
Validation Exercise Observations of Tabletop Exercises

Matthew W. Potts

1 | OBSERVATIONS OF THE TABLETOP EXERCISE GROUP GIVEN NO COMPLEXITY ASSESSMENT TOOL

1.1 | How the group approached their task

The control group started the exercise by introducing themselves and their background. After their introduction, the facilitator prompted further discussion; "If I asked you what does complexity mean in this scenario, how would you both answer?". One participant answered that there are several sources for system complexity including the operational concept and autonomous functionality.

"We have details about the components...but what we are missing is different operational cases...I guess that the system will be used in different stages from a nominal activity or use case to locate different vessels but I guess there will be a lot of different operational use cases that will have to be managed and be operated and there I foresee that we will need some additional resources...this will cause different impacts; first, we will need some operators somewhere, some maintenance, some calibration...some fault detection, we will need to maintain a level of performance so we will need some team to ensure performance is reached and to repair faults quickly, and it is here where I foresee lots of complexity.[VB]"

They went on to add the following; "Especially when we consider these kind of new systems with autonomy, which means the system will have to adapt to its environment, today we do not have clear procedure to sustain this kind of system." The facilitator prompted further by asking "What is the challenge there?" to which the reply was: "we are taking it [from the scenario] that our organisation is able to produce this kind of autonomous system, but to operate this kind of autonomous system we need to frame the autonomous activity and to bring the system into this autonomous activity, and when to do the transition, how to manage the goals from this autonomous period, and as far as I know there are very few systems, I know subways, but this is all that comes to mind.[VB]"

To probe further, the facilitator asked; "If you were part of this project team, or advising this project team, what would you advise them to do about the lack of clarity about operational concept or operational use cases?". The answer to this question was; "I think we can break down and make an analysis of the operational use cases and the kind of activities we are responsible for and [from these] see what are the functions we need to manage in those activities, and then what are the components that carry those functions and what are the performances that need to be reached for each of those activities carried by those components, and then the idea will be to highlight, or make

the complexity appear or solve it, and then I can foresee is then to elaborate on the different scenarios.[VB]"

The facilitator observed that this answer seemed like a typical approach to systems engineering; "Sounds like a fairly traditional approach to systems engineering...would you see any issues for the scenario system in taking such a reductionist, traditional systems engineering approach?" to which the participant replied "Yes, it is a good point...another point of complexity would be the environment where the system is used, and that is a good point, where you will need to listen to potential changes in the environment and these can lead to some iterative induction on the constraints on the performance requirements which could require some additional attention. I was not considering the 'V cycle' activity, I was considering how we do the state of the art, and regulations, some regulation might change, we have some ambiguous relationships to it.[VB]"

The other participant was asked the same initial question, "what does complexity mean to you in this scenario?". The answer provided was that:

"the complexity, as well as the technical, for which there is a lot of information missing, I know there is an architecture diagram, but it hasn't been showing exactly how they interface with each other so from a technical side of things there are some details missing. The other bit of it is the relationships, it has shown there are some companies named in [the scenario] but is to understanding...the culture, not just between different countries but also the culture of the companies themselves...understanding the relationships and also when this system is put together, so there is different layers of stakeholders...you need to understand that do you have good relationships and history with those stakeholders that you can rely on...The other side of it, what is the impact, or the downstream if that happens and how you control the quality of it, and the integration...always falls down on the integration and that is the most complex bit and need a plan and an agreement and a standard way of doing that, and having a good requirement is a good way to do it...to get that shared vision and communication plan and how they work together, how often they are going to check they are using the same language and mean the same thing. The other is the contract and how the contract is actually written, I see it again and again in different industries, if there is unclearness, no clarity, something will fall in between the two and fall between the gaps, I think that complexity, especially with the clarity and the ability to have shared vision, that complexity adds a lot to it and needs to be clarified, either in advance, or have a locked down assumptions, dependencies or what is the risks.[VA]"

They also answered a follow up question of "are there any other perspectives or factors relating to system complexity [outside of the ones we have already mentioned]...Do you have any other factors that you feel would be important to be measured or to be evaluated in some way?" by saying "I think you have the functional stuff but maybe there should be some non-functional stuff as well, performance is another thing that [sometimes gets forgotten with large teams], so performance is important.[VA]"

The facilitator observed that the ideas presented by the participants suggested a reductionist approach to the evaluation of system complexity and probed further, "I get a sense that some of these ideas are interrelated...when you think about system complexity do you tend to think of system complexity as individual views or is system complexity something broader than individual views that you can compartmentalise?". One participant replied,

"For me, I don't see that they can be compartmentalised, if you can one thing it changes another, so for me, for complexity, you need to at least have a consideration of the boundary of the system and how a change will affect another one and have good communication so that other stakeholders know if changes would impact elsewhere. Yes, we can kind of compartmentalise it, but I don't think in the real world you can, and

that is why it makes it complex, any little thing you start moving it moves everything else as well".[VA]

The other participant added; "since we are talking about interdependencies and transverse functions, functionals allocated to several components, yeah interesting yeah. What else to add? [thinking] the social aspects are also valid for a project, it might impact on the complexity of a system.[VB]"

To probe further, the facilitator asked; "If I can remind you both of the sort of exam question, if you were asked to determine the complexity of the MASSS system and its development project, and suggest relevant mitigation where appropriate, how would you both answer that?" From the discussion that followed, it seemed to suggest that the two participants in this group considered system complexity to mean the level of difficulty or risk in a project that does not require any distinct approach to managing.

"I think we can make a matrix of functions versus components, then in relation with the interface, how they are used between the different phases of a project, then identify the ones that are impacted by some of social issues and also identify where we have the most stringent performance constraints and at least highlighting the interfaces which are important.[VB]"

The other participant answered:

"I would ask the stakeholders involved to do some estimations and risks and issues and assumptions. The other thing would be to look at the lessons log and look at previous projects...and with our supplier...to try and see if they can answer the questions and select suppliers accordingly. I think supply chain can do some work to understand their commitment. A lot of the time I see organisation choose the suppliers with the lowest cost which is not necessarily the best for them. And have a good kick off meeting so that everyone is on the same page and [work] together.[VA]"

The facilitator tried to explore further; "What seems to strike me, is that you both seem to advocate an approach that I would recommend for the engineering of any system, I wanted to check if you feel the approach to engineering a complex system would need to be any different then a non-complex system? I'm trying to get a sense of if you see something in this scenario that makes you say 'hold on a second team, we need to do something different here because of something?'".

One participant replied:

"So we are in the complex system, means if it is complex it means that it can be analysed, we have the tools, but it takes effort, time, but we have the tools, I'm not in kind of chaotic system where we cannot control it, so in that sense, yeah, we need a failure analysis, what happens if one of the component fails, if one of environment, stakeholder is doing something that is not expected, we need to analyse the performance at the interface, but we are in complex so we need to be going into this risk analysis as we did and then on the breakdown of different levels of the architecture views of the system, we should be able to 'de-complexify' this system.[VB]"

The other participant was also questioned; "If this [system] is deemed to be particularly complex, would you recommend [the project team] do anything different?" to which the answer was; "I would, I think it is really important to think about our strategy and vision, is it worth to joint venture a partnership or something like that so that you

can shift the dynamic...so we have the same vision and it...drives a different behaviour... Do you want to invest more money to change this project into a product, so that next time you want to sell [the sensors] again you can do so much quicker.[VA]"

The facilitator prompted the group to consider mitigations for the issues they raised, with participants responding that if the problem is complex, that they would approach mitigations in a conventional systems engineering approach, for example, developing operational use cases, identifying functions, assigning components to that functionality, etc.

The facilitator asked participants to provide in the text chat an answer to the question "how complex is the MASSS system and development project?". One respondent answered; "Scale of 10/10 particularly the risks in ensuring the right skillset and good team dynamics.[VA]" which was further explained verbally as:

"in my past experience, it is complex because there are lots of companies and teams involved in it, so the governance is a serious issue, making sure things are there and they are working to the same time and the same level,..., and I think a lot of it is because the communications fall down, or the skills or misunderstand the requirements, and so I think it is very complex and taking care of every part of it, especially the integration part of it, because one team can do really well but if the other doesn't it will affect it....What I see is that sometimes the customer or manager does not see how complex the issue is and they have some unrealistic estimate...[VA]"

The other answered; "Not so complex means 100 engineering". This respondent further explained that when they feel the system is very complex, they would assign a "target solution view" and then plan and deliver some "architecture milestones" to reach the targets.[VB]"

1.2 | How the group approached their evaluation after the scenario inject

After the scenario inject, both participants sought clarity over the impact of the inject on the system boundary and operational scenarios, asking "What is not very clear is how the system is extended...Just wondering what the system boundary is now? [VB]" and "Is it only a new technical configuration or are we talking about some wider use case, wider scope, wider deployment? [VB]". The respondents were then asked, "How would you evaluate the complexity of this system?" to which one participant replied "I think, since the perimeter is not the same, I would take it as a new system and see what components from the initial MASSS to reuse, and how it fits, and what are then the opportunities for reuse but also clearly remove the ones that do not fit and which must be adapted. So once we have to do that yes things become complex because we are biased by the previous system and for sure from a company perspective you want to reuse as much but from an architecture point of view you want the best solution so then comes the discussion about trade off, alternatives. It is not only a line of product question, we are talking about a new system, so we have to take care about that. The mistake would be to consider it as a simple reuse. [VB]"

