A Quantum-Inspired Model for Human-Automation Trust in Air Traffic Control derived from Functional Magnetic Resonance Imaging

Kiranraj Pushparaj, Alvin J. Ayeni, Gregoire Ky, Sameer Alam, V. Vijayaragavan, B. Gulyás, Vu N. Duong
Air Traffic Management Research Institute & Cognitive Neuroimaging Centre
Nanyang Technological University
Singapore

***Abstract* — With a greater proliferation of automation tools in the domain of Air Traffic Management due to exponential growth in air traffic, human factors, and more specifically, trust, becomes a crucial component of Air Traffic Controller (ATCO)-automation teams. An attempt to better represent trust behaviours in ATCOs was made by juxtaposing two philosophies of trust using the principles of superposition and complementarity from quantum mechanics. Neuroimaging evidence of this simultaneous concurrence was demonstrated with use of functional Magnetic Resonance Imaging (fMRI) data. The robustness in this proposed model is higher due to the use of objective data to explain ATCO trusting behaviour under uncertainty. This is an improvement on current models that are context-dependent and based on subjective data.**

Keywords - Air Traffic Management, Human Factors, Neuroergonomics, Quantum Cognition, Trust, fMRI

I. INTRODUCTION

With the expected steady growth of air traffic [22], the Air Traffic Management (ATM) community has been continuously researching new ways of enhancing ATM infrastructure, in order to cope with the ever-growing traffic, while maintaining high levels of safety and capacity. One of the key enablers envisioned by SESAR in its master plan, is the increased automation support of Air Traffic Control Systems [23]. However, with ATM being a safety-critical field, ATCOs have had traditionally low levels of acceptance of these tools. Among the potential concepts mentioned by SESAR in their pipeline to help Air Traffic Controllers (ATCOs) in the future was that of a "Digital Ground Assistant" [20]. The core task of an ATCO is that of ensuring and maintaining separation between all the aircraft under their purview. One of the key elements in providing this service is that of conflict detection and resolution, and this should be the predominant task of this "Digital Ground Assistant." However, regardless of the quality of any automation tool developed, calibration of human-automation trust is crucial to prevent disuse by ATCOs [21].

In one of the earliest models of human-automation trust proposed by [1], trust was defined as the anticipation of a proficient and dependable performance by the automation tool. Conversely, distrust was defined as the anticipation of an ineffectual and unreliable performance by the automation tool. A cursory comparison of the definitions reveals that they are expected to be entirely opposing to each other. This was further corroborated by empirical studies that examined and compared trust to distrust [2], [3]. However, both these studies relied on the self-reports of ATCOs and university students respectively. Questionnaires have inherent limitations, including reflecting the bias of the researcher in the phrasing of questions, ambiguity in the wording of the questions that lead to incorrect interpretations, and difficulty in representing the intricate nature of complex concepts [18]. Furthermore, while subjects may perceive the nature of trust and distrust to be at opposing ends of the same spectrum as indicated in [2] and [3], is that truly the case at a fundamental level?

There is an evident need for the use of more objective data to determine the nature of human-automation trust. While questionnaires have provided a glimpse into the factors that influence trust and human behaviour around technology, the mechanisms and antecedents of trust have not been unearthed thus far. In fact, trust has been one of the least studied human factors in ATM even though it influences ATCO behaviour in a multitude of ways [24]. The answer to elucidating the origins of human behaviour may lie with neuroergonomics, which as the name implies, is the amalgamation of two adjacent fields in research: Neuroscience and Human Factors. This is the next rational step for human factor studies in relation to Air Traffic Management (ATM); to enhance the understanding of fundamental human factors research, by using principles grounded in neuroscience, with trust being no exception [7]. Neuroergonomics, and more specifically, neuroimaging, has the potential to provide a glimpse of the neural basis for trust, and hence, an opportunity to tackle trust-related problems at the very root. This paper aims to do that by presenting a novel quantum-inspired model that characterises trust and distrust on the basis of neuroimaging data.

The chosen framework of analysis for trust was inspired by quantum theory, where superposition of two states occurs under uncertainty. This context is especially appropriate for human-automation trust for ATCOs, since they are required to make important decisions in conjunction with their automation aids in a dynamic environment with the inherent limiting factor of time. Some principles of quantum theory are applied when considering both trust and distrust together as a whole. It is important to point out at this juncture, that the proposed model is not a purely quantum model, but one that derives some quantum elements that are relevant to human-automation trust in ATCOs.

This choice was made because existing models of trust are not robust enough to substantiate this crucial human factor [19]. A quantum-based paradigm may be the most appropriate to change this status quo [11]. Moreover, traditional cognitive models have relied on classical probability theory, which inherently makes these models more deterministic. Human behaviour deviates from these models especially in a complex environment that ATCOs experience daily. Under uncertainty, these traditional models have not been able to explain the resultant behaviour. Quantum cognition models that superimpose a pair of incompatible mental states could potentially account for this uncertainty [12], making them far more reflective of human-automation trust in operational situations. This paper aims to provide a foundation upon which a quantum cognition model can be developed to enrich the understanding of human-automation trust in ATM.

