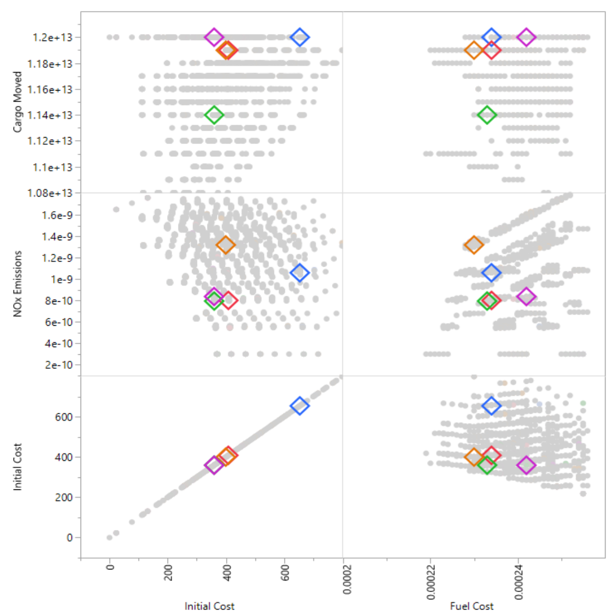

Figure 9. Selected architecture performance plotted on Cost, NOx, Cargo, and Fuel tradespaces.

The selected trade-offs are the most likely that the teams considered in their design walks, according to their statement of design principles (goals), and the attention allocation data collected. Every trade-off is equally weighted, so that a “Global Rank” is calculated simply aggregating the “Trade-off Ranks”.

In cases were the distances are similar, distances to an imaginary point that would be non-dominated have been calculated. That is the case with Team 3 and Team 4’s solutions in Cargo Moved vs. Fuel Cost, and Team 3 and Team 5 in NOx Emissions vs. Fuel Cost. This is shown in **Figure 10**, the intent is to illustrate possible easy ways to calculate relative ranks.

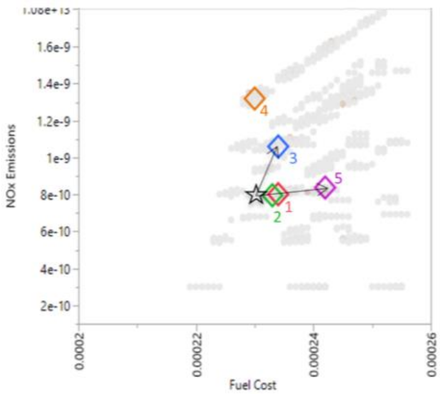

Figure 10. NOx vs. Fuel Cost by Team. Distances to an imaginary non-dominated point in team 3 & 5’s solutions.

Table 11. Performance ranking

Team	Cargo Moved vs. Initial Cost	Cargo Moved vs. Fuel Cost	NOx Emissions vs. Initial Cost	NOx Emissions vs. Fuel Cost	Initial Cost vs. Fuel Cost	Global Rank
	Trade-off Analysis					
	Trade-off Rank					
Team 1	-Dominated by Team 5. -Team 3 better in Cargo Moved. -Team 2 and 4 better in Initial Cost.	-Team 3 and 5 better in Cargo Moved. -Team 2 and 4 better in Fuel Cost.	-Dominated by Team 2. -Team 4 and 5 better in Initial Cost.	-Team 2 and 4 better in Fuel Cost.	-Team 2, 4, and 5 better in Initial Cost. -Team 2 and 4 better in Fuel Cost	2 (Sum = 14)
	3	3	3	2	3	
Team 2	-Dominated by Team 5. -Teams 1, 3, and 4 better in Cargo Moved.	-All other teams better in Cargo Moved. -Team 4 better in Fuel Cost.	Non-dominated	-Team 4 better in Fuel Cost.	-Team 4 better in Fuel Cost.	1 (Sum = 13)
	5	5	1	1	1	
Team 3	-Dominated by Team 5. -Team 1, 2, and 4 better in Initial Cost.	-Team 2 and 4 better in Fuel Cost.	-Dominated by Team 2. -Team 1 and 5 better in NOx. -All other teams better in Initial Cost.	-Team 1, 2 and 5 better in NOx. -Team 2 and 4 better in Fuel Cost.	-All other teams better in Initial Cost. -Team 2 and 4 better in Fuel Cost.	5 (Sum = 20)
	4	2	5	4	5	
Team 4	-Dominated by Team 5. -Team 3 better in Cargo Moved. -Team 2 better in Initial Cost.	-Team 3 and 5 better in Cargo Moved.	-Dominated by Team 2. -All other teams better in NOx. -Team 5 better in Initial Cost.	-All other teams better in NOx.	-Team 2 and 5 better in Initial Cost.	2 (Sum = 14)
	2	1	4	5	2	
Team 5	Non-dominated	-All other teams better in Fuel Cost.	-Dominated by Team 2.	-Team 1 and 2 better in NOx. -All other teams better in Fuel Cost.	-All other teams better in Fuel Cost.	2 (Sum = 14)
	1	4	2	3	4	