The other participant answered;

"I think it was very complex before but now it is definitely more, especially if, we don't know if the company have worked with airborne systems or the company before, if they are new things you don't even have a lot of history that you can look to see you know if you put a risk factor in it, is it twenty percent, is it forty percent and you have to strike new relationships and basically work from the beginning, so yeah, definitely much more, and I think it doesn't help that they want to move forward the delivery date. So yep, that definitely will be challenging....If this company have never worked with these suppliers, or needs the

right engineers with the right experience, maybe you need to leverage some expertise from some other company I don't know.[VA]"

To try and clarify further, the facilitator asked; "I get a sense from the way you have both answered that question that there are no new perspectives on complexity to consider here, so again we have to consider the stakeholder relationships, and maybe that is even more difficult, but that is a factor we have already considered and contractual complexity,...operational complexity...technical complexity. I wanted to check if there was anything here in the inject that made you think we need a new perspective on complexity?". The reply from one participant was "I guess if they have a new system, is this a new system that they have never done before? It is another thing that brings on a lot of risk...," further explaining that "I see another risk in terms of complexity, is that the ambition of the company becomes quite ambiguous, there were starting from the development of a detector to a wider integrator, now it looks at they will be providing natively the integration of their system onto some other ones and you can have the leadership for one project, but to be driven by the service of integrating...you become the project leader for the integration, the owner of the integration for the complete system, then it becomes quite difficult to have it as a normal business case for a company...means they have to instantiate the authority against the different partners and it is going more then leadership and it becomes authority and it needs to be completed embedded within the company in several different aspects of the company that they have to develop.[VB]"

At the end of the session the facilitator asked; "We have used the term risk quite a lot...I wondered if you had a view on how system complexity relates to system risk?". One participant replied "I think it is hand in hand, it is definitely hand in hand, if not proportional possible exponential, in the more complex it is, the more risk...[VA]". The other participant replied "I see the inability of mastering the complexity, for me, will make the system more sensitive to the risk and to the consequence of a realised risk. So, to have holes in the mastering of complexity...so it will increase the impact of the failure. [VB]" The same participant also reported that; "we need to consider the MASSS and inject system as, it is not only the MASSS and the technical system, it is also the company and the enterprise that is delivering the system that should be under the study. [VB]"

2 | OBSERVATIONS OF THE TABLETOP EXERCISE GROUP GIVEN THE "COMPLEXITY PROFILER"

2.1 | How the group approached their evaluation

The group of Thales Group personnel that were instructed to use the "Complexity Profiler" started the tabletop exercise by introducing themselves and describing complexity issues from the scenario that resonated with them. Interestingly, not all of these identified issues fit neatly into the system complexity factors of the tool, for example "certification and regulation considerations", "data fusion", "autonomous behaviour and the impact of autonomy on verification and validation activity", the "number of interfaces" and "organisational expertise".

"The thing that interests me most from the scenario is the interactions between the elements that we control and the elements that are outside of our control, especially when it comes to the PANTHEON operating centre relationship to our own TAU operating centre, and how our purple box and green boxes fit within the other boxes and relate to the other boxes. And the other element that interests me is the novelty of the technology and whether that is going to introduce any additional issues for us.[VE]"

One participant could not provide voice or video to the group session and instead provided text responses.

"The key points that I pick up are: Number of interfaces, Some technology is a high TRL while others are relatively low (i.e., a mix), The environment is quite challenging, There are a number of stakeholders and getting agreements between the parties is always challenging...So, integration as well as some of elements of the design will be tricky and will need thinking about upfront, i.e., how we are going to accept the system....One other point is that we tend to build our systems based on what we know/have done in the past, we tend to focus on the 'easy' bits and miss out on the 'complex' difficult areas. Which never ends up well.[VD]"

Another participant prompted:

"Sensors are quite high TRL so it is ok but what is not really and not really mastered is the data fusion of these sensors so probably something to take into account. Another one is about the autonomous behaviour of the platforms, so when we know that such behaviour is not so simple to manage, in particular, when it comes to verification and validation and also mostly certification and there are probably certification or regulations that we have to follow in such domains, so nobody talks about regulation, could be major issues. So anyway, autonomous behaviour and the validation of such behaviour is at the time something that is absolutely not clear so we probably have here to be very cautious. We have already noticed that there are some different partners, suppliers and other stakeholders, but something that looks a little bit strange is about the power supply, which doesn't appear in the overarching architecture and that power supply seems to be an internal product but it doesn't look like our company is particularly experienced in such products so probably that power supply could generate some questions and problems. So we all know that power supply is not very high tech item, but anyway without power operations are completely closed so even if it is simple it is of paramount importance for operation.[VC]"

Next, the group worked through each system complexity factor in turn, and tried to assign a score. The first suggestion was that the first system complexity factor "Impact of Environment on Solution" should be scored '4', "(Critical) Major physical constraints and operational impact" as "the operational environment is difficult; i.e., having to operate a data link on the platform, and also getting the system to work on such a platform. The equipment will also be subject to a potentially harsh operating environment[VD]".

However, another participant suggested that this factor will not be so difficult, as they saw nothing in the scenario that suggested the environment will have any impact beyond what the organisation has mastered already. The other participant raised an ambiguity they saw with the "Complexity Profiler", "my first question is, what are they talking about in terms of 'environment'? Is it literally the physical actual operational environment or are we talking about all of the context in which it sits?[VE]", answered by another participant who said, according to the guide, "it is talking about the physical environment where the system operates...these physical constraints may have an impact on the system itself or the operational concept and the difficulty to bring that to operation [VD]". This same participant went on to argue that although the issues suggested earlier (e.g., operating in a harsh environment) are valid, that their impact is lower as the organisation has experience delivering such systems, "...knowing that it seems that our company is already used to producing such things [VD]". Another participant went on, "[the scenario] does explicitly state that they have overcome issues with [various issues] so that has been overcome already, but I think there is a question over whether the environment would affect the SATCOM and VFH, the communications between our thing, because

that is outside of our control.[VE]"

At this stage a respondent noted that "such an exercise is really interesting too to capture what is said", "because what is interesting is get not only the result in the profiler but also the rationale.[VC]" The group went on to refine the first entry for the "Complexity Profiler", although one respondent reported "So I don't know what the difference between the hypothesis and rationale here is, I need to go and look at the guide again.[VE]" The participant who originally suggested that the value is a "4" offered that they "would be happy to go to a 3" before adding "But has it been proven to operate on this platform? [VE]".

One participant reported an assumption that in the scenario, the organisation will work with subject matter experts; "in order to mitigate the effect of environment on communications we are going to work with SMEs in this area. However, we have had mixed results in the past.[VE]" The participant went on to raise a concern that "my gut feeling, is that 'demonstrated in an operational environment' doesn't necessarily equate to 'will work in that environment for a prolonged time', does it? I don't know. TRL 7 to me is sort of, yes you have done it once but there is no evidence you can do it over and over. I think I'm with [other participant] on this and I think it is a three[VE]". Another participant replied that:

"...for me the questions is not that much TRL 7, the question is much more about what is not yet tested or proven in the operational environment, like the processing chain which is new, and if we think about data fusion, and also another point which is linked to this environment is the concept of operation, as far as I remember, there is nothing said [in the scenario] about the impact of this environment on the concept of operation. When we are talking about an autonomous platform we don't know at the time if that autonomous platform will withstand high sea levels for example, and if it has an impact on the concept of operation, because nothing is said [in the scenario] about concept of operation. So there is probably a question mark over the impact of environment on the concept of operation...The fact that this is an autonomous platform is not expressed clearly in terms of the concept of operation so there is still a gap. I think this point will also be raised when we discuss the concept of operation. [VC]"

After a question from the facilitator about how the group would like to proceed, whether in agreeing a score or first discussing mitigation actions, the group agreed to score the complexity a '3'. One participant suggested "I think we assign the values and then look at the mitigating factors separately because you find mitigations that address multiple issues [VE]" which was an approach agreed by the group.

