EPiC Series in Computing
Volume 102, 2024, Pages 111–125
Proceedings of Conference on Cognitive and Computational Aspects of Situation Management 2023

A Quantum Decision Approach for Human-AI Decision-Making

Scott A. Humr[12], Mustafa Canan[12], Mustafa Demir[3]
[1]Department of Information Sciences
[2]Naval Postgraduate School, Monterey, CA USA
[3]Ira A. Fulton Schools of Engineering Arizona State University, Mesa, AZ USA
scott.humr@nps.edu, mustafa.canan@nps.edu, mdemir@asu.edu

Abstract

Artificial Intelligence is set to encompass additional decision space that has traditionally been the purview of humans. However, this decision space remains contested. Incongruencies between artificial intelligence and human rationalization processes introduce uncertainties in human decision-making, which require new conceptualizations that capture these distinct types of uncertainties. Hence, developing new ways to model human and artificial intelligence interactions are necessary to account for such uncertainties and improve situation awareness and decision-making. In this paper, we outline current conceptualizations of human and machine rationalities. Next, we offer the concept of rational prediction deviations (via quantum probability theory) for capturing uncertainty in situational awareness. Lastly, we propose a human-in-the-loop construct to explicate how applications of quantum probability theory in decision science can ameliorate situation awareness models by providing a novel way to capture distinct dynamics of decision making.

Keywords—Artificial Intelligence, Decision-making, Quantum Probability, Human-in-the-loop

1 Introduction

Information is a critical input to situational awareness and decision-making [1]. However, contemporary decision-makers are faced with two significant issues that encumber information processing for attaining/improving/sharing situational awareness: (1) making sense of the significant amount of data and (2) servicing additional informational requirements that are difficult to obtain due to a lack or shortage of resources. Recently, adopting artificial intelligence (AI)-based support tools is the solution to overcome these challenges. Augmented by several technologies such as large storage and processing power found in cloud computing (e.g., large quantities of graphics processing units), AI is promising to ameliorate information processing and situational awareness [2]. While state-of-the-art AI achievements show a great deal of promise, they do raise concerns for how to engineer AI into the human decision-making processes and augment situation management.

The structure of the paper is as follows. First, we review human and machine rationalities followed by common conceptions of human decision-making supported by automated decision aids. Next, we discuss quantum models of cognition and how they can potentially improve outcomes by better modeling uncertainty and order effects within human-AI decision-making constructs and conclude by outlining future research.

2 AI Aspirations: Decision Space

AI is becoming more integrated into decision-making processes. In return, humans are increasingly disconnected spatially and temporally from what goes into algorithmic decision-making [3], which introduces new challenges to situation management, e.g., comprehending AI logics [4]. While automated decision-making may reduce human workload since the essence of decision-making entails the choice amongst differing options, automation may also inadvertently increase uncertainty [5], which can be characterized as strategic uncertainty. Hertwig et al. [6] define strategic uncertainty as "uncertainty about the actions of others in interactive situations" (p. 369). In managing interactive situations, actions of others may be difficult to understand or may unintentionally constrain decision options. For instance, if humans sense no degrees of freedom in choosing amongst alternatives, then choice may be an illusion. In such cases, humans may begin to sense a degree of arbitrariness or randomness in decision-making [7].

2.1 Human and Machine Rationalities

The logic of machines and humans have been characterized in separate ways. Machines' logic (e.g., AI) are considered a type of formal rationalization that follows strict procedures by avoiding arbitrariness or situational idiosyncrasies [8]. Such logic is deterministic since the same inputs will always produce the same outputs, given models are not updated. This gives machine logics consistency, reliability, and stability with little to no variance in outcomes. While such regularity is needed in many different contexts, it is not always optimal in situations requiring the application of nuanced judgment and ordering of strategic-level goals.

Conversely, human rationality is different. For instance, it is well-established that human rationality is limited and often exhibits biases [9]–[11]. Human rationality has also been characterized as normative or substantive [8]. Substantive rationality entails reasonableness of evidence [12]. The most distinctive aspect of substantive rationality is the notion that it contains the possibilities of counterfactual reasoning or what the world 'could be' [13]. This aspect of counterfactual rationality is something AI cannot perform [14], which makes human rationality the *sine qua non* for higher decision quality. Undoubtedly, AI can augment human rationality by allowing cognitive offload of certain tasks or by enhancing our computational abilities [15]. Although humans and AI rationalities can complement each other [16], the challenge is how to manage situations in which these two types of rationalities become incongruent.

With AI on the ascendency in many areas, situation management will face challenges for how to integrate humans and AI systems reciprocally. For the near future, human interventions will be critical to sensing the strategic environment and deciding how and when to update algorithmic systems, otherwise, the risk of algorithmic mismatch will increase and invite disaster [17]. Having incongruent perspectives between AI and humans is often situational. Such new conceptualizations and engineering require efforts that are situationally relevant, such as humans-in-the-loop with AI. While most final decisions will rest with humans, both theoretical and empirical research demonstrates that humans can be influenced by AI advice by experiencing inflated uncertainty [18], [19]. Therefore, understanding human conformity to AI advice is necessary for developing and engineering new human-in-the-loop (HITL) concepts.

2.2 Conformity to Machine Logics

Human conformity to the other-than-self, such as society, culture, or human peers is well-established within literature. More recently, scholars have wanted to know if technologies could have similar influences on human conformity. In early research, for instance, Meehl [20] showed that people prefer human-derived decisions, rather than results from a statistical model, despite models outperforming human capabilities [21]. These earlier studies demonstrated that algorithms regularly outperformed human experts, on average, by about ten percent [22]–[24]. On the other hand, a study by Logg et al. [25] found that over the course of six experiments, people adhered more to algorithmic advice than human advice and that forecasting experts relied less on algorithmic advice, which hurt their overall performance.

Biased algorithms also pose numerous problems for perpetuating human proclivities to use AI recommendations inappropriately. For instance, workers can be misled by an AI's erroneous recommendation even when its results are clearly incorrect [26]. Also, users may tend to heavily conform to AI advice even after considering that advice as incorrect [27]. Similarly, people supported by AI-powered decision support systems often accept an algorithm's incorrect suggestions [28]. For these reasons, AI and algorithmic technologies have the potential to influence human behaviors in ways that skew irrational decisions – e.g., interference effects, which may result in suboptimal outcomes and automation bias.