II. Literature Review

A. Trust and Distrust

A lack of human factors experts and ATCOs in the design process have resulted in ergonomic problems with automation tools that were based on suboptimal cognitive models [43]. This can be attributed to the fact that to date, existing cognitive models, have been largely based on empirical data that is subjective and context-dependent [31]. The reluctance to incorporate such models, especially in a field where safety is the paramount concern is understandable. This limitation can be overcome by improving the quality of the data used to conceptualise more versatile cognitive models.

The European Commission’s Flightpath 2050 comprised a vision of aviation in Europe that outlined the importance of lowering the "occurrence and impact of human error, through new designs, training processes, and through technologies that support decision-making” [42]. When considering this in the context of future automation tools that ATCOs utilise, the drivers of automation reliance must be deliberated upon. Trust was found to be a major influence in the way ATCOs depend on their automation aids [41].

There has always been a school of thought that suggested trust and distrust may be entirely different constructs rather than at the extreme ends of a continuum [4], [5]. Anecdotally, it was explained by the phrase “trust but verify”, where the trustor very much authenticates information provided by the trustee, even if there is a high level of trust between them. This was partially supported by neuroimaging studies on human-human trust that followed [6], where distrust was found to have a much larger emotional component as compared to trust. Building trust is more a cognitive process, which explains the close relationship between reliability of a system and trust [25].

A review of neuroimaging studies conducted on human-human trust served as a starting point for this study, in terms of the regions of interest in the brain. The same could not be done for human-automation trust as neuroimaging studies targeting human-automation trust are entirely unprecedented. A comparison study uncovered the common brain regions that were associated with human-human trust in a multitude of trust experiments. These regions were the striatum, thalamus, which were associated with reward, the amygdala, insular cortex and hippocampus, that were associated with uncertainty and risk, the cingulate cortex, which is associated with conflicting mental states, and the frontal cortex, that is associated with mentalising and logic [16]. These brain regions are indicated in Figure 1.

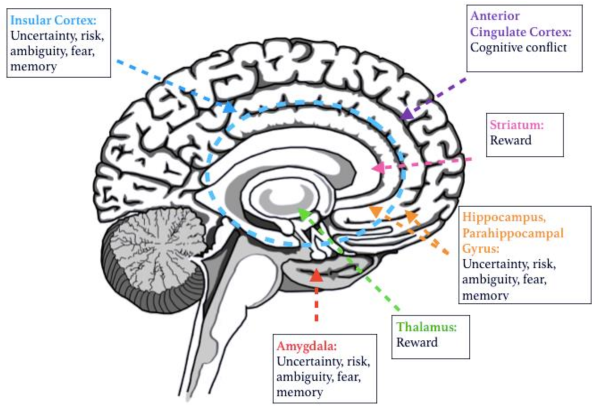

Figure 1: Brain Regions Associated with Trust and Distrust [16]

The true nature of trust and distrust may encompass both these philosophies. Some facets of the relationship between trust may be inverse, while some may be independent of each other. A prudent point to note is that, this prototype of hybrid-trust is discordant with existing literature, especially for human-automation trust. The reason for this could be that the perception of trust that was obtained from studies utilising questionnaires may not have completely captured the complexities and intricacies of trust. However, there is a clear need to demonstrate with objective data, the true relationship between trust and distrust. Yet, it is not clear if such a model can even be represented, since both viewpoints are directly in conflict with one another.

B. Quantum Cognition

In order to interpret such a model, a different approach is required and this paper attempts to create one, inspired from quantum physics. The objective of this study was to make use of the superposition principle taken from quantum physics, that allows for the inherent illustration of the natural cognitive struggles, equivocality or uncertainty that one undergoes during mental processes [26]. Quantum cognition models have been developed to depict irrational decision processes [27], human perception of semantics in language [28], as well as memory, among other cognitive concepts [29]. In fact, there was even a study that examined the trust that a human subject has on photographs, that was explored through the lens of quantum cognition, where the 2 mental states of trust for the authenticity of a picture, and trust for the subject of the picture were superimposed [44]. Likewise, instead of restricting the analysis of trust and distrust to one interpretation, this approach suggests that both incompatible philosophies can coexist and hence be superimposed.

The superposition principle is one of the core ideas specific to quantum mechanics [13]. Indeed, when it is impossible to precisely observe the state in which an object is without measurement, it is considered to be in all possible states at the same time. The most famous example is Schrodinger's cat experiment, in which the cat's state is considered both dead and alive at the same time since its precise state cannot be confirmed. The same principle applies to quantum computing, in which the binary deterministic approach is replaced by quantum superposition. Instead of manipulating bits which will necessary be either 0 or 1, this approach manipulates qubits, which can be both at the same time.