The group moved on to discuss "Operational Concept Stability", with one participant asking: "Is it fair to say that, as long as nothing is said [in the scenario], that the impact on the concept of operation of the autonomous platform is unknown so it maybe a great impact, so this is an unpredictable major evolution, what do you [group] think?[VC]", which was counter argued by another participant who interpreted "it seems relatively stable, I mean this is a concept of operations that these nations and organisations have been doing for a number of years, it doesn't seem like they want to change what they are actually doing, just the way they are doing it.[VE]"

The facilitator asked what value should be assigned to this complexity factor, and the reply was "I don't see anywhere in the documentation where it actually states, other than the expansion of the coverage area, ah, so, here we go, these are the changes, they want to cover areas further from the coast, increase coverage, temporal and spatial resolution. So I think that is probably a '2' because it seems like fairly minor things because they aren't changing the actual operations they are just changing the timings and numbers of those operations.[VE]" These points were agreed by the third participant who offered "I agree. It is more of an integration issue. I would go for a '1' or '2'. [VD]"

The facilitator asked the participant who had originally suggested that the operational concept stability is critical to

reiterate their point “my view was that we don’t know what the impact is of the autonomous vehicle on the concept of operation. So, maybe it will be favourable but maybe also it will be a disaster. But we know the company has mixed results in terms of using this vehicle, but nothing is said [in the scenario] about the impact on the concept of operation.[VC]” After a continued discussion, shown in the appendix, one participant stated:

“[from the scenario] TAU has been conducting maritime surveillance operations for a long time, autonomous capabilities have only recently emerged as a feasible option, so at the time we are only at the feasibility stage, so the impact on the concept of operation is not mastered at all. Of course, what is expected is a gain, as such being able to reach the capability that is expected which is full coverage and so on and capability for 24 hours, seven days a week, but at the time this is not proven. So I would say we simply don’t know [VC]”

to which another participant replied “I guess my question is, does going from manual operation, a manned vehicle to an autonomous vehicle change the fact that they want to do 24/7 days a week? Does it change the mission?” to which the reply was “it may change the mission and also the responsibility. Who is responsible for performing the piloting of the vehicle, even if we have seen there are different modes from manual to full autonomous, but for me this is exactly what we mean by concept of operation, how will you use the different platforms and how you will use the interactions to fulfil the mission.[VE]”

The facilitator asked “what is the risk or issue for the scenario organisation?”, which was answered in terms of increased responsibility by the supplier of the system, and that ultimately it is not well known, “...so the maturity of that concept of operation, while autonomous, I think is still very low.[VC]” The reply to this was “So, I have just read through the guidance, and that is what ‘minor evolution says’, which is ‘the evolution of the operational context is planned and described in one story,...are defined or can be defined accordingly. So I feel like the step from manual to automated you have just done in one story. Yes there are some steps in the story but you are going from manual to automated and we know that there are going to be steps, and if we really wanted to we could probably go through them and define what those steps are in terms of what responsibilities are transferred when in each system capability increment, so when the system autonomy will be transferred, so we probably would go in in terms of a step where we would only allow the system to be operated in remote control mode as opposed to full autonomy mode, or something I don’t know, but I still feel like it is a minor evolution...[VE]”

The other participant supported these claim, “But the concept of operation is well defined and understood. It may be in its infancy, but the components of the system (i.e., subsystems) have been proven in isolation. The organisation has to be flexible to support changes in the CONOPS. I agree, it’s like I said, the ‘building blocks’ of the system are largely proven already in other systems, to one degree or other. It is a ‘2’ for me [VD]”.

There was praise here for the discussion that was enabled by the “Complexity Profiler”, “I think this is a very interesting discussion because I try to take exactly the opposite viewpoint as you, trying to promote the fact that autonomy has an impact on responsibility and for example, there is also another impact which is user experience or at least user interfaces. Nothing [in the scenario] is said about that. So, we have absolutely no clue about this concept of operation and what impact it has on the manning of the system, of the usability of the system and so on[VC]” to which the other participant accepted “based on that I am happy to accept that there are several threads to this...one of those is the actual technology improvement, but there is also the resource, training and people who were going out on boats and are now sat at computers and we don’t have an idea of what their operational awareness needs are and things like that. I’m happy to do a ‘3’ and move on.[VE]”

At this stage the facilitator wanted to probe further around the prevalence of these kind of discussion within

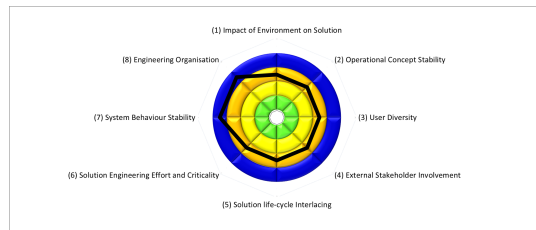


FIGURE 1 Completed “Complexity Profiler” for the MASSS System. See Appendix Q, ??, for details of the rationale for these values.

“business as usual” and asked “my question is if these kind of discussions are fairly typical in an organisation, so I could imagine that these are fairly normal conversations to be having?” to which the replies were “in my opinion, this is not so frequent, just because you can do when you are running the ‘Complexity Profiler’ but is it really so frequent, so systematic for each bid, each project, I am not so sure [VC]” and another “I don’t think these discussions happen often enough [VE]”, conversely another said “I think they are, but the issue (as we are seeing) is that there are different viewpoints and it can be that some of those viewpoints are not heard.[VD]”. “I asked someone on a separate issue earlier, whether we did the...and I asked do we use the ‘Complexity Profiler’ and you know do we have outputs and do we track them for all of the projects and the answer was ‘no, we don’t’. So it isn’t, certainly in [REDACTED] I don’t believe it is used so much. [VE]”

When the facilitator checked that the group were happy to assign a value of ‘3’ to this complexity factor there was agreement, and one participant added “I think one of the things that this proves is that it is really difficult to actually get agreement on some of these things, even when you have got a relatively straightforward scale [VE]”, to which another reinforced that “what is important is not so much to agree on a value of ‘3’ or ‘4’ or whatever, the real important thing is first to capture the rationale, and most of it then to track what will be the action plan and the results of the action plan. And capturing and putting in place the resulting action plan is what makes sense. This is what is mostly not really done. [VC]”

The group continued in the same way as before for the remaining system complexity factors. The completed “Complexity Profiler” is shown in fig. 1 with the completed proforma in Appendix Q, ?? The group spent time discussing their rationale for their value and tried as a group to reach an agreement. During these discussions, time was also spent describing their interpretations of the system complexity factors and their interpretations of the assessment scale for each factor.

2.2 | How the group approached their evaluation after the scenario inject

Unfortunately, due to time constraints with the tabletop exercise, the group could only spend ten minutes on the re-evaluation activity after the “scenario inject” and were asked, “I would instead like to ask you, what sticks out to you about the inject, if anything resonates with you, and how you would approach a re-evaluation”. One respondent answered;

“I quickly ran through all of the categories, it drives everything to a ‘4’...that is my assessment, it changes the CONOPS, it changes the operating environment it changes the design requirement, it changes the number of stakeholders and users, it changes everything. Oh but do it quickly, do it by last week. And we have already got problems, so those areas that we were trying to argue downwards that we were saying

'actually our technology is further on then we thought', well actually no, now we have actually started to see some issues around the design work so that pushes us back up. [VE]"

The group agreed to raise all of the system complexity factors to a value of "4". The discussion turned to mitigating actions. This group did not examine or consider the mitigating actions prompted by the tool, perhaps due to time constraints, and instead the group discussed mitigations such as recommending "a staged deployment of the scenario capability [VE]".

3 | OBSERVATIONS OF THE TABLETOP EXERCISE GROUP OF THALES GROUP PERSONNEL GIVEN THE "COMPLEXITY REGISTER"

3.1 | How the group approached their evaluation

The Thales Group personnel that were instructed to use the "Complexity Register" also approached the tabletop exercise by introducing themselves and then identifying issues that resonated with them from the scenario documentation.

"One interesting thing for me, is they seem to have a product already, if you will, that they are going to use, the front end of a system, but there is also an interface to an existing system and that would be an immediate concern to me because there is obviously ways of working already in the control centre and trying to merge data flow into there. Also in the communications there is nothing in there about security, either data validity or data security from the centres back down to the control centres and slight concerns over the SATCOM VHF aspects, they don't have expertise and they want to get them by an external company, we know how that normally ends up unfortunately. [VF]"

"The things that jumped out to me was that the description of the system that we are offering doesn't have any description of how it meets the measures of effectiveness that the customer has got. I'm unclear about how the AIS integration on the ship works so how do we know how many tracks we have got and it sounds like you are limited by whatever COTS receiver you have got, total number of tracks, doesn't mention that either for the EO tracker or on the radar and it doesn't talk about surveillance volumes either so you don't know how many vehicles you will need, so whether or not we have got a product that is fit for purpose or what it does state if what we have got is a product that is brimming with the latest technology and we'd love to sell it to the customer to show them how good we are at the technology. [REDACTED]. [VG]"

The group started by discussing how best to approach the problem; whether to start entering directly into the "Complexity Register" issues they had identified from the scenario documentation or whether to start by using the aide memoire to identify complexity issues first. The group agreed to enter the issues they had identified from the scenario document first, then use the aide memoire as a "completeness check".