2.3 Human vs. Machine Performance

Humans and machines both have strengths and weaknesses [29]–[32]. One of the first lists put forward was the framework of Machines-Are-Better-At and Humans-Are-Better-At (MABA-HABA) [33], [34]. The Fitts' List was developed at a time when a demarcation between human and machine tasks were more distinct [35], [36]. Newer AI-based systems are better at mimicking humans in the cognitive domain [36], which can add additional complexities to decision-making or augmenting situational awareness. For example, AI is classifying images previously performed by humans, but there still exists little understanding on how AI makes such classification decisions [37]. However, concerns have been raised that such algorithmic systems are easily duped by slight modifications [38]. Understanding how humans make sense of AI outputs is therefore necessary to explicate their role situationally in HITL systems.

2.4 Human-in-the-Loop Decision Making

AI is already supporting human decision-making in many ways and gradually taking over more decision space [39]. In the evolution from "tool" to "partner" and now to "leading" status, the conceptual coverage of human-computer interaction research [30] requires further conceptual work to map out and improve decision and information sciences. For example, critical to adopting AI into the decision-making processes is understanding the distinct roles humans can have within a HITL system to provide additional insights into situational decision-making. Crootof et al. [41] define a HITL as "an individual who is involved in a single, particular decision made in conjunction with an algorithm" (p. 12). Crootof et al. [41] also identified nine distinct roles a human can fulfill for a HITL system provided in Table 1. While such roles can specify oversight to algorithmic decisions, system performance and abilities of individual humans may further determine when, where, and how people are situationally incorporated into a HITL system.

Human Roles	Reason for adding a Human-in-the-Loop
1. Corrective Roles	Improve system performance, including error, situational, and bias correction
2. Reliance Roles	Act as a failure mode or alternatively stop the whole system from working under an emergency
3. Justificatory Roles	Increase the system's legitimacy by providing reasoning for decisions
4. Dignitary Roles	Protect the dignity of the humans affected by the decision
5. Accountability Roles	Allocate liability or censure
6. Stand-in Roles	Act as proof that something has been done or stand in for other humans and human values
7. Friction Roles	Slow the pace of automated decision-making
8. Warm Body Roles	Preserve human jobs
9. Interface Roles	Link the systems to human users

Table 1 – HITL Roles. Adapted from [41].

2.5 OODA Loop

In this paper, we use the observe, orient, decide, and act loop (OODA loop) as a general model for explicating human and AI decision-making. The OODA loop has both a rich history of use and is embedded in the psyche in areas from business to the military [42], [43]. The OODA loop provides an easy and understandable relationship between sensing and acting in any demanding environment [44]. Moreover, the OODA loop is the model observers refer to when discussing automation with humans "in-the-loop," "on-the-loop," or "out-of-the-loop" [45].

The OODA loop consists of four major components. Figure 1 provides a representation of the OODA loop as depicted by its creator, U.S. Air Force Colonel John Boyd. The observation step consists of an agent sensing and taking in information from the environment. This may consist of computerized sensors that generate output or ordinary human sense perceptions. Subsequent to observation is the orientation step where informational content is either pulled or pushed from sensors. The orientation step (i.e., orient) provides the key to understanding human information processing in decision-making and is considered the most critical component of the OODA loop [43]. Orientation contains the five components of genetic heritage, cultural traditions, previous experiences, new information, and analysis/synthesis interconnected to form a fully connected graph structure [46]. The output of the orientation step feeds forward and generates one or more courses of action or decisions that may be viewed as hypotheses in the mind of the agent. The decision step consists of deciding amongst two or more of the generated hypotheses. In the last step, decisions are acted upon, and the results are fed back into the observation step for evaluation and combined with the latest information for the process to iterate.

The five components of the orient phase and their interactions are challenging to operationalize. However, the quirks of the orientation step can be modeled with respect to a subsequent path to the act step, shown in Figure 1. These two paths are synonymous with Kahneman's [21] dual process theory of cognition [47]. As [47] points out, the implicit guidance path of the OODA loop is equivalent to Kahneman's [10] System-1, which consists of fast, intuitive decision making. The implicit guidance and control path goes directly to the act step of OODA loop because an observation matches a known heuristic for taking an action. The OODA loop also accommodates a System-2 approach which consists of a slower, more deliberate decision-making process. Implicit guidance accentuates OODA loop, in return, it demystifies linearly depicted decision processes in situational awareness model put forth by Endsley [48].

The OODA loop also contains several sources of feedback. These feedback loops supply information for the observation step of OODA loop. These nested loops can consist of varying time horizons and various levels of analysis. For instance, the two feedback loops that stem from the act step implies both an immediate action that can be observed and more holistic feedback within the larger environment. As a result, these feedback loops are combined with new information for the process to cycle again.

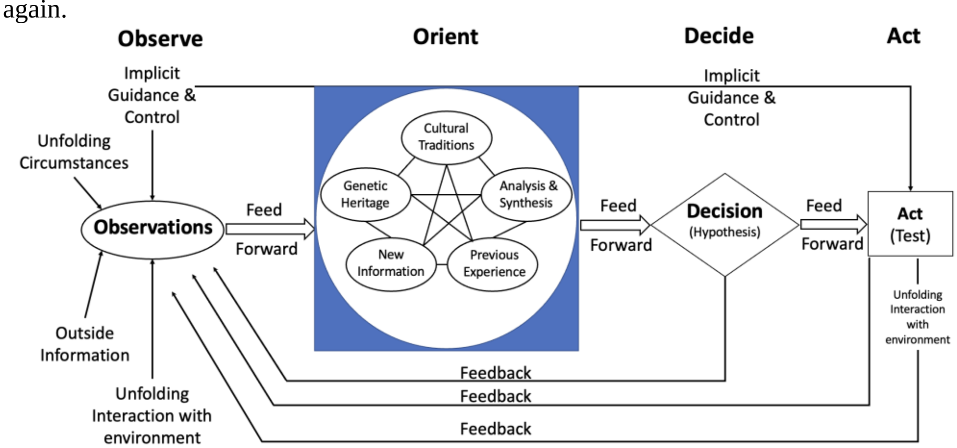

Figure 1 – John Boyd's OODA Loop. Adapted from [43].

2.6 The Orientation Step

Orientation step not only influences all other steps in the OODA loop, but also is an interactive process that additionally influences the character of future orientation [43]. In its current conceptualization, the OODA loop is agnostic to the inputs it receives, whether it originates from computer systems, direct observation, or another human agent. However, it has been shown empirically that people perceive information from AI differently [49]–[56]. Combined with tendencies to anthropomorphize computers in general [57], and AI systems particularly [58], humans may treat AI as more intelligent, which could lead to suboptimal decision outcomes.

The orientation step of the OODA loop is also a black-box process. This is problematic for at least three reasons. First, Boyd never provided justification or a clear explanation for choosing the five components that make up the orientation step. Moreover, it is unclear how such factors are weighted within a decision-making process. Second, it is unclear why the five components are depicted in a fully connected graph structure. While fully connecting all five components of the orient step in a complete graph may convey the notion of complexity, there exists no justification for this or supporting evidence in other bodies of research. Third, the orientation step, as currently conceptualized, does not indicate where one should begin the orientation step. For instance, what would determine first starting at the component of 'genetic heritage' over 'new information,' then subsequently moving to one of the other four components? In other words, what in the orient process triggers the recognition of affordances in the current situation [59]?