According to this performance assessment, Team 1, 2, 4, and 5 performed better than Team 3. However, Team 3 was the only team that stated slightly different design goals giving Emissions Reduction a higher priority than the other teams.

We observe that even on NOx Emissions vs. Initial Cost, Team 1, 2, and 5 perform better on NOx Emissions (and even dominate Team 3’s solution). Another example is with NOx Emissions vs. Fuel Cost, whereby we can see the same on NOx Emissions and both Team 1 and 2 dominate Team 3’s solution. Then, it could be argued that Team 1, 2 and 5’s higher performance ranking than Team 3 is justified.

4. Key Findings

For this maritime cluster expert workshop, a model and simulation of the system, the marine crude oil shipping industry, was effective in enabling teams to enumerate possible design variants and to visualize the tradeoffs of various configurations.

To assess cognitive behavioral aspects, during 4 pilots workshops and one quasi-experiment this research prototyped methods for instrumenting the individual’s attention allocation processes. We also observed decision-making and learning aspects of the attention-decision-learning cycle of influence. These sensors and the quasi experiment platform are work in progress.

4.1. Proposed Hypotheses

This quasi-experiment leads the authors to propose a more formal exploration of five testable hypotheses (shown in **Figure 12**). The units of analysis are individuals and teams.

Do higher-performing teams explore more through their design walks than lower performing teams?

H1: Higher-performing teams enumerate more architectures than lower performing teams.

Do higher-performing teams learn more through their design walks than other teams?

H2: Higher-performing teams encounter more surprises than lower performing teams.

H3: The design walk of higher-performing teams contains more path-dependent sequences of systemic relevance, than the design walk of lower performing teams.

Do teams with clear goals learn more through their design walks than teams with unclear goals?

H4: Teams that agree on clear design goals encounter more surprises than teams with unclear goals.

Do teams that approach design problems from different perspectives learn more through their design walks than teams with narrow perspectives?

H5: Teams that focus their attention on more problem variables encounter more surprises, than teams that spread their attention over fewer problem variables.

Figure 11. Testable hypotheses

The quasi-experimental results – not conclusive by themselves - provide insights for a preparation of scaled experiments with controlled testing of hypotheses. In this case:

H1: Higher-performing teams enumerate more architectures, than lower performing teams.

Teams 1, 2, and 5 enumerated 17, 14, and 18 architectures, respectively, while Team 3 enumerated 6 architectures only. Team 4 is excluded from this evaluation, as they approached the design challenge in a fundamentally different way than the other teams.

H2: Higher-performing teams encounter more surprises, than lower performing teams.

Teams 1, 2, and 5 recorded 14, 7, and 10 surprises, respectively, while Team 3 encountered 4 only. Team 4 recorded 5 insights through their analysis, but once again, it is difficult to compare this figure with the others because of the different approach they followed.

H3: The design walk of higher-performing teams contains more path-dependent sequences of systemic relevance, than the design walk of lower performing teams.

This hypothesis addresses the idea that higher-performing teams learn more through their design walks, than lower performing teams.

This hypothesis could not be evaluated, as only Team 1's design walk was studied in depth.

H4: Teams that agree on clear design goals encounter more surprises, than teams with unclear goals.

In the main experiment, all teams seem to have defined more clear design goals at the outset of the challenge, as opposed to what was observed in the previous pilot experiment, and the attention allocation data confirms these goals, except for Team 3.

While Team 3 mentioned a design goal in minimizing Emissions, we observe that Team 3's attention allocation data indicates a higher focus on Cargo Moved (not Emissions). This finding suggests that Team 3 did not focus on their agreed design goals. This could be verified by reviewing the audio files.