"[talking to other participant] In doing through the "Complexity Register" stuff the way I read it, you can sort of come at it from two angles, there was the steps or the aide memoire at the back. and they both come at it with slightly different angles. I'm not sure how best to integrate those two angles. Did you have a feel for how we would use it for the best? [VG]"

"I was looking at it from the, I know the aide memoire is there, but due to the word count in the aide memoires I'd try and stay away from that and use the spreadsheet column headings as almost, and treat it like a risk review, and say instead of 'there is a risk that' and change that to 'the description'...[VF]"

"In order to use the spreadsheet, the spreadsheet sort of pre-suggests that you have already done an amount of intellectual effort to get to the issues, cause its describe the issue, and the aide memoire helps us get to those issues

doesn't it?[VG]"

"It does. But then to me it seems like it could obscure the thinking a bit.[VF]"

"I'm happy to do it that way though, if we want to jump in and put down the issues we think we have got and then go to the aide memoire and use that as a completeness check? [VG]"

"That is how I would have approached it if I was doing this in my team. [VF]"

"There are some obvious things that jump out from the text so we can get those and then we'll probably be a 60 or 70 percent of the main issues I would have thought. [VF]"

The group proceeded to describe system complexity issues they had identified from reading the scenario documentation using the headings provided by the "Complexity Register" including mitigating actions.

"...the description of the issue from my point of view, I suppose taking your first point, is the fact that you don't know how you are measuring the performance of the system, so there is a lot of things, a to f, of key criteria, but without knowing that in detailed terms, as you said, how do you know the existing sensors and control systems are actually coping with that level, so there is an area of complexity there that we cannot define at this moment in time. [VF]"

"Yeah, I think if i was doing this with more then an hour I'd break that down into the individual bits but with the hour we have got, lets just stick with that [VG]"

"So to me, the impact of that issue is that we won't be able to validate the design so this actually becomes an acceptance point, you know how does the customer actually agree that we have not done something correctly or how do we tell the customer we have done something correctly. [VF]" "Yep [VG]". The facilitator prompted "Is there a suggestion of a mitigation?" to which the reply was "The mitigation would be that we need to break the measures of effectiveness down into the, do some systems engineering on it, some derived requirements and at least do some modelling of the expected performance of the solution against those, even at this ITT stage. [VG]" and "Or, you turn around at this stage and you define what you can, what you are capable of providing, and maybe that would be the way of doing it at the ITT. But yeah definitely I wouldn't want to go any further without knowing what you are signing up for. [VF]" The facilitator prompted the participants for answers to the remaining columns of the "Complexity Register" (What is the extent or scope of the issue? Does the issue implicate other processes or activities? Who should be in the loop for this issue? How likely is the mitigation to succeed? When should consideration of this issue be revisited? Who has responsibility for dealing with this issue?). See APPENDIX X for an extract of the completed "Complexity Register". Part of this process included discussing the question of how likely is the mitigation to succeed? "What do you measure the success? How likely is it to succeed to weed out all of the issues, probably would do, probably do a good job, but would that mean you are going to be over expensive or not meet the customer expectations on the budget and all the other things, probably depends how you define measure of success? [VF]" "Yeah but if you said the measure of success was that we made a good Gate 1 decision, so in Thales speak, that is we know if we think we should bid on it or not, then this mitigation action will ensure that we make a better informed decision. [VG]"

The group continued to identify complexity issues, such as; "So I'd say that there is a stakeholder issue of really understanding the requirements because it seems that PANTHEON are doing all of the procurement activity but the requirement is solely from TAU. So this is now linked to the issue above. So whilst the mitigation from the mitigation above is we need to get to understand what the user needs, this issue is that the way this programme is built, there is always going to be an element of obfuscation between our understanding and the real user needs. [VG]"

Another identified complexity issue was; "So, one suggestion that I've got is the make team buy type of decision you are trying to make on the SATCOM UHF, might get a bit technically detailed, but essentially that is an uncertainty that they don't have a solution for and experience in. So again, in the pre-ITT stage I would want to make sure we have got a good understanding of what that link is and or who is actually going to be providing it. And the customer

might have a preferred solution he's already driving SATCOM UHF back into our existing systems. [VF]"

Another identified complexity issues was; "...the environment the system is being used in, and I don't mean things like the sea state or the temperature, but I mean the social and challenging acquisition nature of the system. So, to me, that could mean social as in there could be an issue with the members of either the TAU nation not liking being monitored, or it could be the nation members of ALPHA and BETA monitoring them, but there is a social issue there and that can add complexity to the solution for clarity of data or sharing of information, people cobbling the boats if they don't like the idea of it. [VF]" "There is definitely something missing in the socio-political environment between ALPHA, BETA and TAU that is not in the text. [VG]" "...it is data transparency basically. You know how much are they actually sharing with PANTHEON as well, you know do they share every single little bit of data they get, how does PANTHEON know that the data processing they have asked us to do isn't going to reduce the clarity of position that only TAU knows about. There are all of these very difficult to put in one line [on the Complexity Register], it becomes a big amorphous blob of data you know, big data. [VF]"

This facilitator later stated "I'm just conscious of time where we have 15 minutes left to finish this exercise, and there was a suggestion earlier that you might be tempted to brainstorm a few issues and then use the aide memoire as some sort of completeness check or some kind of verification activity, I just wanted to check with 15 minutes to go how you would like to move forwards?" to which a participant replied

"I've still got a few more, I don't know whether we want to capture them as grouped up ones, the extent of the integration with the COTS vessel, and there's all kinds of aspects with that, the power control, is the navigation stabilisation data going to be good enough for the sensors we've got, there is probably an hours work breaking that down in its own right. And then there is one about sensor performance, in that they are only TRL 7 and yet you are selling them. And then there is another one about, hidden away in there, about multiple design teams, and the extent to which they can really operate as one design team. But that is probably all the ones that jump out. But I think it would be worth, perhaps rather than doing the detail for those, going back and seeing whether we've picked up all the issues using the aide memoire, what do you think [REDACTED - OTHER PARTICIPANT]? [VG]"

"Yeah, happy to do that. I've just looked at the first one as you were putting that out, and the first one is the number of diverse modules, so it is quite an appropriate one, so the number of subsystems and components and how diverse are they and interconnections, and if you point out on that system diagram, there are things missing from that systems diagram from that architecture diagram, it doesn't mention the C4I system for TAU itself, so it is easy to potentially miss something off. And the encapsulation of the system as well, I think from a level of boundary you are ok because you have the different sensors, but then the data captured from the sensors has to go through the data processing on the boat itself, it has to go through the SATCOM link, it has to go through a groundlink, it has to be displayed by a C4I system, so you don't actually have a lot of data encapsulation there for the data, so the data is probably very [inaudible] so that is probably one area of complexity that might have been brought out. [VF]"

"If we look at A4, [REDACTED - OTHER PARTICIPANT] I don't know about you, but that sticks with the fact that we don't believe that there is enough requirements around the functional and non-functional requirements of the system [VG]", "Yeah, I think we have got that one, that was in the first one [VF]".

"Lifecycle might change, again we've mentioned that quite well, we've mentioned the gate but also further through the engineering development lifecycle, so I think we've considered that. [VG]" "Sorry where are you looking? [VF]" "A6 [VG]" "Yeah I think we've talked about the stability of understanding, agreed. [VF]"

"Ah, users, now this is something that I didn't quite, I couldn't quite get my head around when I looked at the

architecture was, it comes back to the measures of effectiveness I think, I couldn't actually get a handle on whether it actually reduces the number of people of having people on the boats on the briny, it looks like we've got just the same number of people doing data processing and stuff they just happen to be on the land. [VG]" "Suppose it depends on the number of users, you are reducing the number of users at risk by not having them on the boats out at sea, but then those boats are probably going to need more maintenance or inspections when they come back to dock as well.[VF]"

"I think this was a throughout, there are all kinds of elements of the integration that worry me. The power supply sounds like its novel with a link to the C2 but with independent links to knock the power on and off the sensor suite, that power integration into the vessels battery just sounds, anytime you try to break into something thats COTS it is never as easy as you think. [VG]" "But also there you've got some unidentified things, so you don't now what the PANTHEON thing is, what connection system you have to use, say we have our four sensor heads, can they actually communicate with the data processing system or are they all stand-alone, so there are areas we can define ourselves but there are also some areas where we cant define without some input from the customers. [VF]" "Yeah yeah that is a good point, there are some interfaces that are within our gift to define that we don't know and there are some that are outside of our gift to define that we don't know. [VG]"

The aide memoire seemed to prompt consideration of areas that might not have been uncovered. "So the regulatory, B4, I didn't even think of that one. [VF]" "So certainty, you picked up the security if the SATCOM element and I think we've picked up the security of the C4i connections in amongst everything else, I think we haven't talked about whatever maritime autonomous operations regulations might exist in this region, and how mature the regulatory environment might be. [VG]" "Or is this actually ground breaking and they have to change their own laws in order to allow these MUAVs to be able to operate. [VF]"