The current conceptualization of the orientation step in the OODA loop is inadequate for understanding HITL processes when it comes to humans and AI decision-making. To make the orientation step more applicable to decision processes that involve AI, it is necessary to ameliorate the relation between the orientation step and act step with a higher level of abstraction and analysis. Rather than prescribing specific components to the orientation step, abstracting to a higher level can accentuate the characteristic of orientation step in two ways. First, it can provide a more generalized level of analysis that is not constrained by Boyd's five components of the orient step, which currently lack justification. Second, abstracting to a higher level of analysis can provide a more tractable set of concepts for integrating human and machine intelligence. A tractable set of concepts includes only the perspectives necessary to form the full space of decision alternatives while also accounting for distinct types of uncertainty. In particular, including uncertainty within the orientation step, entails the concept of rational prediction deviation for reconciling human and AI situational awareness.

2.7 Rational Prediction Processes

To understand the concept of a rational prediction deviation, we start with a short scenario. Suppose, for instance, a vehicle equipped with AI-driving assistance exhibits an action that is incongruent to a human's intention or expectations that results in the human experiencing uncertainty (e.g., unexpected acceleration, braking, or steering correction). This uncertainty can be understood as a dichotomy of reference frames; these reference frames are classified as observer and participant frames. From the participant's reference frame, AI accesses more information, and a human can reduce this uncertainty by accessing the information in the environment either by pulling information or having AI push this information. This type of uncertainty is known as epistemic uncertainty, which is reducible by accessing more information from the environment. The second type of uncertainty is associated with the observer's reference frame. In this case, the human agent knows everything about the system, yet, knowing more does not reduce this type of uncertainty, which is called ontic uncertainty [60]. This type of uncertainty introduces what we call a rational prediction deviation between the human and AI-driver assistant. This deviation is observed when the human cannot process multiple perspectives associated to the event simultaneously [61].

Next, suppose before receiving any information from another information processing agent (e.g., AI), the human agent has his/her own perspective concerning the phenomenon in a situation. For simplicity, this perspective involves two decision outcomes, A and B; the human agent is probabilistically closer to deciding A ($Prob(A) \gg Prob(B)$). Subsequently, the human agent receives conflicting information about the phenomenon in question from an AI. This information is categorized as conflicting information because AI recommends "B," while the human $Prob(A) \gg Prob(B)$. As a result, the received information introduces a difference in rational prediction, or rational prediction deviation, because the two perspectives cannot be processed simultaneously [61], [62]. Yet, if two perspectives can be processed simultaneously, it means that $prob(A \cap B) = prob(B \cap A)$ is held true and perspectives are called compatible (e.g., obeys the commutative property). If the two perspectives cannot be processed simultaneously, $prob(A \cap B) \neq prob(B \cap A)$ is not held true (e.g., noncommutative), then the perspectives are identified as incompatible, which gives rise to what we term as a rational prediction deviation.

When perspectives are incompatible, three outcomes are observed: (1) a joint probability distribution cannot be formed; (2) non-negligible systemic order effects are observed; and (3) a violation of the law of the total probability occurs. However, critical to the modeling of a cognitive phenomenon is that the axioms of classical probability theory (CPT) do not support incompatible perspectives (i.e., $\mathrm{prob}(\mathrm{A} \cap \mathrm{B}) \neq \mathrm{prob}(\mathrm{B} \cap \mathrm{A})$) and a different system of probability must be used to model such cognitive phenomena [63].

One approach that has sought to address uncertainty with evidence accumulation models is Dempster-Shafer Theory (DST). DST operates on a mathematical property of conjunctive pooled evidence [64]. An assumption also built into DST is based on an independence of sources of such evidence [64]. However, modeling evidence accumulation in DST does not account for interaction and context effects when information is received in a particular order (e.g., being notified of a win before a loss). Additionally, DST models do not consider uncertainty inherent in the mind of a decision-maker, but only the environment. Models such as DST do consider aleatory uncertainty, but do not effectively model this uncertainty that is experienced by a decision-maker. An effective way of modeling aleatory uncertainty is taking into account ontic type of uncertainty, which is rather a more accurate determination of the source of aleatory uncertainty because it typifies the stochastic interactions and incompatibilities that may arise from different states of the world [65]. To address the shortcomings of DST, we offer to improve models with the quantum models of cognition for decision-making.

3 Quantum Models of Cognition for Decision-Making

The application of quantum probability theory (QPT) axioms provides a novel ontology for decision-making concepts for several reasons. First, the concept of measuring a system recognizes that observations change the system [66]. Second, QPT axioms are applicable to decision science because, contrary to conventional wisdom, judgments "create" rather than records what existed before a judgement [62]. Third, research has consistently demonstrated violations of CPT in human judgment and decision choices. As a result, violations of CPT make it difficult for developing combinations of HITL-AI systems because incongruencies will eventually emerge and give rise to uncertainty, which may result in poor decision-making. Equally important, assuming the generalization of one probabilistic model, may result in oversight and incompleteness in AI decision support for situational awareness. To move towards improving decision-making quality with AI, we must understand how violations of CPT are situationally manifested. For this, we turn to a discussion on several examples that show violations of classical probability in HITL-AI systems.

It is conjectured that QPT can help improve HITL-AI outcomes by anticipating situational dynamics that can induce interference effects that arise through human and machine interactions within decision-making. The combination of different outcomes due to the nature of different probabilities brought about by interference effects can prevent understanding potential emergent behaviors [68]. Therefore, to leverage the advantage of QPT, incompatibility of different perspectives must be ascertained [69].

One way to visualize the incompatibility of different perspectives in QPT is to use a Hilbert space. Hilbert spaces provide a geometric visualization to study both judgment and decision-making [70]. A Hilbert space is an N-dimensional vector space and the state vector, $|S\rangle$ is projected on to orthogonal basis vectors, A and B, that represent mutually exclusive choices (e.g., yes/no, true/false, buy/sell).

Furthermore, additional perspectives can be overlayed and rotationally offset from one another to represent different perspectives (i.e., Human, AI) as shown in Figure 2. Figure 2 demonstrates the different results one obtains depending on the order perspectives are taken. The difference between Figure 2(a) and Figure 2(b) displays graphically how commutativity is not obeyed in QPT, which is indicative of order effects and thus demonstrated by the differences in the lengths of the green projection. The QPT approach can therefore mathematically model situations in which incongruent perspectives can reveal different outcomes when the order of perspectives engenders different contexts to each other in information processing.

