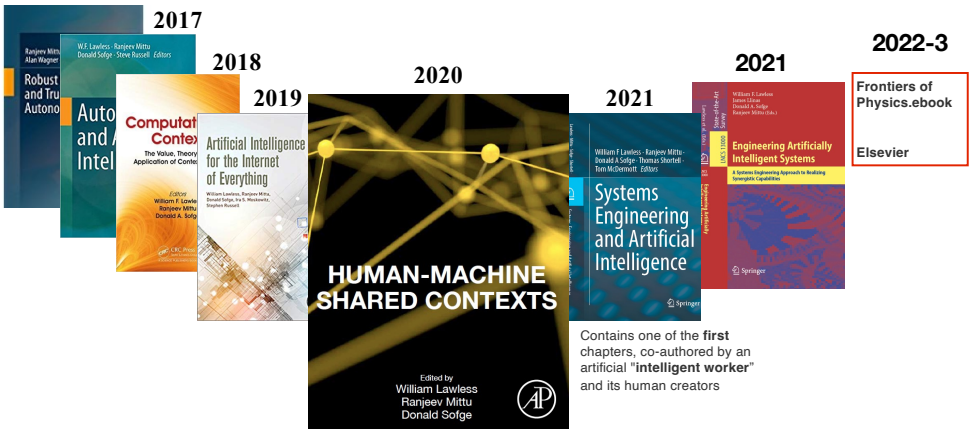
Risk determination versus risk perception: From misperceived drone attacks, hate speech and military nuclear wastes to human-machine autonomy

W.F. Lawless (w.lawless@icloud.com) & Don Sofge
(donald.sofge@nrl.navy.mil)

1

Background: Our research team's status

2016 2017 2018 2019 2020 2021 2021 2022-3



Contains one of the **first** chapters, co-authored by an artificial "**intelligent worker**" and its human creators

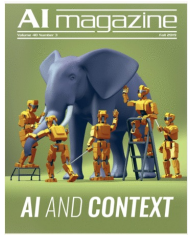
Frontiers of Physics.ebook

Elsevier

Human-Machine Shared Contexts **has been nominated by Elsevier for the ASIS&T "Information Science" Book of the Year in 2020**
(*https://www.asist.org*)

2

Background: Calls for Special Issues



- Special Issue *AI Magazine* (2019),"Computational context for human-machine teams" (Lawless et al., 2019)
  - e.g., **Uber's self-driving car killed a pedestrian in 2018**: *NTSB* (2018): The car saw the pedestrian ~ 6s ahead, interlock prevented braking1.2s ahead, car's classification oscillated; the **human operator** took the wheel < 1s before impact and applied brakes ~ 1s after impact
  - The human operator was distracted, and the car did not alert its human teammate earlier nor ask for help; ***both were poor team players***

- **Closing May 14th*: Frontiers in Physics***: Interdisciplinary Approaches to the Structure and Performance of Interdependent Autonomous Human Machine Teams and Systems (A-HMT-S) (Lawless, Sofge & LoFaro); https://www.frontiersin.org/research-topics/25455/interdisciplinary-approaches-to-the-structure-and-performance-of-interdependent-autonomous-human-mac;

- **Closing Jul 31st**: ***Entropy***: "An Entropy Approach to the Structure and Performance of Interdependent Autonomous Human Machine Teams and Systems"; see at https://www.mdpi.com/journal/entropy/special_issues/Human_Machine_Teams

3

AI is changing society & war: Swarms versus A*utonomous* H-M Teams imply closed versus open systems

- **Closed Systems: Often do not generalize**
- **Open Systems: Examples**

Russians believed its base in Syria was "*impregnable*" (Grove, 2018)

  - **Swarms killed 2 Russian soldiers** at their Russian airbase in Syria on New Year's Eve (Grove, 2018, WSJ), injuring 10 and damaging 6 planes ... Russia denied any deaths, US disavowed involvement.
  - **Highly-maneuverable hypersonic missiles dramatically reduce reaction times increase risks**
  - **In the future, can an A-HMT-S defeat the threat posed by "*Skynet*" slave swarms?**
    - Generalizing from Lincoln's (1838, 1/27) Lyceum speech: "Yes … As a nation of freemen, we must live through all time, or die by suicide."
  - **If yes, will humans give authority to intelligent autonomous machines, teams & systems?**
    - **Invited article: *AI Bookie Bet* in *AI Magazine*** (Sofge et al., 2019): In 5 years, **yes**, a machine (e.g., commercial airliner, train, ship, car) will be authorized to take responsibility from its human operator. Pro: W.F. Lawless; Con: Ranjeev Mittu; Referee: Don Sofge
  - **Editors** (7/5/19), ***NYT*** **and the United Nations reject "autonomous" systems as potentially immoral**