However, the combinative model introduced in this paper requires more than two possibilities. Consequently, a 2-qubit basis has been chosen in order to superimpose both approaches. This basis contains four vectors which, with regards to trust and distrust modelling, can be interpreted as follows:

- $|00\rangle$, as pure distrust
- $|11\rangle$, as pure trust
- $|01\rangle$, as reciprocal trust
- $|10\rangle$, as reciprocal distrust,

where pure trust and pure distrust are independent of each other, while reciprocal trust and reciprocal distrust are entirely dependent on each other. Subsequently, the quantum state, $|\psi\rangle$, can be written as such:

$$|\psi\rangle = \alpha|00\rangle + \beta|11\rangle + \gamma|01\rangle - \gamma|10\rangle \qquad (1)$$

With α, β, γ complex numbers corresponding to the probability amplitudes. In the case of reciprocal trust, positive trust is synonymous with negative distrust, and vice versa. Due to this inverse relationship, the coefficients for $|01\rangle$ and $|10\rangle$ must be necessarily opposed, which is indeed a slight divergence from a purely quantum model. Furthermore, according to the Born rule, the modulus of each of those coefficients are the respective probabilities of each outcome. Consequently, those coefficients are related by the following relationship:

$$|\alpha|^2 + |\beta|^2 + 2|\gamma|^2 = 1 \qquad (2)$$

One of the founding principles of quantum theory, and the foundation of the superposition principle is that of complementarity, which suggests that any quantum system requires two or more mutually exclusive states [30]. The origin of this principle can be traced back to the acclaimed physicist, Niels Bohr, who emphasised that this was the distinguishing feature between classical probability theory and quantum theory [10]. These mutually exclusive states can then be superimposed to reflect the principle of superposition.

It is this very same principle that can account for the simultaneity of two incompatible cognitive trust models. Both sets of contradictory literature on human-automation trust may indeed be right under different contexts. The specificity of these environments is very much unclear at this moment. Furthermore, evidence of this concurrence may prove difficult to obtain experimentally through observed behaviour, if it is even possible. However, underlying mechanisms in the brain will be reflected under suitable conditions if neuroimaging techniques are used whilst subjects perform cognitive tasks that induce both pure trust and reciprocal trust.

Previously, models on cognition have largely relied upon 2 types of framework:

1. "Heuristic" framework, and a
2. "Rational" framework [14].

Quantum cognitive models borrow elements from both, by acknowledging the logical behaviour of humans, but also accounting for unpredictable differences that induce natural constraints in the process of decision making. One of the most relevant examples of these natural constraints for ATCOs is the time pressure, that can change their trusting behaviour. Decreased time pressure will encourage a more rational approach, while increased time pressure will likely prompt a quicker, but less accurate heuristic approach. A quantum model of trust has the potential to represent both types of behaviours, making it more robust.

C. Neuroimaging

Neuroimaging continues to be one of the most popular tools used in neuroergonomics [7] due to the large variety of accurate and precise neuroimaging techniques available, each with its own advantages and disadvantages. The versatile properties of neuroimaging have enabled its use in various types of studies. Tools such as functional Near Infrared Spectroscopy (fNIRS) that uses light wavelengths to compare the density between oxygenated haemoglobin vs de-oxygenated haemoglobin and Electroencephalogram (EEG) which measures the electromagnetic signals of the brain during specific tasks, are able to provide excellent temporal resolution for time-sensitive tasks. Tools such as the Positron Emission Tomography (PET) and Transcranial Doppler Sonography (TCDS) on the other hand, detect the changes of blood flow diffusion into the various anatomical regions of the brain and provide excellent spatial resolution [15].

Neuroimaging techniques have in fact been used when attempting to understand human factors such as cognitive workload, attention, and vigilance in the context of ATC [31]. The success of those studies provides a foundation for further endeavours into neuroergonomic studies in ATM.

Even though the EEG may be the most practical neuroimaging technique to carry out human-in-the-loop experiments in an environment reflective of conditions experienced by an ATCO [31], for the purposes of this study, the identification of precise regions of interest in the brain is crucial in demonstrating the superposition of pure trust and reciprocal trust. As such, spatial resolution was given the

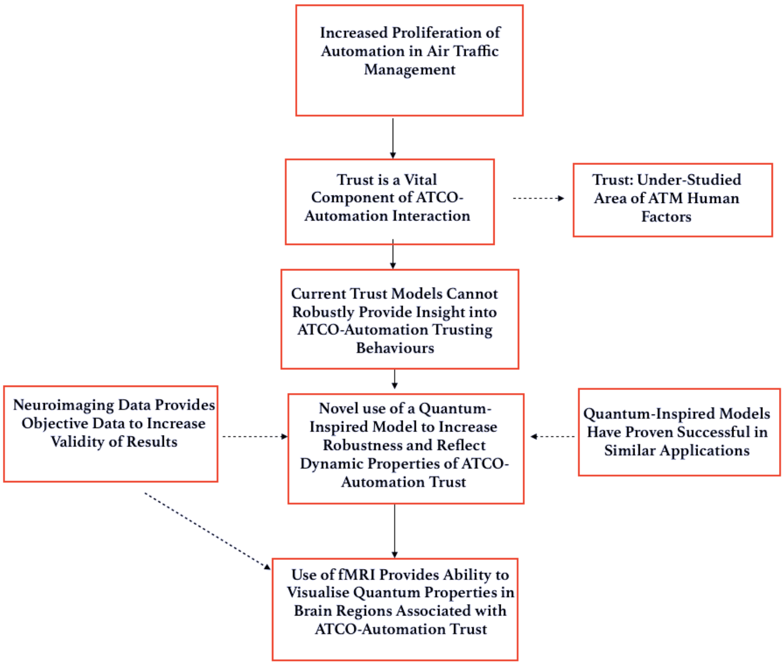

Figure 2: Formulation of Approach

priority and the technique that provided the best spatial resolution, while still being compatible with the experimental design was that of functional Magnetic Resonance Imaging (fMRI).