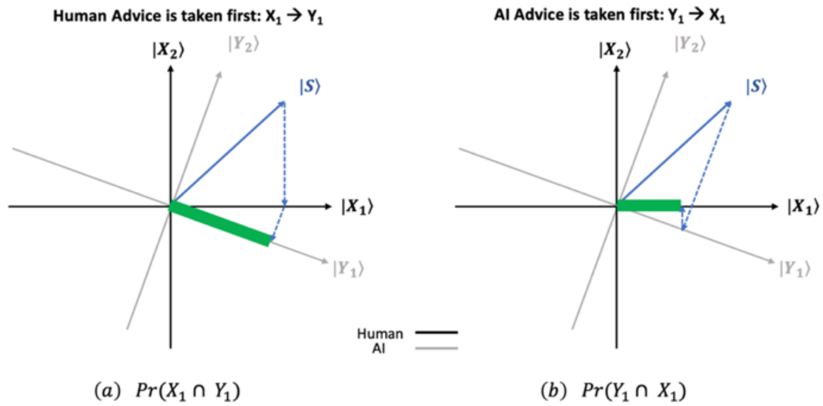

Figure 2 - Order Effects as demonstrated with QPT. The probabilistic difference between two situations is the difference between the length of square of the green bar in (a) and (b); this is conceptualized as a rational prediction deviation. Adapted from [69].

In QPT, events are represented as subspaces in Hilbert space. For example, in Figure 2, a system, S, is represented with the bases, $|X_1\rangle$ and $|X_2\rangle$, which are the subspaces of a two-dimensional Hilbert space. The cognitive system $|S\rangle$ is represented as the superposition of the subspaces $|X_1\rangle$ and $|X_2\rangle$, $|S\rangle = a|X_1\rangle + b|X_2\rangle$, in this equation, "a" and "b" symbolize the primitive amplitude, which is represented with complex numbers, which form the foundation of the probability calculation within QPT. The probability of an event, $|X_1\rangle$, is calculated by projecting the cognitive state vector |S⟩ on the subspace $|X_1\rangle$ (i.e., dashed lines). A projection operator can be written as $P_{X_1} = |X_1\rangle\langle X_1|$. If P_{X_1} operates on a superposition vector $P_{X_1}|S\rangle = |X_1\rangle\langle X_1|(a|X_1\rangle + b|X_2\rangle) = \langle X_1|a|X_1\rangle|X_1\rangle$, the inner product is $\langle X_1|X_2\rangle = 0$ because of orthogonality; the probability of event A is calculated as $p(X_1) = |\langle X_1|a|X_1\rangle|^2 = a^2$.

In real decision-making environments, decision makers may not explicitly question whether they trust information from an AI agent. However, the order in which information is received sets the context for subsequent information processing. For instance, while making decisions, decision makers receive information (from an AI or human) concerning the phenomenon of interest. After receiving information, the decision maker rationalizes the source of the information. In this rationalization process there are two scenarios. In the first scenario, the self, and the other (e.g., AI) perspectives (P) are commutative, which means $(P_{Self}P_{AI}) - (P_{AI}P_{Self}) = 0$ or close to zero. In this type of situation, there is no negligible conflict between the two perspectives for decision-makers. In the second scenario, the self and the AI perspective are incompatible when taken in different order. This means that $(P_{Self}P_{AI}) - (P_{AI}P_{Self}) \neq 0$. In this situation, depending on magnitude of the conflict between two perspectives, a decision maker experiences a rational prediction deviation, which means $\left(P_{Self}P_{AI}\right) - \left(P_{AI}P_{Self}\right) > 0$. Rational prediction deviations may be resolved through a number of different ways. Combinations of inquiring how AI justifies its decision(s) or searching out additional information can resolve such incongruencies. This does not mean that a rational prediction deviation must go to zero to decide. Decision makers will likely have different thresholds depending on their experience, context, and the nature of the decision (e.g., high-risk vs. low-risk scenarios).

In the case of having a single perspective, the state of the human can remain at this state and update his/her probability values concerning the two possible outcomes. When the decision process involves multiple perspectives to represent an other-than-self – in this case of an AI's perspective – the system state can start alternating between two different perspectives as shown in Figure 2. Modeling the situation shown in Figure 2 can result in extreme subjective predictions if the perspectives are incompatible, which can result in complex situations [61]. For these reasons, order effects can be more generalized in QPT. In the case of human systems, an initial judgment generates the context for subsequent decisions or categorizations, thus potentially making events non-commutative.

If humans can access the same phenomenon, he/she may be able to resolve the incongruent perspectives problem. However, when a human does not have direct observation of a particular phenomenon and relies on only an AI assessment and recommendation, it places the human in the AI loop. Figure 3 shows A HITL construct where the human does not have access to the phenomenon and is put in the AI decision-making loop for one or more of the nine reasons by Crootof et al. [41]. In these types of situations, interactions with an AI may introduce an ontic type of uncertainty while attempting to reduce epistemic uncertainty (e.g., accessing more information from the environment). More importantly, the human can be directed to a specific decision via these interactions [19].

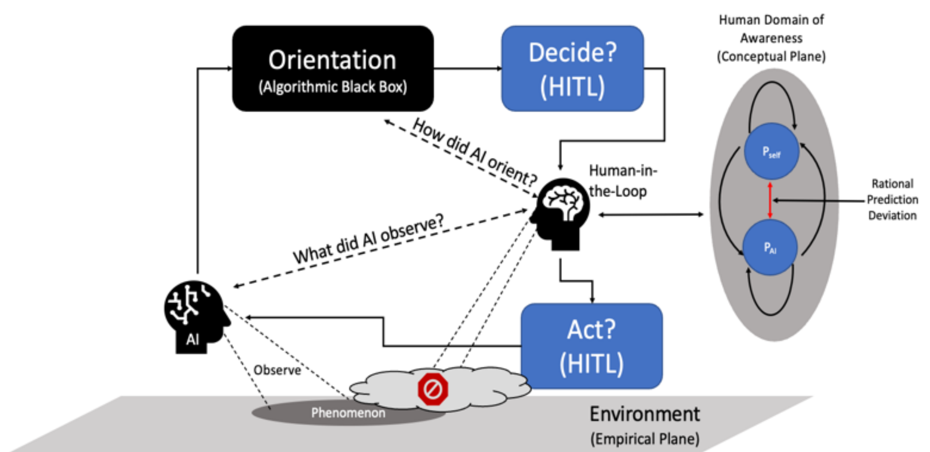

Figure 3 - Multiple Loops. The human agent has no access to the phenomenon and is reliant upon an AI's assessment and recommendation.