4

DoD's Joint Risk Framework = logic (*2016*)

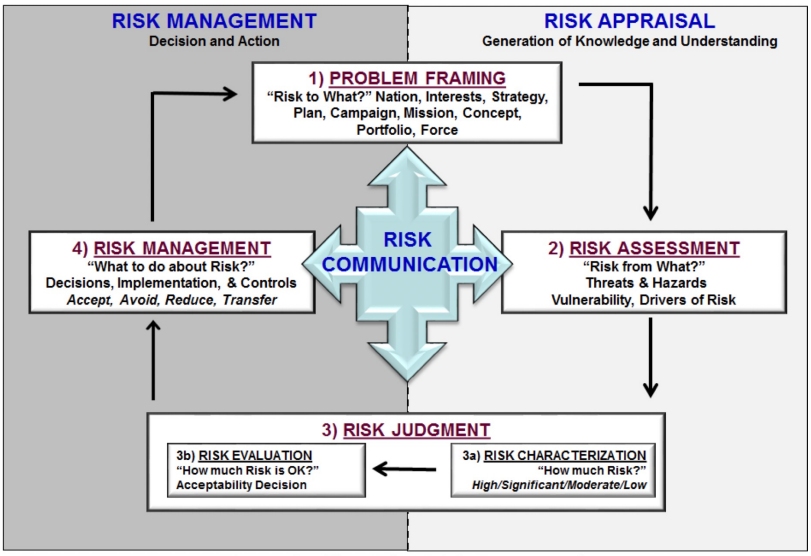


Figure 3. The Joint Risk Framework

5

But logic didn't prevent the Afghanistan drone strike in 2021

- **What went wrong?**
  - **On August 26, 2021, 13 American service members were killed**, 15 more injured by ISIS suicide bomber …
  - **On August 29, 2021, a DoD drone strike in Afghanistan killed 10 civilians, a tragic mistake**
    - Gen. Said, USAF: **highly emotional state** may be prevented by "**red teams**"
- How to distinguish perceptions -> driven by highly **emotional *entanglements*** in the rush to act from the need to act quickly?
  - e.g., USS Vincennes 1988; USS Greeneville 2001; USS Fitzgerald & USS McCain 2017
- **Literature**: Logic & rationality fail facing **uncertainty & conflict** (Mann, 2018); e.g., **games**
  - **Interview**: General Zinni: **war games -> "preordained proofs"** (Augier & Barrett, 2021).

6

DoD (2021): Risk determination versus subjective perception

DoD's rational new *Methodology* assesses subjectivity, but not emotion nor red teams

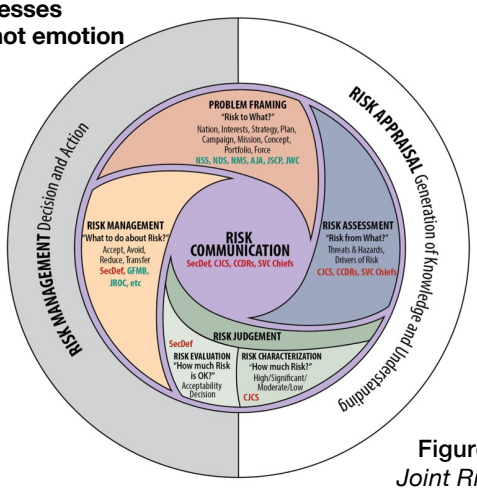


Figure 2. CJCSM 3105.01A. *Joint Risk Analysis Methodology* (*p.* B2, Encl. B)

Putting AI in the Critical Loop: Assured trust and autonomy in human-machine teams. AAAI Symposium, March 21, 2022