There was also some confusion over one of the entries in the aide memoire; "So what does B5 mean then [facilitator]? [VG]". "So I believe on the following page encapsulation is explained a little better. So I think that the gist of it is that, so you could imagine, that maybe this platform, you say well it is just a COTS platform it is just a black box, we will just buy it, it is fine, it has some well defined interfaces, we will just buy it, plug and play and off we go. But there might be some difficulty in there, in the same way the system deployed in its operational environment, the system might have to interoperate with other systems, where on the face of it, it should be well, 'well of course, these two systems will have to interoperate, that will be fine, they can be treated as black boxes', but maybe sometimes that is a dangerous assumption to make. So I think encapsulated can be replaced with the term 'black boxed'.", "Ok, ok [VG]", "So I hope that helps, and it says an interesting point to me which says that this is a vague term." "But I think I get what it means, is the system making use of black box technology with well defined interfaces, and the answer I can see here is I cant see anything that fits into that category as I look at the architecture, everything looks like its got a development of an interface for the context that we are putting it in. [VG]"

"And [REDACTED] I think for B6 this shouldn't actually change too much during the project, because its the technical aspects so as long as the requirements we set at the outset, they shouldn't be changing too much during the project, they should only be changing due to an operational change or a lack of clarification in the earlier stages? [VF]" "Yeah I guess there is a risk that the, because those countries C4I systems will be doing more then just interfacing with our system that they'll have a lifecycle of their own won't they [VG]", "Yeah, so you could have an outstep of technologies or an outstep of interfaces [VF]", "yeah, which you might see as a risk or you might see as an opportunity to make money down the line. [VG]"

3.2 | How the group approached their evaluation after the scenario inject

After the scenario inject, the group documented the main system complexity issues they perceived had arisen. "I guess what we are trying to do here is now, is how can we use the Complexity Register which we did at ITT time to re-asses complexity deltas and any delta actions right? [VG]", to which the facilitator replied "Correct. I wonder if a good place to start would be things from the inject that really stand out to you or really resonate with you from your prior experience.". One participant provided; "The fact that they didn't admit that there was a planned upgrade to the PANTHEON C4I system and they now need to, or now they want it done or improved to improve the search and rescue capability. They want the inclusion of the airborne interoperability with MASSS and increased operability and functionality, those three things stuck out like a sore thumb and they'd probably turn around and go 'oh and there is no extra cost right because you are already doing the work?' ...And if the information available for any of those three is in the same state as when we did our original ITT then they haven't a clue what they are doing and they haven't got a refined requirement set. So alarm bells basically. [VF]"

The other participant suggested that the guidance material for the "Complexity Register" prompted consideration of compounding issues.

"So I think the way I would use the register here, if we were doing this for real, and I don't know if it works here or not though, is to take each one of those injects as a new complexity issue. So be it that they want it faster, be it that they got a C4I, be it that we have not emergent and persistent issues with the vessel ICDs, or be it that we have got a new interoperability requirement with a new airborne system or be it that we have got a new towed array to integrate, and then, because I think I saw in the guidance, the concept of compound issues, where one does another. So if you took all the new ones and then I think I put them in a matrix against the old ones and see if they stand in their own right or are they contributors to existing complexity issues. And then perhaps go through and modify the impact and the mitigations...Having them in a table like this perhaps doesn't lend itself...You know it probably would if we were all sat around a white-board wouldnt it, but the environment doesnt lend itself all that well. But we could do couldnt we, there is only five new things, we could just taje them and do a quick, does it make them a bit worse, does it make them a lot worse, or does it not affect any of the existing issues we've already got and then say what are the new mitigation strategy. [VG]"

This approach was queried by the other participant; "So you are treating this almost like a risk review where you are refreshing the risk and saying how the risk has changed during the course of the project, so treat it the same way, has the complexity got worse, has it been mitigated. Yeah ok...[VF]"

"So lets create the, how did they say it, it was essentially about the interfaces to the vessel. oh adequate design information, so lets call it 'vessel design information inadequate to progress the hosted system design, so that'll be the C2, the power supply, the EO, the radar and the comms, but other then that its probably fine [laughs]'. [VF]" "So rather then look at the impact in its own right, [facilitator] can we somehow go back and say which of the existing complexity issues it related to? [VG]" The group then went through the original complexity issues raised and updated these entries in the "Complexity Register". For example, for the first raised complexity issue, one participant offered "I know we've mentioned this a few times but this is almost like a risk register approach, you would take this one [first entry in the Complexity Register] as worsening wouldn't you? With mitigation not working. [VF]", "yeah it is analogous to that but using complexity as a way to think about risk isn't it?[VG]". They then decided to examine compounding effects of the new developments from the inject on their previous evaluations, the results of which can be seen on

the extracts from the completed “Complexity Register”.

4 | OBSERVATIONS OF THE TABLETOP EXERCISE GROUP OF NON-THALES GROUP PERSONNEL GIVEN THE “COMPLEXITY REGISTER”

4.1 | How the group approached their evaluation

The group of non-Thales Group personnel that were instructed to use the “Complexity Register” also began the tabletop exercise by introducing themselves and identifying issues from the scenario documentation that seemed pertinent to them. However, for this group, the first participant to introduce themselves also raised the question of what the term “system complexity” means to the group; “[the scenario] sounds like an interesting thing, notable that the users were part of the scope from the customer but are going to be the main part of the complexity, it is going to be interesting in terms of how we interact with that. But also what our collective definitions of complexity is, might be an interesting point to start.[VJ]”

Another participant introduced themselves, and added “in some respects it looks an awful lot like a lot of tasks that you would be asked to get involved with as a systems engineer. So someone in your business has gone and done a bit of business winning or business development and said, we are going to do this, and before you as the systems engineer have had a chance to get involved, they go, we are going to tie all these systems together, bundle them up and make it work. And I think there is a big element of that here. I think the big concern that leaps out at me is there’s a reasonable nice sort of picture for the grown ups, systems architecture, does it have all the information in it it needs? No, nowhere near. But it is an interesting start and I think it’ll tease out some of those complexity issues...[VH]”

Another participant introduced themselves; “It does look quite familiar, walking into some of the projects we walk into, resonates with multiple stakeholders with lots of interests and what they want to get out of it is what it feels like to me with lots of different companies and countries, and that is going to add different levels of complexity, not just in terms of the system but actually what the stakeholders want to get out of it, so I’d be interested in when we use this tool, how does that pull through? [VL]”.

Another participant introduced themselves; “I agree with the earlier comments that it looks like a salesman pitch for what the system would look like, there isn’t an architecture there is just some little pictures in boxes, so that’s not really giving any indication of complexity and the descriptions are using all the right words but are not highlighting in my mind where the difficult problems are going to lie, so it’s often the non-functional requirements and that sort of thing where systems engineers will get caught, the functionalities are normally fairly straightforward. So it’s those sorts of areas where I would be looking for gotchas, particularly safety and security would be one of the first ones I would worry about and then security. This thing will be driving around by itself, yeah.[VI]”

This group then discussed what system complexity meant to them, the question originally asked by the first participant to introduce themselves. “I don’t know how other people feel, but is it worth quickly talking about what we think complexity is, so when we are talking about complexity we are talking about the same thing?[VJ]”, another participant replied “...I had to ask two questions in my introduction which were triggered what you just said, so it would help if I at least pose the questions for all of us to keep in mind. So as [REDACTED] just mentioned, one of my questions was how do you define system complexity in the context that are discussing, because inaudible, I found a definition in several places inaudible, and the other question was, who is going to finance the creation of the system complexity register, because when you work on bids and proposals the scope is much more restricted inaudible, what the realistic scope of what the complexity register addresses which can be your own system of interest up to the system of systems that are not in your control, and how are you going to fund this complexity evaluation if inaudible,

I mean the Complexity Register is as good as the people working on it and the scope inaudible of the Sol.[VK]"

"So I agree that the Complexity Register is as good as the people filling it in and the scope of the Sol, in terms of the payment, isn't the scenario that we are being paid or employees who have been tasked to fill it in. So yeah, while there is potential complexity in the budget, but the budget for filling in the register seems to be outside of the scope of the study that we are doing at the moment, 'cos effectively we are paid in the study.[VJ]"

"The reason is I've been in scenarios where this is your scope boundary, don't go outside this scope, and from a complexity evaluation it is not good enough.[VK]"

The group continued to discuss what "system complexity" means to the participants. "I mean, for me, there is a number of threads to it, and you'll soon learn that threads is a term that I use a lot in systems engineering, its those kind of interoperability threads across the architecture, thats the kind of thing that number one leaps out at me; how many complex little hops have you got there from things tootling around out in the ocean and someone in the ground station pulling it all together and making it into a meaningful set of information. And I think the other one is inside those systems, do I need all of these blocks, do they contribute to the overall picture, and is inserting one of them in in terms of this sales pitch going to get me into trouble. Those two are my big things that leap out on me on it. Its not just necessarily about this is a component with lots of subcomponents or this is a component that does lots of functionality, its sort of system of systems thinking going on here. [VH]"

"So my definition in my head I think disagrees slightly and agrees slightly, so I've always thought of complexity being where you cant directly model all of the inputs, this is the equation, these are all of the inputs that leads to this output, ie you have to dance with a system rather than directly model every single thing. So when [REDACTED] was talking about safety and security and some of those non-functional requirements, they seem to be more like the complex things, whereas when you talk about the threads I define as complicated but not complex..[VJ]", "yeah yeah fair point [VH]", "which is why I just thought it was worth us talking about complicated versus complex, complicated yes we can model and understand and yes it might be difficult painful expensive but complex is what are those emergent properties that we don't expect or are harder to model. [REDACTED] I detect that you have similar thoughts from your introduction or...[VJ]"