This lack of epistemic transparency prevents the human agent from fully grasping AI recommendations and decisions. More importantly, the axioms that are used to build AI systems may not be held true in complex situations, in which the violations of the law of total probability can be observed [18]. Consequently, an AI system, that is built with the axioms of CPT, may not be able to form a joint probability between its own perspective and the human perspective. Russell [71] characterizes this situation as, "rationality for two," which becomes impossible with each agent trying to second guess the other. When AI advice is incorrect, probabilistically a human agent perceives this answer by representing it with a particular perspective. As a result, decision-making with AI calls for modeling approaches that can capture not only uncertainties, but how to better engineer human and AI interactions to improve shared situational awareness. The following provides an example based on recent empirical findings.

As decision-making becomes increasingly supported by AI-based decision aids, the HITL could become subject to a number of different incompatible events as seen in Figure 4. These effects, suggested in the research, include order effects and concept categorization [9], [62], [72], [73]. These can generally be classified as interference effects. Cognitive processes can give rise to interference effects when attempting to process events that results in incompatible cognitive representation/perspectives [61], [74]. For instance, AI may provide annotated images or video as an input to the HITL process for building situation awareness. However, if a human screener understands the object to be of a different nature (e.g., AI says, "school bus;" human says, "military transport"), a human can experience an interference effect, because of these incompatible perspectives, a human (depending on the mental variety of the individual) may not process both perspectives (human's and AI's perspectives) simultaneously. In complex situations, these types of events may further be exacerbated by a host of additional inputs to situation awareness and decision-making.

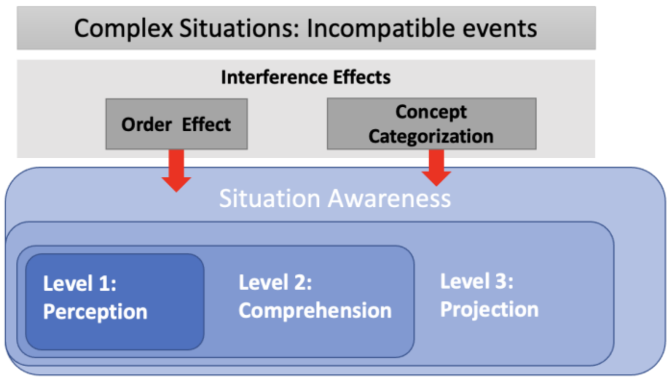

Figure 4 – Incompatible events in complex situations

In such cases, classical modeling techniques, such as Markov evidence accumulation models, require modeling the human's uncertainty with a definite state at any given time. However, such modeling of uncertainty does not truly capture a human's vacillation within their mind. Moreover, asking a question about a state can induce or create a definite state where none existed before [62]. Therefore, situation awareness techniques such as the Situation Awareness Global Assessment Technique (SAGAT) [75] can create states that may not have previously existed, as opposed to recording the awareness; as a result, in situations involving interference effects, the awareness between the two decision makers, can be significantly different.

Recent research has captured oscillatory behaviors in participant choices over time. Research performed by Kvam et al., [60], suggests that human behaviors can be dampened or bolstered in two-stage decision processes. However, conventional models can fail to capture such behaviors in a comprehensive way. An experiment that closely followed [60], suggests similar results when decision-making was supported by AI. The experiment outlined in [76] (forthcoming) elicited oscillatory behavior and a bolstering effect in the choice condition (explicitly agree/disagree with AI) over the no-choice condition (not eliciting whether agree/disagree with AI) shown in Figure 5.

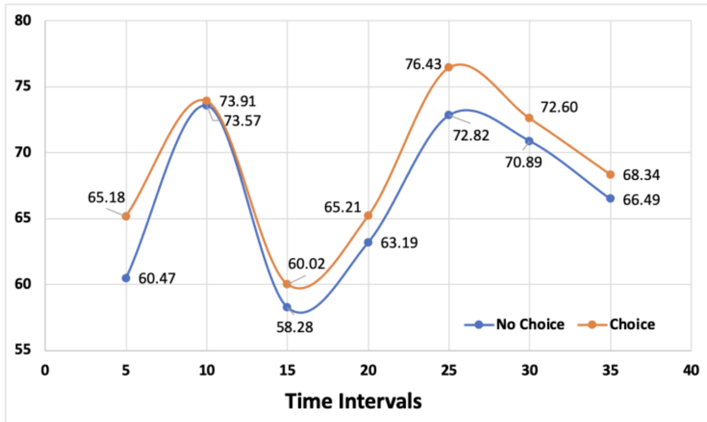

Figure 5 – Delegation strength over time intervals

These demonstrated behaviors pose issues for modeling human decision-making supported by AI, but this is not all. In the same experiment by [76], violations of total probability were observed with various threshold values of delegation [77] (forthcoming). Table 2 shows the violation of total probability, which cannot be captured with CPT. As a result, in these types of situations, the models of situation awareness and decision-making require techniques that can capture these behaviors in order to engineer human and AI interactions in a comprehensive way.

Timing	Condition	P(Agree)	P(Delegte|Agree)	P(Disagree)	P(Delegate|Disagree)	Total Probability (Delegate) - Choice	P(Delegate) - No Choice Condition	Pr(Delegate) minus Pr(Delegate) > 0
5 Sec	Choice	0.7097	0.5966	0.2903	0.1389	0.4637	0.4118	-0.0519
10 Sec	Choice	0.8495	0.7089	0.1505	0.2143	0.6344	0.6097	-0.0247
15 Sec	Choice	0.6231	0.6173	0.3769	0.2143	0.4654	0.4559	-0.0095
20 Sec	Choice	0.6833	0.6042	0.3167	0.2360	0.4875	0.4926	0.0050
25 Sec	Choice	0.8566	0.7225	0.1434	0.2895	0.6604	0.6182	-0.0422
30 Sec	Choice	0.8327	0.7143	0.1673	0.1556	0.6208	0.5941	-0.0267
35 Sec	Choice	0.7907	0.6716	0.2093	0.1296	0.5581	0.5257	-0.0324

Table 2 – Violations of total probability

4 DISCUSSION

In this paper, we discussed the different considerations for how human and AI situation awareness and decision making is conceptualized within HITL constructs. If AI takes over additional awareness/decision space as projected in the literature and media, researchers will need to carefully consider how humans and AI are integrated for improving situational awareness and decision-making.