The fMRI was proposed given its inherent ability to highlight the brain regions most activated for specific cognitive tasks [8]. This is accomplished through the comparison of Blood Oxygenation Level Dependent Contrasts (BOLD) signals before and during the completion of tasks. Indeed, as our brains focus on specific cognitive tasks, the brain regions implicated in the completion of these tasks subsequently require more neurons to be activated, which in turn, require more energy in the form of blood/oxygen. The increase of oxygen in these associated brain regions are then compared between when the task is being performed versus when that specific task is not being performed, allowing a glimpse into the brain regions that are most activated when ATCOs are interacting with their automated tools, and visualising if any of these regions are associated with trust and/or distrust [8].

III. Research Problem

Trust is a crucial human factor that governs the way ATCOs use their automation tools, especially when it comes to state-of-the-art equipment. Its importance magnifies with the increased quality and quantity of automation tools being developed to be deployed in ATM. However, present trust models have not been able to fully represent ATCO trusting behaviours, especially under uncertainty. A more robust quantum inspired cognitive model is required to address this particularly in ATM, since ATCO tasks involve accounting for numerous dependent and independent variables in a stochastic environment. Objective fMRI data, that offers a glimpse into the fundamental nature of ATCO trust is used to test the validity of this model, as shown in Figure 2.

IV. Methodology

The study consisted of 5 participants (2 former professional ATCOs and 3 student ATCOs) one female student, and 4 males in total (2 student ATCOs, and 2 former ATCOs). One of the male students' data was excluded from the study despite being scanned due to the subject's difficulty in understanding the task, thus resulting in anomalous results. Final experimental design culminated in a final count of 4 participants (2 former professional ATCOs and 2 student ATCOs). All participants were right-handed and mean age was 39.75 years old with a range of 41 years.

The simulation was created using the ATS-Cap software which was recorded and coded using E-Prime which allowed the ATS-Cap scenarios to be followed up by an on-screen prompt (Figure 4) in which subjects had to decide whether to accept or reject the automation's provided advisory. The ATS-Cap software provided a set of 5 conflict detection scenarios between a pair of aircraft of varying degrees of difficulty. The simulation of 5 scenarios were based on an arbitrary airspace with sparse traffic density as shown on Figure 3, making it harder to predict if the aircraft identified would truly experience a loss of separation. However, all the advisories provided were accurate.

As previously mentioned, fMRI provides the capabilities to provide in real-time, evidence of the most activated brain regions in respect to specific tasks being performed, providing insight into that region's functionalities for the tasks being completed. Indeed, by using fMRI whilst professional ATCOs and student ATCOs were resolving potential aircraft conflicts

ISSN 0770-1268

via the ATS-Cap software, it is evident if the areas traditionally associated with trust and/or distrust are activated, or combination of the two, which would indicate quantum properties.

Figure 3: Ambiguous Traffic Condition in Scenario

In order to investigate how much trust and/or distrust ATCOs have for their autonomous tools, a novel research paradigm utilising both fMRI and the ATS-Cap software was developed, which was able to simulate realistic aircraft flight plans, in addition to including the waypoints subjects must follow, as well as airports along the route. ATCOs were confronted with potential aircraft conflicts and decided whether to accept or reject the advice given by their automated autonomous tools in the form of a prompt advisory. The videos depicting the potential conflict scenario was projected for approximately 2-3 minutes. After being presented, subjects were asked if they accepted or rejected the automated autonomous tool's advisory, as shown in Figure 4.

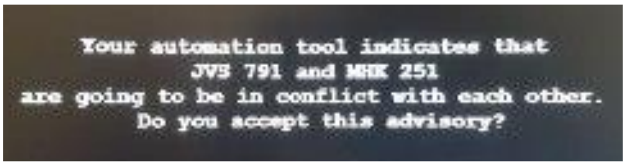

Figure 4: Prompt Advisory Example

The experiment took place at the Cognitive Neuroimaging Centre at Nanyang Technological University (NTUin the Siemens 3-Tesla MAGNETOM Prisma MRI scanner (Siemens Medical Solutions, Erlangen, Germany). The protocol consists of the following sequences a) Localizer scan, b) 3D T1 MPRAGE for co-registration, with the acquisition parameters: TR/TE= 2400/2.28ms, Slice thickness (ST)= 1*mm*, Number of slices= 208, FOV= 256×256*mm*, Matrix size = 256×256. c) fMRI sequence with a measurement of 409 for the ATC simulation. The rest of the acquisition parameters are TR/TE= 2000/28*ms*, Slice thickness (ST)= 2.5*mm*, Number of slices= 64, FOV= 250×250*mm*, iPAT= 2, Matrix size = 100×100. The experimental time for the simulation was 13:51. The flight simulation was programmed in E-Prime 2.0 software and was synchronized with the fMRI scans. Whilst in the fMRI, subjects viewed the simulation through a mirror that displayed the stimuli from a screen and head motion was restricted using padded clamps.