"Yeah so there is this thing about measures of complexity isn't there and there is a whole load of mathematics around it which is way too complicated for my small brain. But if you look at it all, it relates to the number of descriptions of things, so picking up what you've just said, if you can describe something adequately with a small number of descriptions it is probably simple, if you can describe something adequately but it requires a large number of descriptions it is probably complicated, so its just a volume thing between simple and complicated, but as you say, when you get into complex and chaotic, if you are using a Cynefin type thing, then those descriptions become less productive, so there will be important things about the system that you cant simply describe, you've got to find other ways of trying to uncover what they are, it just becomes too difficult to take into account all of the different factors and often the answer with complex systems become counter intuitive, so it isn't just a lot of simple descriptions and you just need more and more of them, you need a lot for a complex system but they will still be inadequate and then the chaotic, the complex adaptive side, is that things are changing on you all the time, so as your trying to build up those sets of descriptions that makes up the problem and the solution tractable so you can do something with it, everything has moved on by the time you are getting on top of it and you have to start again. And that for me feels like chaos. And for me, it is often where you have got people and organisations and that sort of thing is where you'll often get most of the complex problems because people change, the personnel might change but peoples perceptions of things change. That was my six pennith worth.[VI]"

"So, yes, similar to what [REDACTED] was just saying there, complex is I think where people are involved, in terms of what they use it for, what they want it for, their viewpoint and then the personal side of an interpretation. So when

people get involved I think it becomes complex and then the second and third order would be if they have particular motives for the systems, we tend to a lot of the time use complex when we are talking about the technology side, but thinking about what [REDACTED] said, that might be more complicated as opposed to complex.[VL]"

"So if we are relatively happy with that difference between complex and complicated, and looking at that yellow but of what our task is to determine the complexity of the MASSS system, and looking at the tool that [facilitator] has provided us which is a register...it sounds like our job is to identify where areas that I might introduce and make this a complex system and therefore it is hard to gauge a bid and therefore our company might not want to bid on it and therefore this is about working out what is going to be complex and are their areas of complexity that make it too complex for us to want to bid on, does that feel right in terms of our task to what other people understood it as at a high level? [VJ]"

The group agreed on this understanding and decided to each work individually on adding as many ideas to the register as they can and then regroup to discuss their contributions. The first entry discussed was "what are the emergent behaviours at the system boundaries" which was further elaborated as "really this was about following that thread through for boats out at sea to someone analysing it, and it is kind of on that border between complex and complicated, so some of the things might be complicated, it is just a [pain] to deal with it but you'll churn through it and it'll be fine, the thing that I was perhaps more interested in around it was I guess those emergent behaviours that you are going to find as you start to consider a System of Systems and how the systems within it start to work with each other. So, you are going to have multiple operators along this chain and they are going to be operating different bits of the system in the ways they see fit, you are going to have data spat out of one interface that might not be compatible with the start end of the next interface, that in itself might cause some issues with system functionality so, buried away in one of the, the green document I think it was, was something about a...rotating C band multifunction passive array radar, well great, cool, I don't know if my ground station can pick up signals from that, what behaviours is that going to cause, is the operator going to have to do some inputting that causes some trouble. Non functional constraints, I think some of the other guys have starting touching on that, particularly security and safety, but there is a whole bunch of things, systems operating together in a scary environment such as the ocean that might not be compatible with each other, so really there is a huge amount of interoperability going on and the complexity that arises from that is well, people have to be able to get that data all the way from point A to point Z, in this case point A to B is easy, point A to B to C to D is complicated and point A to B to Z is complex. I think, in some respects I'm quite an old fashioned systems engineer so my response to this one is start by having more rigorous systems architecture. The one that we have got is clearly put together by someone who thinks about physical, come back a few steps, what are we actually trying to achieve with this bid? And you come back to the green document and someone has written, I referred to it earlier as a sales pitch, pull that apart, turn that into your system architecture and put the two things together and start uncovering all of these emergent parts. I didn't get any further then that, other then identifying sort of who is responsible for it, well you have to put a lead solution architect who sits over it all and there is a lot of overlap of systems engineering as systems engineering and almost systems engineering as engineering management, and if you don't have that lead over the top of it all, same as if you don't have the functional interaction, if you don't have that organisational interaction then you'll start to fall apart quite quickly, and yeah that is about as far as I got on that one. [VH]"

The next complexity issue described was "Architecture does not partition between safety critical and safety impacting functions" which was further elaborated as "Yeah so it was just what we were saying earlier, its often the functionality of the system looked fairly straightforward, as described, I was thinking about how the bits together. So in this case, there is definitely some safety critical functions to do with the navigation of the vessel so it doesn't run into other things and that sort of thing, and then those informational parts of the system where you are just providing

information back, and yet the architecture hadn't made much accounting of that, it just grouped things together based on whether they were sensors or software and processing type stuff. So I didn't really think we had an architecture at that point, as the architecture was probably as we said, a salesman description of it rather than an architects description. [VI]" The facilitator probed a mitigation action which was described as "to develop an architecture". Note, this participant later raised in the interview that they found it challenging to enter more text into the "Complexity Register" due to their connectivity, which perhaps explains why no additional information was recorded for this entry.

The next complexity issue described was "...security requirements not clear - what is secure enough - hard to define and always more you can do" which was further elaborated as "and the potential issue that could have is client acceptance costs could spiral, there could be issues down the line, there is an operational impact from a security thing. And I suppose from this initial system design stage the action plan to mitigate is that thorough understanding of the system architecture and detail it, what will be delivered in the bid, what secure enough actually means. But this point is kind of a sub-point on the first three points around actually you need to go back to the green document and create a proper systems architecture and go from there and resource and define things like security to be good enough. And actually, the point below about safety, I think it should be, safety is similar to security in a way and you need to define what is safe enough and it comes from that architecture. So safety and security, similar in a way. There is one kind of technical one, which is a lot more detailed than we have just spoken about so far, but the only thing that jumped out from the system spec as not being particularly clear how it worked, was the Electro-optic system, and going down a slight rabbit hole, the EO looked like it probably used some machine learning of some kind to train it, it talked about data classifications and it was a real pain to create. There could be a real impact of not knowing why the system has identified an object as being friendly, not friendly, lost. So it could be an issue if there was a fault or something happened or a missile or something ridiculous was launched based on this, and it could be hard to go back and validate the fact. What the action plan is to mitigate that, I don't know, I think it is a problem we are all going to have over the next few years is the amount of ML and AI that is used to make things, so I'd welcome any thoughts on how we are going to deal with that complexity of not knowing why that property has emerged in the future. [VJ]" These points led to further discussions by other participants in terms of safety, security and AI implications which were updated onto the "Complexity Register".

This group did not seem to use the aide memoire until prompted by the facilitator. However, once prompted, the group used the aide memoire to identify additional risks and issues arising from the complexity of the scenario system. The facilitator then prompted "So we have got another 10 minutes left of our original hour here, a suggestion might be to re-look at the Complexity Register guidance, whether the aide memoire or the steps detailed in the user guide, alternatively you are welcome to continue this kind of brainstorming, I am happy with either."

"I'm just looking through the aide memoire and seeing which bits of the Complexity Register we haven't thought about. We've thought a lot about what some of the issues are and what the impacts are and what are some of the potential ways to solve those impacts. There is other bits in terms of looking at the diversity of subsystems and how many connections exist and I think we've spoken about that a lot. Functional and non-functional we have spoken about a lot... [VJ]"

"Encapsulation is probably one thing that we haven't addressed...So can we just at a high level draw what are the high level self contained blocks of the system which we don't need to get into the details of, but these are the ones that we need to be wary of. [VJ]"

"Yeah so there is a big red block at the right hand side of the system architecture that says 'provided by the customer' and more or less we are being directed 'consider that as a black box, don't really poke into it' and it is about knowing enough about the complexity of it to say can I legitimately do that to make everything work with what I need to make work. You've got this red customer box with your own purple box inside it, always fun. I kind of go back to

the green paper on that, because it has got lots of what I would politely call weasel words in it and they say things like 'you are going to have an approach to maximise the flexibility adaptability and durability of the system' and us as systems engineers look at that and go, so what? What do you mean by that? How flexible is flexible? What does flexible mean? They say the system architecture is orientated towards an open architecture. Brilliant. Well, how are you going to measure that? And that is complexity concern number two for me really, as someone has clearly sold this and they have said a load of things, is it even implementable? [VH]"

"Yeah and there is a whole massive question around what is the user community and there is a nice question in the aide memoire going 'how many and how diverse are the users of this system? And to what extent is their intended use of the system stable and well understood?' We have got a very nice idea of they want to drive the boats around and send the pictures back to them and take control, but what conditions? Are they being shot at by the missile falling out of the sky or have they got a nice warm cup of tea and air conditions and how much training have they had, how experienced are they, how much control do we want to let them have. There are lots of questions around that user community that could be interesting. Right back to [REDACTED]'s point at the start that users, people, make things complex. [VJ]" and "Equally who has ever written a requirement for operability that was measurable? [VH]"