Significant work lies ahead for developing HITL-AI systems. How to incorporate humans into AI-augmented situational awareness and decision-making will take on many different forms. It is clear from the research that humans and AI systems will continue to engage in shared decision-making [78], [79]; the questions will be what decisions are ceded to AI and how will organizations align their different rationalities. However, using QPT in the design of HITL-AI systems also opens the door for reevaluating previous research. For instance, earlier research such as the belief-adjustment model [80] that found order effects due to recency bias or weighting information based on temporal arrival, could be revaluated with QPT. Without capturing the correct dynamics, HITL-AI systems will exacerbate decision cycles as engineers attempt to reconcile human and AI rationalities. Future research will need to address how HITL-AI systems operate as time pressures increase and what may be done to improve decision-making in high-tempo and ethically significant operations with comprehensive frameworks.

5 Summary

QPT and similar efforts to formulate a concept of Quantum Decision Theory, have provided novel results that can better model uncertainty and human decision-making behaviors. Applying QPT to human and machine situational awareness models is still at a nascent stage of development at the human-machine dyad level [68], [81]. In short, QPT modelling can ameliorate interactions by providing a novel way to capture diverse types of uncertainty within human-AI decision systems and therefore, advance human-machine engineering efforts for improving decision-making and situation awareness.

Acknowledgment

This research is supported by Military Sealift Command Award N0003323WX00531.

References

[1] J. R. Galbraith, "Organization Design: An Information Processing View," *Interfaces*, vol. 4, no. 3, pp. 28–36, 1974.
[2] E. Brynjolfsson and T. Mitchell, "What can machine learning do? Workforce implications," *Science*, vol. 358, no. 6370, pp. 1530–1534, Dec. 2017, doi: 10.1126/science.aap8062.
[3] V. Bader and S. Kaiser, "Algorithmic decision-making? The user interface and its role for human involvement in decisions supported by artificial intelligence," *Organization*, vol. 26, no. 5, pp. 655–672, Sep. 2019, doi: 10.1177/1350508419855714.
[4] P. D. Harms and G. Han, "Algorithmic leadership: the future is now," *Journal of Leadership Studies*, vol. 12, no. 4, pp. 74–75, Feb. 2019, doi: 10.1002/jls.21615.
[5] N. Luhmann, *Organization and decision*. Cambridge, United Kingdom ; New York, NY: Cambridge University Press, 2018.
[6] R. Hertwig, T. J. Pleskac, and T. Pachur, *Taming uncertainty*, 1st ed. MIT Press, 2019. Accessed: Dec. 04, 2022. [Online]. Available: https://mitpress.mit.edu/9780262039871/taming-uncertainty/
[7] K. Brunsson and N. Brunsson, *Decisions: the complexities of individual and organizational decision-making*. Cheltenham, UK: Edward Elgar Pub, 2017.
[8] D. Lindebaum, M. Vesa, and F. den Hond, "Insights from 'The machine stops' to better understand rational assumptions in algorithmic decision making and its implications for organizations," *AMR*, vol. 45, no. 1, pp. 247–263, Jan. 2020, doi: 10.5465/amr.2018.0181.
[9] J. Jiang and X. Liu, "A quantum cognition based group decision making model considering interference effects in consensus reaching process," *Computers & Industrial Engineering*, vol. 173, p. 108705, Nov. 2022, doi: 10.1016/j.cie.2022.108705.
[10] D. Kahneman, *Thinking, Fast and Slow*, 1st edition. New York: Farrar, Straus and Giroux, 2013.
[11] X. Tan, J. Zhu, and T. Wu, "Dynamic reference point-oriented consensus mechanism in linguistic distribution group decision making restricted by quantum integration of information," *Group Decis Negot*, vol. 31, no. 2, pp. 491–528, Apr. 2022, doi: 10.1007/s10726-022-09775-0.
[12] D. Fogal and A. Worsnip, "Which reasons? Which rationality?," *Ergo*, vol. 8, no. 0, Art. no. 0, Dec. 2021, doi: 10.3998/ergo.1148.
[13] R. Suddaby, "From the editors: what grounded theory is not," *The Academy of Management Journal*, vol. 49, no. 4, pp. 633–642, 2006.
[14] J. Pearl and D. Mackenzie, *The book of why: the new science of cause and effect*. New York, NY: Basic Books, 2018.
[15] R. Frantz, "Herbert Simon. Artificial intelligence as a framework for understanding intuition," *Journal of Economic Psychology*, vol. 24, no. 2, pp. 265–277, Apr. 2003, doi: 10.1016/S0167-4870(02)00207-6.
[16] G. Kasparov, *Deep thinking: where machine intelligence ends and human creativity begins*. PublicAffairs, 2017.
[17] J. Johnson, "Automating the OODA loop in the age of intelligent machines: reaffirming the role of humans in command-and-control decision-making in the digital age," *Defence Studies*, vol. 0, no. 0, pp. 1–25, Jul. 2022, doi: 10.1080/14702436.2022.2102486.