In order to visualise which brain areas were most activated when completing the task, ATCOs made their decision to either accept or reject the automation's suggestions via a binary remote controller that they had access to whilst in the fMRI device. Since a specific ATC automation device was not

Figure 5: MRI Scanner With Subject Conducting Task

available to be used for this experiment, the custom prompt was not provided by a corporeal automation tool, but the participants were unaware of this to preserve the integrity of this study.

V. Results and Discussion

Data analysis was conducted using the Connectivity Toolbox (CONN) [32] and Statistical Parametric Mapping 12 (SPM12) [33]. fMRI pre-processing was performed using the default processing pipeline provided by CONN, which included correction for head motion artefacts, temporal and spatial normalisation in Montreal Neurological Space (MNI) and brain smoothing using a Gaussian kernel with an isotropic kernel of 8 mm.

To account for random artefacts associated with spiking and motion, which could lead to false correlations, CONN's artefact detection feature was used which identifies principal components associated with white matter and cerebrospinal fluid (CSF) for each subject. White matter, CSF, and realignment parameters were entered into CONN as first-level analysis confounds, which were then band-pass filtered to [0.008 to 0.09Hz], thus normalising the data.

Functional connectivity between any two brain regions can be attributed to connectivity within a network, or connectivity between two separate networks [45]. Thus, seed-based connectivity and by extension, a seed-based correlation analysis (SCA) is the customary method of exploring functional connectivity within the brain. Dependent on the time series of the seed voxel or the primary brain Region of Interest (ROI), seed-based connectivity is calculated as the degree of correlation between the time series for all other voxels in the brain. As such, to test the brain activation hypotheses, seed-based functional connectivity analysis was performed using the CONN toolbox utilising the standard weighted general linear model (GLM), which can be thought of as an extension of the linear regression statistical technique.

More specifically, The GLM technique used in fMRI experiments consists of the same conceptual equation ($Y = X\beta + \varepsilon$) as a simple linear regression example ($Y = a + bX + \varepsilon$). The GLM states that Y(which in the case of fMRI, represents the measured fMRI signal from a single voxel as a function of time) can be expressed as the sum of one or more experimental design variables (X), each multiplied by a weighting factor (β), plus random error (ε)[40]. This technique was applied across all 4 subjects to ensure that only the brain regions that showed statistically significant activation and an effect size of larger than (d= 0.6) across all 4 subjects were reported in the analysis.

The basis of brain activation comparison were regions of the brain associated with either pure trust, pure distrust, reciprocal trust, and distrust, which were then compared based on a ROI to ROI analysis of the regions most associated with these areas according to the literature [16]. Within these regions, only those regions in which there was significant activation were included, based on a False Discovery Rate (p-FDR) value of less than 0.05. Though the Bonferroni method would be better suited to minimise the likelihood of Type 1 errors, its usage would be too conservative leading to many missed findings, especially considering the low sample size of this study.

The first part of the brain that was used as a seed to compare brain activity was the Anterior Cingulate Cortex, as it had the highest association with cognitive conflict [16], [34], [35]. The brain activity for the duration of the scenarios was analysed in addition to the brain activity during the prompts, which were the conflict detection advisories, to capture the cognitive conflict that ATCOs experience under operational conditions.

The results are shown in Table 1. Subsequently, since the insular cortex network showed frequent significant activations (during 4 out of 5 scenarios and 3 out of 5 prompts), the insular cortex, which is highly associated with uncertainty and risk [36], [37], [38], [39], was then used as the seed for a similar analysis with the same raw data to reflect the functionally correlated areas and the results are shown in Table 2.

The different regions of the brain that showed a significant amount of activation during this study are the insular cortex network, amygdala, putamen, nucleus accumbens, Anterior Cingulate Cortex (ACC), and Posterior Cingulate Cortex (PCC). The type of trust that they are likely to be associated with, on an inferential basis are as follows:

- Insular Cortex: Associated with risk and uncertainty [16] and thus, likely to indicate Pure Distrust.
- Amygdala: Associated with negative emotional salience [16] and thus, likely to indicate Pure Distrust.
- Putamen: Associated with reward [16] and thus, likely to indicate Pure Trust.
- Nucleus accumbens: Associated with reward [16] and thus, likely to indicate Pure Trust.
- ACC and PCC: Associated with cognitive conflict [16] and thus, likely to indicate reciprocal trust and distrust.

Additionally, based on the literature, the thalamus and the hippocampus have been found to influence trust.