H5: Teams that focus their attention on more problem variables encounter more surprises, than teams that spread their attention over fewer problem variables.

In this quasi-experiment, the higher performing teams explored more of the tradespace than lower performing teams.

4.2. Discussion and Lessons

The team performance ranking method used in this thesis project should be reviewed, and more precise algorithms developed. Rather than only surprises, one can explore more types of learning events, as we have seen that not only unexpected results can trigger reflection and learning. Consolidation of insights might also be considered as a learning event.

Machine audio analysis proved infeasible in the experiment conditions, making "manual" analysis the only way to index the recordings. Alternative methods for audio analysis should be developed. Sentiment analysis of audio files could be implemented, whereby validation of surprises could be obtained. Possibly, non-disruptive video recording tools could be tested for capturing/ validating team's mood correlating it to the timestamp of surprises.

All steps of the experiment procedure (incl. registration, pre-survey, and post-survey) should be integrated within one platform that makes the experiment participation seamless. Further work should also focus on the scalability and reproducibility of experiments in an industrial setting for collection of larger amounts of data.

4.3. Future Work

Based on ongoing and future work we will report in detail on the chosen approach to instrumentation, early results, implications for subsequent rounds of experiments and the consistency of observations with other recent literature. We will also lay out the specific testable hypotheses we intend to test before launching an at-scale, reproducible experiment.

5. References

[1] B. R. Moser and R. T. Wood, "Complex Engineering Programs as Sociotechnical Systems," in Concurrent Engineering in the 21st Century, Springer, 2015, pp. 51–65.

[2] X. Ge and L. Leifer, "Design Thinking at the Core: Learn New Ways of Thinking and Doing by Reframing," presented at the ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2017, p. V007T06A025-V007T06A025.

[3] D. A. Schon and M. Rein, Frame reflection: Toward the resolution of intractable policy controversies. Basic Books, 1994.

[4] J. Carroll, J. Sterman, and A. Marcus, "Playing the maintenance game: How mental models drive organizational decisions," Debating Ration. Nonrational Elem. Organ. Decis. Mak., pp. 99–121, 1998.

[5] Y. Engeström, "Learning by expanding: An activity-theoretical approach to developmental research," 1987.

[6] G. Mantovani, "Social context in HCl: A new framework for mental models, cooperation, and communication," Cogn. Sci., vol. 20, no. 2, pp. 237–269, Apr. 1996.

[7] I. Nonaka, "A Dynamic Theory of Organizational Knowledge Creation," Organ. Sci., no. 1, p. 14, 1994.