7

Hate mail: Human perception versus machine determination

- In an interview, Sheryl Sandberg, Facebook's COO, claimed FB's **algorithms detected 91% of 1.5 million** posts detected for violating its hate policy.
- However, the software engineers and scientists at Facebook reported to *BuzzFeed* (Mac & Silverman, 2020):
  - **1 of every 1,000 pieces of content — or 5 million of the 5 billion pieces of content posted daily to Facebook** — violates its rules on hate speech.
  - **However, with AI and third-party help, Facebook was "deleting less than 5% of all of the hate speech** posted to Facebook"
- FB's claims versus its reality leave the public under-informed.
  - The under-informed are vulnerable to the **risk of manipulation** (Chalmers, 2022)

8

Can Social Science, a closed system study of the individual, generalize to open systems to determine the risk of autonomy? No, it's insufficient

Leading Theory	Leading theory and theorist	Theory invalidated by:
Self-Esteem	Diener (1984); hailed by the American Psychological Association (1987) as "important" to success	Baumeister et al., 2005
Ego-Depletion	Baumeister & Vohs, 2007	Hagger et al., 2016
Implicit Attitudes Theory (racism)	Greenwald et al., 1998	Blanton et al. (Tetlock), 2009
Superforecasters	Tetlock & Gardiner, 2015	Brexit & Trump, 2016
Honesty (2012), PNAS	Lisa, Ariely & Bazerman et al., 2012; Ariely, chief developer & promoter of the "Honesty" scale, used fabricated data	Berenbaum, 2021, Editor in Chief, PNAS, retracted original article about "honesty."

Nosek's (2015) **Replication Project attempts to fix the validation problem, while ignoring the inability to generalize**!

9

Can we determine risks for autonomous human-machine teams with closed system models (e.g., Shannon, rational, games)?

- **Yes (*comms*) and No (*orgs*). Shannon IT assumes i.i.d. data** (*independent*), implying **no social effects** (Schölkopf et al., 2021)
  - Conant (Shannon): Teams work best under minimal communication (*i*)
  - But zero Shannon info => min *i* for teams -> *swarms of slaves*:
    - however, 'Knowledge ... [is useless to] a slave'; Frederick Douglass (1892, p. 103)
  - **Contradicting Shannon, the best science teams are *interdependent*** (Cooke & Hilton, 2015; Cummings, 2015; Bisbey et al., 2019)
- **Can rational theory work in open systems?**
  - **Rational theory -> observations of social data aggregated ≠ recreate social**
  - **Rational theory *fails* facing conflict or uncertainty (Mann, 2018)**
    - **Machine Learning = context dependency => closed system (e.g., Uber)**
    - **Interdependence = state dependency => open system ~ Von Neumann IT**

10

Starting over with open systems: A history of "inter-dependence" (*i*): Quantum mechanics, Social Psychology, Systems Engineering & Science of Teams

- First, **Schrödinger** 1935 (p. 555) describes quantum theory:
  - … the **best possible knowledge of a *whole* does not necessarily include the best possible knowledge of all its *parts*** … [because they are not] ***independent***
- **Lewin** (1951), founder of Social Psychology: asserted that the "whole is greater than the sum of its parts."
- From the ***Systems Engineering Handbook*** (Walden et al., 2015), "A System is a set of elements in interaction" (Bertalanffy, 1968) where systems "… often exhibit emergence, … meaningful only when attributed to the whole, not to its parts" (Checkland, 1999).
- **But if parts of a whole team are not independent** (e.g., Boltzmann's reduction in *dof*)**, can a state of interdependence among complementary parts confer a thermodynamic advantage to the whole**? ***Yes*****!** (Cooke & Lawless, 2021).
- **Social systems are autonomous, interdependent & open** (Jones, 1998). ***However,* what is interdependence?**

Interdependence Theory: Arbitrarily separated, *i* produces 3 effects: bistability, measurement problem & non-factorability

- **Bistability** (individual <-> teammate; two-sided stories; multitasking (MT) in orthogonal roles; competition; deception)
- **Measurement problem: M(*i*) collapses bistable information from orthogonal roles ~ *zero* correlations,** or inner product $\langle a, b \rangle = 0$ => two-sided stories collapse to the dominant one)
- **Non-factorable information** (*endless* debates; divorce; Brexit; quantum interpretations; e.g., Weinberg, 2017); ***Tradeoffs*** tested in free markets ($\partial I_{bistable} / \partial t$), courts, ***"red teams***"