- Thalamus: Associated with reward [16] and thus, likely to indicate Pure Trust.
- Hippocampus: Associated with learning and memory [16] and thus, likely to indicate Reciprocal Trust and Distrust.

However, there was no significant activation of these two areas in this study. Nevertheless, the quantum effect could still be observed in subsets of the data. For example, during Scenario 3, when using the Insular Cortex as the seed, there was statistically significant activation of the insular cortex network, the ACC, and the putamen, which signifies simultaneous activation of pure distrust, reciprocal trust, and pure trust respectively. This activation is illustrated on Figure 6, where the magnitude of activation, that can be represented by α, β, and γ from (1) is also evident. Other instances of this superposition can be seen during Scenario 1 as well, when using the Insular Cortex as the seed. When using the ACC as the seed, quantum superposition can be observed during Prompt 2 and Prompt 5, as well as during Scenario 5. This suggests that the quantum model proposed is indeed valid.

TABLE I. AREAS WITH SIGNIFICANT EFFECTS

Seed: Anterior Cingulate Cortex Network			
Scenario Number	***Areas where p-FDR < 0.05***	***Prompt Number***	***Areas where p-FDR < 0.05***
Scenario 1	Insular Cortex Network	Prompt 1	Insular Cortex Network
Scenario 2	-	Prompt 2	Insular Cortex Network, Amygdala, Putamen
Scenario 3	Insular Cortex Network	Prompt 3	-
Scenario 4	Insular Cortex Network	Prompt 4	-
Scenario 5	Insular Cortex Network, Putamen	Prompt 5	Insular Cortex Network, Nucleus Accumbens

TABLE II. AREAS WITH SIGNIFICANT EFFECTS

Seed: Insular Cortex (Atlas)			
Scenario Number	***Areas where p-FDR < 0.05***	***Prompt Number***	***Areas where p-FDR < 0.05***
Scenario 1	Anterior Cingulate Cortex Network, Insular Cortex Network	Prompt 1	-
Scenario 2	-	Prompt 2	Insular Cortex Network
Scenario 3	Anterior Cingulate Cortex, Insular Cortex Network, Putamen	Prompt 3	-
Scenario 4	-	Prompt 4	-
Scenario 5	Insular Cortex Network, Nucleus Accumbens	Prompt 5	Posterior Cingulate Cortex Network

The concurrence of the different trust states under uncertainty and ambiguity will influence use, misuse and disuse [21]. The weight of the coefficients of each state will determine both the conformance and the type of use by ATCOs. For example, if pure trust is the dominant element, misuse is more likely to occur as the ATCO becomes more comfortable with their automation tool and may even border on over-confidence, likely due to over-trust. However, if all the elements are equally weighted, the cognitive conflict will be the main driver behind the conformance decision. This is yet to be demonstrated empirically, but this model is able to account for this type of behaviour.

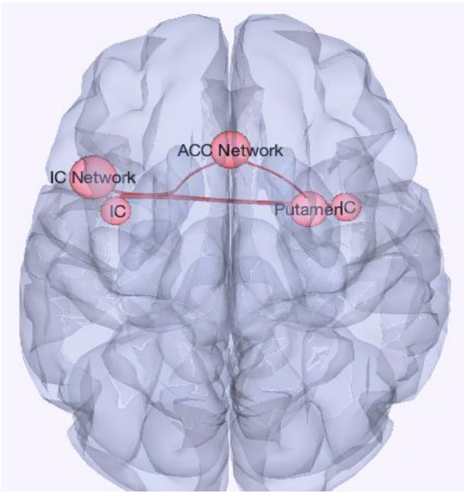

Figure 6: Incidence of quantum superposition

It must be noted that regions of the brain associated with pure distrust are dominant for both seeds initially. This could potentially suggest that participants had a predisposition to distrust automation tools under uncertainty and ambiguity, when automation characteristics are unclear. However, this changed as the simulation progressed, with the later scenarios and questions reflecting greater elements of trust, both pure and reciprocal. Furthermore, the amygdala showed significant activation only once across the simulation, even though it is very strongly associated with distrust [6]. This could be attributed to participants being cognizant that the experiment is based on a simulation thus, no real consequences were attached to making an incorrect decision.

A two-sample t-test was also carried out to compare the brain activity of the ATCOs and student controllers, but this failed to yield any significant results. Though, this should be taken with consideration to the small sample size (two subjects per group).

Despite not observing quantum effects throughout the simulation, the data demonstrated aspects of quantum characteristics in segments of the simulation, which has not been explored in previous literature. Moreover, the data obtained here is far more objective than self-reports or questionnaire-based models that are more subjective in nature, and less robust.

VI. Limitations and Future Work

It must be emphasised that the proposed model is not a purely quantum model. For a system to be considered as a quantum model, it must exhibit the superposition principle, complementarity as well as the uncertainty principle [30]. Even though complementarity and the superposition principle have been established in this model, it is unclear at this point whether the uncertainty principle is applicable in this instance. Further research is required to confirm the validity of the uncertainty principle and authenticate this as a pure quantum cognitive model.