[8] E. Salas, D. Reyes, and A. L. Woods, “The Assessment of Team Performance: Observations and Needs,” 2017, pp. 21–36.
[9] E. Salas, D. E. Sims, and C. S. Burke, “Is there a ‘big five’ in teamwork?,” Small Group Res., vol. 36, no. 5, pp. 555–599, 2005.
[10] S. W. J. Kozlowski, G. T. Chao, J. A. Grand, M. T. Braun, and G. Kuljanin, “Capturing the multilevel dynamics of emergence: Computational modeling, simulation, and virtual experimentation,” Organ. Psychol. Rev., vol. 6, no. 1, pp. 3–33, Feb. 2016.
[11] S. W. Kozlowski and K. J. Klein, “A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes.,” K J Klein W J Kozlowski Eds Multilevel Theory Res. Methods Organ. Found. Ext. New Dir. Pp 3-90 San Franc. CA US Jossey-Bass, 2000.
[12] G. Chao et al., “Macrocognition in teams: Agent-based interventions and emergence of team knowledge,” presented at the GT Chao & JR Rentsch (Chairs), Building shared knowledge in teams: Shaping macrocognitive processes. Symposium presented at the 28th Annual Conference of the Society for Industrial and Organizational Psychology, Houston, TX, 2013.
[13] Y. Jin and R. E. Levitt, “The Virtual Design Team: A computational model of project organizations,” Comput. Math. Organ. Theory, vol. 2, no. 3, pp. 171–195, 1996.
[14] B. R. Moser, “The Design of Global Work: Simulation of Performance Including Unexpected Impacts of Coordination across Project Architecture,” PhD, University of Tokyo, Tokyo, Japan, 2012.
[15] M. Tambe, “Teamwork in real-world, dynamic environments,” in Proceedings of the Second International Conference on MultiAgent Systems, 1996, vol. Proceedings of the Second International Conference on MultiAgent Systems, pp. 361–368.
[16] T. Alrashed et al., “An observational study of usability in collaborative tangible interfaces for complex planning systems,” Procedia Manuf., vol. 3, pp. 1974–1980, 2015.
[17] S. W. Kozlowski, G. T. Chao, C. Chang, and R. Fernandez, “Team dynamics: Using ‘big data’ to advance the science of team effectiveness,” Big Data Work Data Sci. Revolut. Organ. Psychol., pp. 273–309, 2015.
[18] W. Pan, W. Dong, M. Cebrian, T. Kim, J. H. Fowler, and A. S. Pentland, “Modeling dynamical influence in human interaction: Using data to make better inferences about influence within social systems,” IEEE Signal Process. Mag., vol. 29, no. 2, pp. 77–86, 2012.
[19] A. Pentland, Social physics: How good ideas spread-the lessons from a new science. Penguin, 2014.
[20] A. D. Christian, “Simulation of Information Flow in Design,” Massachusetts Institute of Technology, 1995.
[21] G. Stompff, F. Smulders, and L. Henze, “Surprises are the benefits: reframing in multidisciplinary design teams,” Des. Stud., vol. 47, pp. 187–214, 2016.
[22] C. E. Shannon, “Communication in the presence of noise,” Proc. IRE, vol. 37, no. 1, pp. 10–21, 1949.
[23] A. Berglund and L. Leifer, “TRIPLE-LOOP-LEARNING: AN INSTRUMENTATION MODEL FOR ENGINEERING DESIGN INNOVATION EDUCATION,” presented at the DS 83: Proceedings of the 18th International Conference on Engineering and Product Design Education (E&PDE16), Design Education: Collaboration and Cross-Disciplinarity, Aalborg, Denmark, 8th-9th September 2016, 2016.
[24] C. Argyris, “Single-loop and double-loop models in research on decision making,” Adm. Sci. Q., pp. 363–375, 1976.
[25] A. W. Woolley, C. F. Chabris, A. Pentland, N. Hashmi, and T. W. Malone, “Evidence for a collective intelligence factor in the performance of human groups,” science, vol. 330, no. 6004, pp. 686–688, 2010.
[26] N. Sonalkar, K. Jablokow, J. Edelman, A. Mabogunje, and L. Leifer, “Design Whodunit: The Relationship Between Individual Characteristics and Interaction Behaviors in Design Concept Generation,” presented at the ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2017, p. V007T06A009-V007T06A009.
[27] M. Iluz, B. Moser, and A. Shtub, “Shared Awareness among Project Team Members through Role-based Simulation during Planning–A Comparative Study,” Procedia Comput. Sci., vol. 44, pp. 295–304, 2015.
[28] B. R. Moser and R. T. Wood, “Design of Complex Programs as Sociotechnical Systems,” Concurr. Eng. 21st Century, p. 197, 2015.
[29] D. R. Lessard, R. Miller, and others, “The shaping of large engineering projects,” Int. Handb. Mega Proj. Edw. Elgar Chelten., pp. 34–56, 2013.
[30] Y. J. Kim, D. Engel, A. W. Woolley, J. Lin, N. McArthur, and T. W. Malone, “What Makes a Strong Team? Collective Intelligence Predicts Team Performance in League of Legends,” presented at the Academy of Management Proceedings, 2016, vol. 2016, p. 13564.
[31] T. W. Malone and M. S. Bernstein, Handbook of collective intelligence. MIT Press, 2015.
[32] N. Hirschi and D. Frey, “Cognition and complexity: an experiment on the effect of coupling in parameter design,” Res. Eng. Des., vol. 13, no. 3, pp. 123–131, 2002.
[33] P. T. Grogan and O. L. de Weck, “Collaborative design in the sustainable infrastructure planning game,” presented at the Proceedings of the 49th Annual Simulation Symposium, 2016, p. 4.
[34] V. A. Sakhrani, “The Objective Value of Subjective Value in project design negotiations,” Eng. Proj. Organ. J., vol. 6, no. 2–4, pp. 78–91, Oct. 2016.
[35] Shinnosuke Wanaka and Kazuo Hiekata, “Alternative Fuels for Shipping: A Study on the Evaluation of Interdependent Options for Mutual Stakeholders.,” presented at the 13th International Marine Design Conference, Helsinki, Finland, 2018.