Theory: *i* = Bistability

When no predators are nearby, forests become unhealthy forests (Carroll, 2016)

Predators nearby produce a healthy forest (Carroll, 2016)

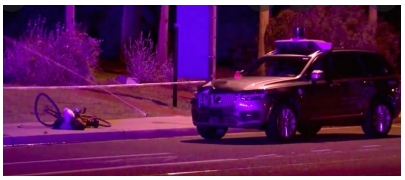
- **Bi-stable illusions tell us that:**
  - **"… the visual system chooses only a single interpretation at a time, never a mixture."** (Eagleman, 2001)

- **Teams can multi-task, individuals cannot** (Wickens, 1992)

Bistability implies *duality,* but closed systems assume risk perceptions and actions are 1:1

- **Why is social science hard to *replicate*** (Nosek, 2015)? ***Possibly because*:**
  - **Cognitive concepts imply 1:1 *implicit behaviors*** (e.g., implicit racism, in Greenwald & Banaji, 1995; not valid, in Blanton et al., 2009); or,
  - **Behaviors imply 1:1 *implicit cognitive beliefs*** (e.g., *Inverse Reinforcement Learning,* in game theory; Amadae, 2017; praxeology, or action by von Mises; Thagard, 2019).
- Why don't self-reported observations of behavior = actual behavior? (Zell & Krizan, 2014);
  - ***Despite women reportedly taking HIV prevention pills 95% of the time, => drug failed, the secondary measure of effective drug levels in their blood when reported was <26%*: "There was a profound discordance between what they told us … and what we measured,"** infectious disease specialist Jeanne Marrazzo said. (Cohen, 2013)
  - ***Interdependence* is respectable again** (e.g., see National Acad. Sci., in Cooke & Hilton, 2015; Endsley, 2021), **driving society-technology evolution** (Ponce de Leon et al., 2021)

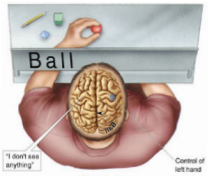
Pentagon Briefing: tragic drone attack driven by one *misperceived risk* (DoD, 2021)



Independent agents (Uber car & operator) who do not share perceptions -> fatality

Theory: M(*i*) -> Measurement problem = collapses 2-sided stories to one => the dominant risk perception

Measurement in social systems affects what's measured => *state-dependency* (Davies, 2020)



Authoritarian & gang one-sidedness
-> de-evolution; e.g., N Korea

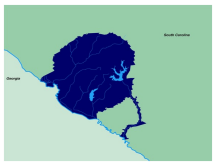
Gazzaniga (2011)**: the left split-brain did not see what the right half saw**

Putting AI in the Critical Loop: Assured trust and autonomy in human-machine teams. AAAI Symposium, March 21, 2022 15

15

US DOE's one-sided mandates before 1983 saw no risks to its practices that grossly contaminated the environment across the U.S.

Before 1983, DOE as sole authority -> nuclear waste mis-management
(Lawless, 1985)



Public awareness in 1985 stopped DOE's use of cardboard boxes; *public awareness* in 2000 accelerated the closure of its old radwaste burial ground

- **DOE's nuclear waste management today = 2-sided factions, information processing & better decisions**

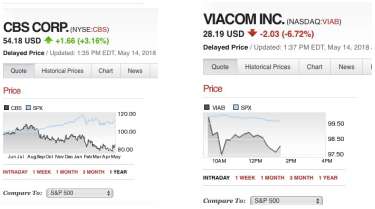
- DOE promoted minority (consensus) rules -> conflict on Hanford's Citizen Advisory Board, but not on the majority-ruled SRS-CAB (Bradbury et al., 2003)

***Policy & trust:* At DOE SRS today** (majority rules), **with competing oversight and debate by several groups** (e.g., DNFSB)**, DOE's decisions significantly improved** (Lawless et al., 2014; Akiyoshi et al., 2021).

Putting AI in the Critical Loop: Assured trust and autonomy in human-machine teams. AAAI Symposium, March 21, 2022 16

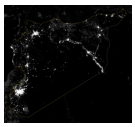
16

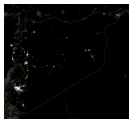
Endless fights are common, but who is right?