[18] E. Jussupow, K. Spohrer, A. Heinzl, and J. Gawlitza, “Augmenting medical diagnosis decisions? An investigation into physicians’ decision-making process with artificial intelligence,” *Information Systems Research*, vol. 32, no. 3, pp. 713–735, Sep. 2021, doi: 10.1287/isre.2020.0980.
[19] L. Snow, S. Jain, and V. Krishnamurthy, “Lyapunov based stochastic stability of human-machine interaction: a quantum decision system approach.” arXiv, Mar. 31, 2022. doi: 10.48550/arXiv.2204.00059.
[20] P. E. Meehl, *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. in Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. Minneapolis, MN, US: University of Minnesota Press, 1954, pp. x, 149. doi: 10.1037/11281-000.
[21] C. Longoni, A. Bonezzi, and C. K. Morewedge, “Resistance to medical artificial intelligence,” *Journal of Consumer Research*, vol. 46, no. 4, pp. 629–650, Dec. 2019, doi: 10.1093/jcr/ucz013.
[22] R. M. Dawes, “The robust beauty of improper linear models in decision making,” *American Psychologist*, pp. 571–582, 1979.
[23] W. M. Grove, D. H. Zald, B. S. Lebow, B. E. Snitz, and C. Nelson, “Clinical versus mechanical prediction: a meta-analysis,” *Psychol Assess*, vol. 12, no. 1, pp. 19–30, Mar. 2000.
[24] H. Mahmud, A. K. M. N. Islam, S. I. Ahmed, and K. Smolander, “What influences algorithmic decision-making? A systematic literature review on algorithm aversion,” *Technological Forecasting and Social Change*, vol. 175, p. 121390, Feb. 2022, doi: 10.1016/j.techfore.2021.121390.
[25] J. M. Logg, J. A. Minson, and D. A. Moore, “Algorithm appreciation: People prefer algorithmic to human judgment,” *Organizational Behavior and Human Decision Processes*, vol. 151, pp. 90–103, Mar. 2019, doi: 10.1016/j.obhdp.2018.12.005.
[26] Y. Liel and L. Zalmanson, “What if an AI told you that 2 + 2 is 5? Conformity to algorithmic recommendations,” presented at the In International Conference on Information Systems 2020, Dec. 2020. [Online]. Available: https://aisel.aisnet.org/icis2020/hci_artintel/hci_artintel/17.
[27] P. Schmidt, F. Biessmann, and T. Teubner, “Transparency and trust in artificial intelligence systems,” *Journal of Decision Systems*, vol. 29, no. 4, pp. 260–278, Oct. 2020, doi: 10.1080/12460125.2020.1819094.
[28] Z. Buçinca, M. B. Malaya, and K. Z. Gajos, “To trust or to think: cognitive forcing functions can reduce overreliance on ai in ai-assisted decision-making,” *Proc. ACM Hum.-Comput. Interact.*, vol. 5, no. CSCW1, pp. 1–21, Apr. 2021, doi: 10.1145/3449287.
[29] J. C. F. de Winter and P. A. Hancock, “Reflections on the 1951 Fitts list: do humans believe now that machines surpass them?,” *Procedia Manufacturing*, vol. 3, pp. 5334–5341, Jan. 2015, doi: 10.1016/j.promfg.2015.07.641.
[30] S. W. A. Dekker and D. D. Woods, “MABA-MABA or abracadabra? Progress on human-automation co-ordination,” *Cognition, Technology & Work*, vol. 4, no. 4, pp. 240–244, Nov. 2002, doi: 10.1007/s101110200022.
[31] P. A. Hancock and S. F. Scallen, “The future of function allocation,” *Ergonomics in Design*, vol. 4, no. 4, pp. 24–29, Oct. 1996, doi: 10.1177/106480469600400406.
[32] R. S. Nickerson, “Man-computer interaction: a challenge for human factors research,” *Ergonomics*, vol. 12, no. 4, pp. 501–517, Jul. 1969, doi: 10.1080/00140136908931076.
[33] M. Cummings, “Man versus machine or man + machine?,” *IEEE Intelligent Systems*, vol. 29, no. 5, pp. 62–69, Sep. 2014, doi: 10.1109/MIS.2014.87.
[34] P. M. Fitts, “Human engineering for an effective air-navigation and traffic-control system,” National Research Council, Oxford, England, 1951.
[35] A. De Santis, B. Siciliano, A. De Luca, and A. Bicchi, “An atlas of physical human–robot interaction,” *Mechanism and Machine Theory*, vol. 43, no. 3, pp. 253–270, Mar. 2008, doi: 10.1016/j.mechmachtheory.2007.03.003.
[36] J. S. Metcalfe, B. S. Perelman, D. L. Boothe, and K. Mcdowell, “Systemic oversimplification limits the potential for human-AI partnership,” *IEEE Access*, vol. 9, pp. 70242–70260, 2021, doi: 10.1109/ACCESS.2021.3078298.
[37] S. T. Mueller, “Cognitive anthropomorphism of AI: how humans and computers classify images,” *Ergonomics in Design*, vol. 28, no. 3, pp. 12–19, Jul. 2020, doi: 10.1177/1064804620920870.
[38] D. Heaven, “Why deep-learning AIs are so easy to fool,” *Nature*, vol. 574, no. 7777, pp. 163–166, Oct. 2019, doi: 10.1038/d41586-019-03013-5.
[39] M. Frank, P. Roehrig, and B. Pring, *What to do when machines do everything: how to get ahead in a world of ai, algorithms, bots, and big data*. John Wiley & Sons, 2017.
[40] J. S. Wesche and A. Sonderegger, “When computers take the lead: The automation of leadership,” *Computers in Human Behavior*, vol. 101, pp. 197–209, Dec. 2019, doi: 10.1016/j.chb.2019.07.027.
[41] R. Crootof, M. E. Kaminski, and W. N. Price II, “Humans in the loop.” Rochester, NY, Mar. 25, 2022. doi: 10.2139/ssrn.4066781.
[42] J. Hasík, “Beyond the briefing: theoretical and practical problems in the works and legacy of john boyd,” *Contemporary Security Policy*, vol. 34, no. 3, pp. 583–599, Dec. 2013, doi: 10.1080/13523260.2013.839257.
[43] F. P. B. Osinga, *Science, strategy and war: the strategic theory of john boyd*. London: Routledge, 2006. doi: 10.4324/9780203088869.
[44] D. Blair, J. O. Chapa, S. Cuomo, and J. Hurst, “Humans and hardware: an exploration of blended tactical workflows using john boyd’s ooda loop,” in *The Conduct of War in the 21st Century*, Routledge, 2021.
[45] W. C. Marra and S. K. McNeil, “Understanding ‘the loop’: regulating the next generation of war machines,” *Public Policy*, vol. 36, p. 47, 2013.

[46] I. T. Brown, *A new conception of war: John Boyd, the U.S. Marines, and maneuver warfare*, First edition. Quantico, VA: Marine Corps University Press, 2018.

[47] N. J. Johnson, "Boyd's real OODA-loop and fencing," 2014. [Online]. Available: https://ooda.de/media/nicholas_j_johnson_-_boyds_real_ooda_loop_and_fencing.pdf

[48] M. R. Endsley, "Toward a Theory of Situation Awareness in Dynamic Systems," *Hum Factors*, vol. 37, no. 1, pp. 32–64, Mar. 1995, doi: 10.1518/001872095779049543.

[49] R. Gelles, D. McElfresh, and A. Mittu, "Project report: perceptions of AI in hiring," College Park, MD, Oct. 2018. [Online]. Available: https://anjali.mittudev.com/content/Fairness_in_AI.pdf

[50] D. Kaur, S. Uslu, K. J. Rittichier, and A. Durresi, "Trustworthy artificial intelligence: a review," *ACM Comput. Surv.*, vol. 55, no. 2, pp. 1–38, Mar. 2022, doi: 10.1145/3491209.

[51] R. Kocielnik, S. Amershi, and P. N. Bennett, "Will you accept an imperfect ai?: exploring designs for adjusting end-user expectations of ai systems," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, Glasgow Scotland Uk: ACM, May 2019, pp. 1–14. doi: 10.1145/3290605.3300641.

[52] J. D. Lee and K. A. See, "Trust in automation: designing for appropriate reliance," *Hum Factors*, vol. 46, no. 1, pp. 50–80, Mar. 2004, doi: 10.1518/hfes.46.1.50_30392.

[53] J. Meyer and D. Remisch, "An experiment on the impact of information on the trust in artificial intelligence," in *HCI in Business, Government and Organizations: 8th International Conference, HCIBGO 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings*, Berlin, Heidelberg: Springer-Verlag, Jul. 2021, pp. 600–607. doi: 10.1007/978-3-030-77750-0_39.