Furthermore, there were only 4 participants in this study, as it was designed to be a proof of concept, rather than an extensive exploration of trust. More participants will be required to reach statistical significance, and more functionally related brain regions. A more diverse group of participants will also enhance the robustness of the experimental design and results. That would only augment the understanding of the underlying antecedents of human-automation trust in ATCOs, which is still a largely unexplored area in research.

It must be noted that all inferences of brain regions to pure trust, pure distrust, reciprocal trust and distrust were based on existing literature of human-human trust. It is uncertain whether the same brain regions are associated with the antecedents of human-automation trust. Moreover, the link between ROI and type of trust was based on deduction and inference. Further neuroimaging studies are required to compare and contrast human-human trust and human-automation trust, and to validate the precise constituents of this quantum-inspired model.

VII. Conclusion

Traditional models of trust have not been robust enough to reliably infer ATCO behaviour, when utilising automation tools under uncertainty. An innovative model, that leverages on quantum properties, that can better encompass the spectrum of ATCO trusting behaviour with automation is proposed. fMRI data was used to demonstrate the dyadic nature of trust, where pure trust and distrust coexist with reciprocal trust and distrust.

Acknowledgment

This work is partially supported by NTU-CAAS Research Grant M4062429.052 by Air Traffic Management Research Institute, School of MAE, NTU, Singapore. Ethics approval for this research was granted by the Nanyang Technological University (NTU) institutional review board (IRB) (NTU IRBIRB-2018-12-002). A participant information sheet was provided and signed consent form was collected. At no point were participants asked to reveal names or other identifiable information, and only anonymised data was collected.

We would also like to thank Dr. Vimalan Vijayragavan and the entire neuroimaging team from the Cognitive Neuroimaging Centre (CoNiC) for helping us carry out the study in a timely manner.

REFERENCES

[1] Muir BM. "Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems," Ergonomics. 1994 Nov 1;37(11):1905-22.

[2] Goillau, P., Kelly, C., Boardman, M. and Jeannot, E., "Guidelines for trust in future ATM systems – measures," 2003

[3] Jian JY, Bisantz AM, Drury CG. "Foundations for an empirically determined scale of trust in automated systems," International Journal of Cognitive Ergonomics. 2000 Mar 1;4(1):53-71.

[4] Lewicki RJ, McAllister DJ, Bies RJ. "Trust and distrust: New relationships and realities," Academy of management Review. 1998 Jul 1;23(3):438-58.

[5] McKnight DH, Choudhury V. "Distrust and trust in B2C e-commerce: Do they differ?" In Proceedings of the 8th international conference on Electronic commerce: The new e-commerce: innovations for conquering current barriers, obstacles and limitations to conducting successful business on the internet 2006 Aug 13 (pp. 482-491). ACM.

[6] Dimoka A. "What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study," Mis Quarterly. 2010 Jun 1:373-96.

[7] Parasuraman R, Wilson GF. "Putting the brain to work: Neuroergonomics past, present, and future," Human factors. 2008 Jun;50(3):468-74.

[8] Smith, S. M. (2004). "Overview of fMRI analysis," The British Journal of Radiology77(suppl_2): S167-S175.

[9] Nakahara M, Ohmi T. "Quantum computing: from linear algebra to physical realizations," CRC press; 2008 Mar 11.

[10] Bruza PD, Wang Z, Busemeyer JR. "Quantum cognition: a new theoretical approach to psychology," Trends in cognitive sciences. 2015 Jul 1;19(7):383-93.

[11] Arfi B. "Resolving the trust predicament: A quantum game-theoretic approach," Theory and Decision. 2005 Sep 1;59(2):127-74.

[12] Wang Z, Busemeyer JR, Atmanspacher H, Pothos EM. "The potential of using quantum theory to build models of cognition," Topics in Cognitive Science. 2013 Oct;5(4):672-88.

[13] Dirac PA. "The principles of quantum mechanics," Oxford university press; 1981.

[14] Bruza PD, Hoenkamp EC. "Reinforcing trust in autonomous systems: A quantum cognitive approach," In Foundations of Trusted Autonomy 2018 (pp. 215-224). Springer, Cham.

[15] Parasuraman R. Neuroergonomics: "Brain, cognition, and performance at work," Current directions in psychological science. 2011 Jun;20(3):181-6.

[16] Riedl R, Javor A. "The biology of trust: Integrating evidence from genetics, endocrinology, and functional brain imaging," Journal of Neuroscience, Psychology, and Economics. 2012 May;5(2):63.

[17] Madsen M, Gregor S. "Measuring human-computer trust," In 11th Australasian conference on information systems 2000 Dec 6 (Vol. 53, pp. 6-8).

[18] Demetriou C, Ozer BU, Essau CA. "Self-report questionnaires," The encyclopedia of clinical psychology. 2014 Dec 29:1-6.

[19] Chien SY, Semnani-Azad Z, Lewis M, Sycara K. "Towards the development of an inter-cultural scale to measure trust in automation," In International conference on cross-cultural design 2014 Jun 22 (pp. 35-46). Springer, Cham.