Long-running fights impact profits 2016-18

Non-factorable information: 2-sided factions & conflicts are costly





Endless civil-war disrupts social structures: Syria 2012 versus 2014

- **Perturbations of poorest teams generate too much Shannon *I* (e.g., CDM -> poor Russian teams; in WSJ)**:
  - Interdisciplinary science teams are the worst at performing team science (Cummings, 2015)
- **Perturbations of the best teams act as a unit (Schrodinger, 1944) ~ little information generated, => subadditivity:**
  - **Von Neumann's Subadditivity: $S(\rho_{AB}) \leq S(\rho_A)+S(\rho_B)$**

If *interdependence* -> "bewilderment" (Jones, 1998, p. 33), what can theory predict?

- Closed Systems: Homo economicus is a theoretical abstraction that some economists use to describe a rational human being
  - Rudd [2021]: "Mainstream economics is replete with ideas that "everyone knows" to be true, but that are actually arrant nonsense" (e.g., surveys of inflation ≠ inflation)
- **Previous *Closed-System* Predictions on *Redundancy*:**
  - **Social network analysts**: "As *redundancy* increases, the network becomes more efficient …" (Centola & Macy, 2007, p. 716)
  - **NAS: team size: "many hands make light work,"** but that team size remained *unsolved* (Cooke & Hilton, 2015)
- **My *Open-Systems* Prediction: Redundancy reduces *i* & MEP**

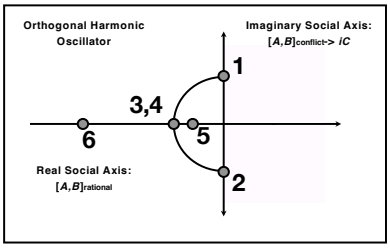
Results: Facing uncertainty in open systems, how to reduce risk?

- **1st test: Of top oil firms (Kullback-Leibler divergence), contradicting social network theory & NAS, redundancy impedes *interdependence* & free choice** (Lawless, 2017a, *JMP*).
  - **e.g., in 2018, Exxon had 1/8th as many employees as Sinopec**
- **2nd test: Of militaries around the globe (K-L divergence): Authoritarians increase redundancy & corruption** (Lawless, 2017b, *Frontiers in Physics*).
- **3rd test: Intelligence "tunes" a team's interactions to reduce risk & increase MEP**
  - **Support for MEP based on a nation's patents produced v HDI from UN data for 19 MENA nations & Israel: school and patents** (r =.62, p < .05)

- **Orthogonality** (Lawless, 2019, *Foundations of Science*)
  - **Finding (2019): Schooling (knowledge training) ~ total patents**
  - **Finding (2001): Flight training (not schooling) = Top fighter pilots**

Open System Models: Unless prevented, in shared contexts, humans facing uncertainty debate alternatives to reduce risk ("red teams")

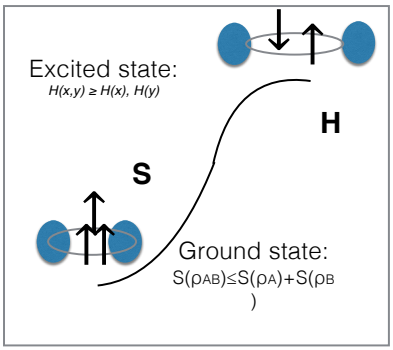
Model-1: Opposing beliefs serve to challenge traditions to improve behavior, leading to more evolution in free, but not closed, systems



A. Minority rules (consensus-seeking; ***easy to model***) **-> minority-control impedes alternatives; + stability, *independence*;** e.g., Asch's (1951) line study -> poor decisions **(e.g., China; Cuba; N. Korea)**

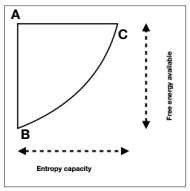
B. Majority rule (debate builds factions; + instability; but checks & balances control factions; ***hard to model***); e.g., best legal decisions: "***informed assessment of competing interests***" (Justice Ginsburg, 2011)

Model-2 (SEP): **Ground** ($S_{Whole} \leq \sum S_i$) **vs. excited** ($S_{Whole} \geq \sum S_i$)



Lawless, 2020, *Entropy;* 2021, *Informatics;* excited state calculations from Kang, 2002, in Lawless, 2002

Open Systems, *Notional Model*: An organization provides Helmholtz *free energy* "from an external source … [to maintain its] dissipative structure" (Prigogine, 1977) by offsetting its waste and products produced. We illustrate with a notional diagram of *free energy* from Gibbs (closed systems).