[54] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?': explaining the predictions of any classifier," *ArXiv*, Aug. 2016, Accessed: May 07, 2022. [Online]. Available: http://arxiv.org/abs/1602.04938

[55] G. Veletsianos, A. Doering, and C. Miller, "Conflicts in human-computer interactions: a framework for designing effective message exchange between humans and pedagogical agents," *Annual Proceedings-Anaheim*, vol. 2, no. 29, p. 6, 2007.

[56] S. Zuboff, *In the age of the smart machine: the future of work and power*, Reprint edition. New York: Basic Books, 1989.

[57] C. Nass, J. Steuer, and E. R. Tauber, "Computers are social actors," *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 72–78, Apr. 1994.

[58] D. Proudfoot, "Anthropomorphism and AI: Turing's much misunderstood imitation game," *Artificial Intelligence*, vol. 175, no. 5, pp. 950–957, Apr. 2011, doi: 10.1016/j.artint.2011.01.006.

[59] S. A. Kauffman and A. Roli, "What is consciousness? Artificial intelligence, real intelligence, quantum mind, and qualia." arXiv, Jun. 29, 2022. doi: 10.48550/arXiv.2106.15515.

[60] P. D. Kvam, J. R. Busemeyer, and T. J. Pleskac, "Temporal oscillations in preference strength provide evidence for an open system model of constructed preference," *Sci Rep*, vol. 11, no. 1, Art. no. 1, Apr. 2021, doi: 10.1038/s41598-021-87659-0.

[61] M. Canan and A. Sousa-Poza, "Pragmatic idealism: towards a probabilistic framework of shared awareness in complex situations," in *2019 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, Apr. 2019, pp. 114–121. doi: 10.1109/COGSIMA.2019.8724208.

[62] J. R. Busemeyer and P. D. Bruza, *Quantum models of cognition and decision*, Reissue edition. Cambridge: Cambridge University Press, 2014.

[63] F. Farahmand, "Quantum Cognition: A Cognitive Architecture for Human-AI and In-Memory Computing," *Computer*, vol. 56, no. 4, pp. 135–138, Apr. 2023, doi: 10.1109/MC.2023.3242056.

[64] K. Sentz and S. Ferson, "Combination of evidence in Dempster-Shafer theory," Sandia National Lab. (SNL-NM), Albuquerque, NM (United States); Sandia National Lab. (SNL-CA), Livermore, CA (United States), SAND2002-0835, Apr. 2002. doi: 10.2172/800792.

[65] P. D. Kvam and T. J. Pleskac, "A quantum information architecture for cue-based heuristics," *Decision*, vol. 4, pp. 197–233, 2017, doi: 10.1037/dec0000070.

[66] L. Floridi, *The philosophy of information*. OUP Oxford, 2013.

[67] E. M. Pothos, O. J. Waddup, P. Kouassi, and J. M. Yearsley, "What is rational and irrational in human decision making," *Quantum Reports*, vol. 3, no. 1, Art. no. 1, Mar. 2021, doi: 10.3390/quantum3010014.

[68] R. G. Lord, J. E. Dinh, and E. L. Hoffman, "A quantum approach to time and organizational change," *AMR*, vol. 40, no. 2, pp. 263–290, Apr. 2015, doi: 10.5465/amr.2013.0273.

[69] T. Darr, R. Mayer, R. D. Jones, T. Ramey, and R. Smith, "Quantum probability models for decision making," presented at the 24th International Command and Control Research & Technology Symposium, ICCRTS, 2019, p. 20. [Online]. Available: https://static1.squarespace.com/static/53bad224e4b013a11d687e40/t/5dc42d54e8437d748186b031/1573137749772/24th_ICCRTS_paper_8.pdf

[70] M. Canan, "Non-commutativity, incompatibility, emergent behavior and decision support systems," *Procedia Computer Science*, vol. 140, pp. 13–20, 2018, doi: 10.1016/j.procs.2018.10.287.

[71] S. Russell, *Human compatible: artificial intelligence and the problem of control*. Penguin, 2019.

[72] T. Holtfort and A. Horsch, “Social science goes quantum: explaining human decision-making, cognitive biases and Darwinian selection from a quantum perspective,” *J Bioecon*, vol. 25, no. 2, pp. 99–116, Aug. 2023, doi: 10.1007/s10818-023-09334-w.
[73] F. Xiao, “CEQD: A Complex Mass Function to Predict Interference Effects,” *IEEE Transactions on Cybernetics*, vol. 52, no. 8, pp. 7402–7414, Aug. 2022, doi: 10.1109/TCYB.2020.3040770.
[74] R. Blutner and P. beim Graben, “Quantum cognition and bounded rationality,” *Synthese*, vol. 193, no. 10, pp. 3239–3291, Oct. 2016, doi: 10.1007/s11229-015-0928-5.
[75] M. R. Endsley, “Direct Measurement of Situation Awareness: Validity and Use of SAGAT,” in *Situational Awareness*, Routledge, 2011.
[76] S. A. Humr, M. Canan, and M. Demir, “Temporal evolution of trust in artificial intelligence-supported decision-making,” in *Human Factors and Ergonomics Society*, Washington, DC: SAGE Publications, Oct. 2023.
[77] S. A. Humr, “Temporal evolution of trust in artificial intelligence-supported decision-making,” Doctoral dissertation, Naval Postgraduate School, Monterey, CA, 2023.
[78] L. M. Blaha, “Interactive OODA Processes for Operational Joint Human-Machine Intelligence,” presented at the In NATO IST-160 Specialist’s Meeting: Big Data and Military Decision Making, NATO, Jul. 2018. Accessed: Jun. 06, 2023. [Online]. Available: https://www.sto.nato.int/publications/STO%20Meeting%20Proceedings/STO-MP-IST-160/MP-IST-160-PP-3.pdf
[79] K. van den Bosch and A. Bronkhorst, “Human-AI Cooperation to Benefit Military Decision Making,” in *In NATO IST-160 Specialist’s Meeting: Big Data and Military Decision Making*, NATO, Jul. 2018. Accessed: Jun. 06, 2023. [Online]. Available: https://www.karelvandenbosch.nl/documents/2018_Bosch_etal_NATO-IST160_Human-AI_Cooperation_in_Military_Decision_Making.pdf
[80] V. Arnold, P. A. Collier, S. A. Leech, and S. G. Sutton, “Impact of intelligent decision aids on expert and novice decision-makers’ judgments,” *Accounting & Finance*, vol. 44, no. 1, pp. 1–26, 2004, doi: 10.1111/j.1467-629x.2004.00099.x.
[81] D. Mortimore, M. Canan, and R. R. Buettner, “Two probability theories and a garbage can,” *Comput Math Organ Theory*, Jun. 2023, doi: 10.1007/s10588-023-09378-3.