"There is this issue of if this system is complicated, and it is to do with situational awareness, which is a cognitive function, and it is to do with how well decision makers in the TAU operating centre and the PANTHEON operating centre, how useful or how accurate their mental picture is of what is going on out there, when all of their sensors are bobbing around using surveillance radars and novel EO sensors, each one of those things will add a degree of abstraction to their understanding of what is going on. And then we have delays coming over the SATCOM over the information getting back. So there needs to be some understanding of how the operators or users of the system understand what the system is telling them. [VI]" "Yeah and I think as a subset of that, going back to the green document, there is nothing about training, or trained, or what level of training will be required, and that fits really nicely under that, we have got no idea of, yeah they might be experts who have been looking at maritime surveillance for over 50 years...[VJ]"

Another participant added system of system considerations; "Would that add into the complexity of if you have spotted something that could be an issue with the architecture, so either safety or the security, the complexity of who do we need to resolve it, will they, how do we engage and influence that person? Is that a second order type of complexity, have we seen that kind of thing before. You know we have got something wrong with the safety side but we need to influence someone further down the chain to fix it but it might not be their biggest priority, I think that adds some sort of underlying complexity of this particular programme of work. [VI]" To which the facilitator probed "So are you almost talking there [REDACTED] about a System of Systems perspective, where a system is provided by PROCULUS GROUP, there might be an issue in the wider System of Systems, but PROCULUS might not be able to address that issue themselves because it is not in their remit or responsibility..." "Yeah yeah it is not in their box, in their little circle of control and it is in someone elses to fix, but that other box it isn't their highest priority or they haven't got the funds to fix it, but it is going to affect what we are trying to deliver. There is a level of complexity of how we influence...who do we need to engage, you know levels of influence vs levels of interest and that kind of thing, adds that kind of undertone for this kind of system. [VL]"

4.2 | How the group approached their evaluation after the scenario inject

After the scenario inject there was a discussion around whether the complexity of the scenario system fundamentally changes or if instead only the time pressure and risk exposure had changes. However, there were disagreements here, with one respondent arguing that the system was fundamentally more complex following the scenario inject.

"So the obvious things are, some of the potentials for complexity have been realised in terms of our worry about people we don't know providing us with information, not knowing about how things are going to work and the pressure on things working, and there is just now more subsystems and more system components to it. And I think because we analysed the system at such a full system of interest, high level of what potential sources of complexity are, I don't think it significantly or fundamentally changes a lot of our earlier analysis from before. Do people agree, disagree? [VJ]"

"Yeah yeah I'm with you. I see the word towed array and I immediately get scared, I mean do you really want a towed array, all the attendant stuff that goes with a towed array? [VH]"

"So for me it changes the baseline at such a key point in the programme [VI]"

"Yeah yeah exactly, at CDR you are getting on for being ready to manufacture and deliver and all of a sudden, oh just add a subsystem it'll be fine [VH]"

"Yeah just add a subsystem and can you make it quicker please [VJ]"

"So maybe the complexity doesn't change, well there is less to confirm where the complexity is, but the level of risk and the opportunity for things to go wrong has increased from the increased subsystems and the time pressure, and oh by the way you probably have to integrate this aircraft in the future, which you probably knew already you would have to implement things into this system as you go along, so hopefully you have designed the architecture in such a way it could deal with through life updates and further integrations but in terms of the sources of complexity, I don't they have changed, or the way you would mitigate them it is just time isn't it. [VJ]"

"Yeah and I think I'd be going back to the requirements as well, saying well your original requirements were to do this and do you think the introduction of this new subsystems meets them, and again I'd be saying towed arrays are good for detecting low frequency things or stuff on the seabed, are you really after those in your original system of interest? Probably not. Plus you have got to process the bloody thing and that is the real pain. Yeah, I agree with you [REDACTED], 100 percent. [VH]"

"So I'm going to disagree slightly, it feels to me the project is very much more complex, we have a lot more thread going on and they are at different levels of maturity, so we've got a whole load of complexity now of change churning into the project. But then, I agree with [REDACTED] in so far as our original prescriptions of why the project was as complicated as it was qualitatively hasn't changed. So, and yeah it looks more complicated, more complex doesn't it. So there is maybe something here that the qualitative measures that we have identified don't give you an accurate measure of the total complexity of the system. If I was the chief engineer of this thing I'd be a lot less sanguine about it at CDR then I would have been at contract award, and I wouldn't have been that happy about it at contract award. [VI]"

"Yeah and I think your point about maturity and different levels of maturity of the subsystems and the various components of it is interesting. Now, trying to combine things at different levels of maturity, or TRL, could introduce complexity into the system and we've just brought in a whole bunch of stuff that could do that and I don't think we'd captured that previously. [VH]"

This entry was added to the "Complexity Register" while another participant added "And the teams doing it will have made a whole set of decisions abed on their understanding at the time added to the fact that there is a whole chunk of stuff they don't know because they aren't getting it from the boat manufacturer. So there has been that problem they have been encountering but they would have made a whole set of decisions based on the information they had at the time, and a lot of decisions will now be wrong yet we don't know that they are wrong because people make those decisions there is all of that covert, tacit stuff that people do when they do design, and you cant write it down because people would never finish it. [VI]"

The discussion continued to identify sources of system complexity and proposed mitigations. Interestingly, this

group seemed to strongly believe that the mitigating actions they had suggested prior to the scenario inject had a significance chance to “head off” some of the issues that materialised in the scenario inject or would have at least placed the organisation in a strong position to be able to deal with what was uncovered from the inject.

“Yeah and I suppose if we think back to our original complexity analysis and [REDACTED]’s point of time and how time is doing to change things, hopefully we would have clocked that the requirements are going to change in time and they are probably going to change what they wanted to do. So that would have put the property on our system design of it would need to be I say reconfigurable but it needs to have a data format that would allow you to introduce other platforms and other sensors into, to maybe a solution we have is ok we can build the system according to the original spec we have and introduce those variants that can do ... in the future. So if we’d done a decent job of designing the original MASSS then it might have been relatively straightforward, so I suppose the actual solution for us, is check how we have designed MASSS to date, has it got the ability to integrate new things, did we do the architecture in a way that is modular and progressive or is it one big system, black box and hope for the best? [VK]”

“I suppose if you have done that architecture with the people side, you know if we had fleets within fleets, you know what type of training do people need, can they cut between different types to support and operate it, again similar thing of can everyone just maintain and operate it or are we adding complexity of additional skills that if you bounce between the fleets you have skills fade or can’t trade down between the two, therefore additional costs of instead of three people to cover it off you need six because three people can’t switch between platforms. [VL]”

Again, near the end of the scheduled time the facilitator prompted the group to reconsider the aide memoire “...another reminder that you might want to consider looking at the complexity register user guide or the aide memoire, alternatively, I’m more than happy to see us continue to flesh these ideas out.” “Shall we have a quick look at the aide memoire and see if there is anything that we haven’t considered? [VJ]” “Yeah [VH]”.

“So there are three parts, related to the technical system, related to the technical system operating in its operational environment and related to the project that realises the technical system. We have spoken about the technical system quite a lot, about the number of different components and the connections, I think. [VJ]”

“I think we’ve touched on the project quite a lot as well [VH]”

“Yeah, in terms of the operational environment that the technical system will be deployed into, I’m not sure much has changed? I suppose there is the environmental factors of how well the system is perceived because of the media incidents with obsolete equipment causing issues, so there is a change in how successful the system is required to be if implemented. If it is a spectacular failure from the start it is going to be more negative than it would have done [VJ]”

4.3 | Interview Findings

Here we report findings from the semi-structured interviews with participants which asked questions around; perceived challenges in evaluating system complexity, how confident participants were in their evaluations, perceived aims and effectiveness of the various complexity assessment tools, feedback on the tool themselves (positive and negative features, opportunities for improvement), and willingness to utilise the various tools in the future.

4.3.1 | Challenges in evaluating system complexity

The first question asked to all participants was; “What challenges did you find when trying to evaluate the complexity of the system?”. Unsurprisingly, given the nature of the study, participants generally reported challenges relating to the scenario such as; not having enough time to assimilate all of the information available, not having access to the required stakeholders to discuss concerns or to gain further information, not having previous working relationships

with other participants, etc.