[20] SESAR Joint Undertaking. "SESAR Solutions Catalogue 2019," Third Edition 2019.

[21] Parasuraman R, Riley V. "Humans and automation: Use, misuse, disuse, abuse," Human factors. 1997 Jun;39(2):230-53.

[22] EUROCONTROL. "Challenges of Growth 2018," June 2018.

[23] "European ATM Master Plan Edition 2015.", 2015

[24] Langan-Fox J, Sankey MJ, Canty JM. "Human factors measurement for future air traffic control systems," Human factors. 2009 Oct;51(5):595-63

[25] Dzindolet MT, Peterson SA, Pomranky RA, Pierce LG, Beck HP. "The role of trust in automation reliance," International journal of human-computer studies. 2003 Jun 1;58(6):697-718.

[26] Wang Z, Busemeyer JR, Atmanspacher H, Pothos EM. "The potential of using quantum theory to build models of cognition," Topics in Cognitive Science. 2013 Oct;5(4):672-88.

[27] Pothos EM, Busemeyer JR. "A quantum probability explanation for violations of 'rational decision theory," Proceedings of the Royal Society B: Biological Sciences. 2009 Mar 25;276(1665):2171-8.

[28] Bruza P, Kitto K, Nelson D, McEvoy C. "Is there something quantum-like about the human mental lexicon?" Journal of Mathematical Psychology. 2009 Oct 1;53(5):362-77.

[29] Brainerd CJ, Wang Z, Reyna VF. "Superposition of episodic memories: Overdistribution and quantum models," Topics in Cognitive Science. 2013 Oct;5(4):773-99.

[30] Lahti PJ. "Uncertainty and complementarity in axiomatic quantum mechanics," International Journal of Theoretical Physics. 1980 Nov 1;19(11):789-842.

[31] Aricò P, Borghini G, Di Flumeri G, Bonelli S, Golfetti A, Graziani I, Pozzi S, Imbert JP, Granger G, Benhacene R, Schäfer D. "Human factors and neurophysiological metrics in air traffic control: a critical review," IEEE reviews in biomedical engineering. 2017 Apr 12;10:250-63.

[32] Whitfield-Gabrieli S, Nieto-Castanon A. "Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks," Brain connectivity. 2012 Jun 1;2(3):125-41.

[33] Ashburner J, Barnes G, Chen C, Daunizeau J, Flandin G, Friston K, Kiebel S, Kilner J, Litvak V, Moran R, Penny W. "SPM12 manual," Welcome Trust Centre for Neuroimaging, London, UK. 2014 Jun 26.

[34] Botvinick MM, Cohen JD, Carter CS. "Conflict monitoring and anterior cingulate cortex: an update," Trends in cognitive sciences. 2004 Dec 1;8(12):539-46.

[35] Botvinick M, Nystrom LE, Fissell K, Carter CS, Cohen JD. "Conflict monitoring versus selection-for-action in anterior cingulate cortex," Nature. 1999 Nov;402(6758):179.

[36] Critchley HD, Mathias CJ, Dolan RJ. "Neural activity in the human brain relating to uncertainty and arousal during anticipation," Neuron. 2001 Feb 1;29(2):537-45.

[37] Huettel SA, Stowe CJ, Gordon EM, Warner BT, Platt ML. "Neural signatures of economic preferences for risk and ambiguity," Neuron. 2006 Mar 2;49(5):765-75.

[38] Krain AL, Wilson AM, Arbuckle R, Castellanos FX, Milham MP. "Distinct neural mechanisms of risk and ambiguity: a meta-analysis of decision-making," Neuroimage. 2006 Aug 1;32(1):477-84.

[39] Kuhnen CM, Knutson B. "The neural basis of financial risk taking," Neuron. 2005 Sep 1;47(5):763-70.

[40] Monti MM. "Statistical analysis of fMRI time-series: a critical review of the GLM approach," Frontiers in human neuroscience. 2011 Mar 18;5:28.

[41] Higham TM, Vu KP, Miles J, Strybel TZ, Battiste V. "Training air traffic controller trust in automation within a NextGen Environment," In International Conference on Human Interface and the Management of Information 2013 Jul 21 (pp. 76-84). Springer, Berlin, Heidelberg.

[42] European Commission. "Flightpath 2050: Europe's Vision for Aviation," Report of the High Level Group on Aviation Research Publications Office of the European Union Luxembourg, 2011.

[43] Wise JA, Hopkin VD, Garland DJ. "Handbook of aviation human factors," CRC Press; 2016 Apr 19.

[44] Bruza PD, Hoenkamp EC. "Reinforcing trust in autonomous systems: A quantum cognitive approach," In Foundations of Trusted Autonomy 2018 (pp. 215-224). Springer, Cham.

[45] Joel SE, Caffo BS, Van Zijl PC, Pekar JJ. "On the relationship between seed-based and ICA-based measures of functional connectivity," Magnetic Resonance in Medicine. 2011 Sep;66(3):644-57.