"I think that understanding the stakeholders. Obviously, in a real situation I would know which company actually bidding for the rest of the systems, but without knowing that, you don't know the culture and I think as much as the technical bit can be quite challenging. The integration between the stakeholders is very key. [VA]"

"Most importantly, were the constraints of time available, as well as the limited amount of information available. And lastly, access to the stakeholders that we know that we wanted to, you know, ask the right questions in the context of [inaudible], you know, so the people in the field really know what the reason for the project, what the drivers are. [VK]"

"So in general it was assembling the information in a way that lent itself well to a structured assessment of the complexity and an I, I guess that's the point. It's always that way, the information doesn't nicely line itself up in a way that patterns become abundantly clear immediately...And then the third bit was about how we did it. Absolutely, doing it collaboratively is the right thing to do. Doing it collaboratively, over Skype, with people that you don't know, and understand their background and where they're coming from. So the fact that [REDACTED] and I had never worked together before made that harder than it would be in a real life situation where you'd be part of a bid team or part of a project team, and you know each other and have a report. [VG]"

"From the basic aspects of the lack of the knowledge and the little bit information we had and the short amount of time to read into it if I'm honest, but then also trying to use the tool so trying to make sure we didn't just use our own opinions and own experiences and actually use the tool as an approach and I did find the aide-memoire bit we did at the end quite useful. [VF]"

"kind of procedural challenges in terms of that storming, norming, forming new group dynamics, who's actually leading, how are we doing this? And it ended up being a bit of a rambling discussion, half using the tool half, not using the tool, rather than being organized and going like this is how we're going to approach it. This person is leading it, let's make sure actually ticking off the various different things as we go through. However, I do think that we talked about a lot of things and got into a lot of detail and maybe that was a good way to approach it to the start. And if we've been doing this for real, we then might have gone right we had a good brainstorming session, let's actually now look at it from a little bit more organized perspective and go through see what we've missed and try and try and be a bit more quantitative about what we think the complexity is rather than just a qualitative discussion...[VJ]"

Challenges in complexity evaluation were also related to the specific tool used by participants. One participant using the "Complexity Profiler" reported challenges around ambiguous definitions of the system complexity factors and their scoring criteria; "...some of the guidance around the different levels wasn't as clear as it could have been, and trying to understand exactly what information fits into which category wasn't always clear...[VE]"

Participants from groups that used the "Complexity Register" and participants from groups that used the "Complexity Profiler" reported challenges around whether system complexity evaluation should be qualitative or quantitative, how system complexity is defined, and establishing the level of abstraction that the system should be considered at for the evaluation.

"Challenges I think came into two phases, so the first phase was what I kind of expected to happen when we started with a group that I haven't worked with before was which is what is complexity and agreeing on that common definition. Hence, I was really keen to ask that question and try and have that conversation right at the start of what is complexity? What do we mean by complexity? What do we think a complex versus a complicated system was? And happily we came to an answer very quickly on it, but there was some slight difference of opinion of what it was or maybe lack of knowledge from it. [VJ]"

"I think the big one really was the same one that always comes up. What's complexity and what is complication. And I think what you find is that there's things that are on the border of complex and complicated. Actually, I think thinking back to that system architecture that was put together [in the scenario], that was a sort of thing that was right on the border of it. You know where one person might say 'well, there's nothing complex about it, it's just a number of things that work together and you just have to do the slog of putting it together' and another person would say, 'Oh no, that's very complicated. I expect to see all sorts of weird stuff happening across that.' So I guess for me that was challenge number one. [VH]"

"The challenges I found was, something that [REDACTED] brought up, what people define as something complex and complicated, so having a common lexicon, a common understanding what you mean by complex and complicated, but once those ground rules, I suppose, were established it made things easier. And then once you established, you know what was complicated, complicated, or complex, again, people's perceptions of what is complicated and complex again, so some people focused a lot more on the technology side of a complex system as opposed to people, processes and technology. So again, it is that balance of what people from their previous life experiences, what is complex? But we spent a lot of time talking about those sort of things. Yeah, so that was the key issue to start with.[VL]"

"In terms of examining the complexity of a system, I think I mentioned this at the time, is it's really hard to come up with a quantitative assessment of complexity, like what is complexity? What is good enough? And I think we discussed a lot about where system boundaries are or the environmental impacts are on the Sol is as you start to include those, the environment and environment affects, complexity is going to increase. And it depends how much that you take into account and how much you include time and money and a simple system, if you go all, everything's going to change in 10 years time, we don't quite know how, could be complex or have complex variables acting on it. [VJ]"

4.3.2 | Self-reported confidence in system complexity evaluation

We asked every participant "How confident did you feel after the exercise that you had effectively evaluated the complexity of the system?". The control group, who were given no complexity assessment tool, reported mixed confidence in their assessment where one was not very confident and the other was reasonably confidence, as both said they needed additional information (e.g., from finance, supply chain, human resources):

"I think there's a lot of information ... in the normal world you will have whole pack of documentation come with it and obviously, without the presence of the supply chain, the HR to making sure the resource you need, ... with finance, you know the sorts of risks and the margins that we're looking at... but also involving with procurement and legal to making sure that you know all those little things. [VA]"

The group who used the “Complexity Profiler” generally reported confidence in their assessment:

“I would say our assessment, where we positioned the system, our assessment was, I feel confident in our assessment, it effectively shows there is a lot of uncertainty around the system. [VE]”

“...I thought we did fairly well...I think pretty much all the areas considered, because it was a large system of systems really and because of that, for me it was complex. So I thought that a three was, overall, if you wanted to summarise it, you know, grouping them all together a three out of four would probably be, you know, straightforward. [VD]”

“I think more or less when we do such exercise in real life, you should feel more comfortable because you will have on the paper what you probably had already in mind so you will probably discover not so much, but at least it will help to put on the paper a global vision that you can share, the important thing there is to make explicit some of the topics. What I have seen in projects many, many times from the very beginning you already know what will be the main [inaudible] that you have to solve, [inaudible], so at least you don't address really the topic and you don't find a solution. So it is important to make explicit and to share that with the team and to find the appropriate answer. So you should more comfortable except if you discover something which is completely new. Or you discover an area that you are not comfortable with. So let's say for example, you discover that you have very high issue of ergonomics or human factors and you have nobody in your team that is able to do that you would probably have to contract to find the right resources outside, so it may bring difficulties but then at least you manage [the challenge]. [VC]”

The intervention groups who used the “Complexity Register” generally reported that they were fairly confident, and would have been even more confident if they were given additional time and access to subject matter experts:

“Not in its totality, however, I think that was largely due to time constraints. I feel had we kept going for another hour, actually, from the level of information and the level of complexity in that system architecture as it were, I think I would come away feeling like we've done a good, if I had I really been a peer reviewer. I think I'd come away feeling like I've done a pretty good job. [VG]”

“Caveat that by the time we had and the information we had, I think we did OK actually. I think it did start to open up avenues of where we would quite happily carry on, you know, [REDACTED] joking said “we could spend an hour on interfaces”, well, quite easily an hour on interfaces. It was only the tip of the iceberg really, but I don't think that's necessary with tool, I think it was just a time frame. I do question how detailed you could go with that level of tool as well. I think it might be good for capturing the high level, and maybe the non-technical aspects of what we're complexity is. But being able to drive it down to the technical “ones and naughts” levels, I think then it would start to struggle. [VF]”

“[And how confident did you feel after the exercise that you had effectively evaluated the complexity of the system?] Oh well, it's a bit like that old joke, isn't it? How do you know an engineer is finished designing something? 'cause he's dead. But yeah, I think, I keep coming back to design reviews, but I think as you move through the maturity of your product, I think you'd be going back and back to that Complexity Register saying I'm confident I've addressed these issues. Or ‘Oh crap, something else has arisen’ or and I'm in the middle of this problem, I've done something...” [VH]

"So I feel we'd identified a number of different sources of potential complexity, evaluated what the complexity level of the MASSS system was, not that confident, but yes, we said there are potential sources of complexity in the system, we had a number of ways that we thought we could reduce the complexity of it, but maybe this is my lack of experience with trying to assess and quantify what complexity is and I don't know where this as an engineering discipline of trying to assigned whether this system is complex or not. So I don't know, and that might be down to my lack of experience, rather than anything else...The exercise we did, I think I would have had a better handle of the complexity of the system if we had more time, as we got into that, ok we've done that brainstorming phase which is effectively what we did, and lets try and think how complex this system is. And I think rather than going, overall this system is this level of complex, thinking about it now, I could have been happier to go right, this system is complex in these areas, and then you can assign something to go right we think the system is particularly complex in its use cases with users and interactions with users but its technical aspects are not complex, I'm not saying that's true for the system we did, but you could break it into different areas of complexity. Which, yeah I suppose we did as we were going through the areas as identifying the areas of complexity [VJ]"

"I believe in given one hour we had for the team and the time we had to prepare then, I know I had probably even lesser time, apologies for the delay, but I think at least, if not almost all the so called surprises where already considered in the complexity register. So I think it was very effective and I believe it can be made more effective by giving some clear, shall we say, objectives as examples, so that we don't limit the scope. So what is going to help in defining the project scope? What is going to help in defining the resources available? What is going to be available in ensuring that the requirements are technically feasible for implementation and you know what is going to be available that the requirements are not conflicting? Have we addressed the safety and security of the system up front...Given the time constraints, I think we did very well. We knew most of the surprises and most of the things that came up would have been already known before, so not being a surprise if the mitigation actions had been incorporated, and possibly some of it may not have been captured in writing, but in terms of what we had discussed as a team, I don't believe we would have had new surprises. [VK]"