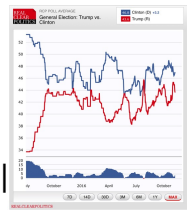
4th test: Metrics for complementarity

- **A new equation leads to several discoveries: C ≈ ΔSEP * ΔMEP**
  - **First, tradeoffs: ΔSEP <–> ΔMEP**
  - **Second, for maximum performance:** $lim_{\Delta SEP->min} \Delta MEP \rightarrow max$
  - **Third,** $\Delta SEP_1 > \Delta SEP_2$ **=> a *vulnerability* in** SEP_1 **that decides competition**
    - **For example:**
      - **Enforced cooperation increases systemic vulnerability to risk & the need to *steal* innovations** (e.g., China, in Baker, 2015; Ratcliffe, 2020).
      - **Monopolies increase organizational vulnerability to risk & the need to *steal* innovations from clients** (e.g., Amazon, in Mattioli, 2020, *WSJ*).
      - **Redundancy in organizations increases vulnerability to shocks** (e.g., oil collapse in 2020)
  - **Fourth, to enhance deception: Minimize** $SEP: Team \rightarrow Unit \Rightarrow lim_{dof->unit} ln(Team) = minimum$
  - **Fifth, to test risk perceptions & break emotion entanglement ("group think"), red teams -> compromise**

Speculation: A model matching experience

- 1. MT = orthogonal roles => fewer *dof* (e.g., marriage; team; restaurant).
  - e.g., The "*performance of a team is **not decomposable** to, or an aggregation of, individual performances*" (Endsley, 2021, p. 11).
  - **Tensors** = factorability => ***decomposable***, no interdependence
- ***2.* From portfolio theory** (Markowitz, 1952):
  - Individuals MT poorly (Wickens, 1992), but teams MT perfectly
  - *lim n->N team* r_{MT} -> 0 (e.g., orthogonal pairs: conductor-pianist)
- **3. Superordinates = team goals -> *"asabiyya"*** (Ibn Khaldun ~ 1400)
  - *lim n->N team* $r_{Superordinate}$ -> 1 (e.g., a unit's "fierce cohesion")
- 4. **Emotion**: Ground states (good team fit) v excited states (poor team fit)
- 5. **Words into music**: **https://melobytes.com/en/app/melobytes** (see the words: "conservative" and "liberal")
  - 3-block harmonic oscillators at **resonance** -> ***MEP***
- 6. **Crystal**: an ion is at higher entropy than a crystal
  - Intelligence "tunes" teammates to fit with each other -> ***-SEP, +MEP***

Mapping Linear Algebra to evidence with IBM Quantum Lab



- For dI/dt, **rotations** of neutrals driven by L-R beliefs$X|$ $\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}$; results in **limit cycles**
- Neutrals in **superpositions** -> ***i* (-> a state of dependency)** $H|0\rangle=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$, where $\begin{pmatrix}a\\b\end{pmatrix}$ -> $|a^2|=|b^2|=\frac{1}{2}$
- Future: Provided minimum available energy for least structural cost (SEP), a "bio-cell" at resonance maximizes work (MEP)

Conclusions: Risk Determination versus Risk Perceptions for Autonomous HMTs & Systems

- ***1st*: A perfect team at ground state => intelligence tunes a team's χ among its parts (not rational), low emotion, least structural entropy ~ team fitness (ordered, like a *bio-crystal*) -> MEP** (Martyushev, 2013), **like focusing a telescope** (Lawless, 2021) -> resonance in information flow ($\partial I/\partial t$)?
  - **Static perceptions cannot determine risks, dynamic skills, team fit, performance nor *interdependence;* only competition (debate) can identify perfect or dysfunctional teams, motivating hit-or-miss mergers & alliances**
- **2*rd:* Subadditivity of *i* => the best teams evolve, resilient to shock, but perfect teams are irreproducible => select team members by trial & error**
- **3th: Vulnerability theory => shocks harm teams with more redundancy**.
- **4th: Little generalizes from Science of individuals or H-H teams to science of A-HMTs; however, the converse may be true, that the mathematics of A-HMTs generalize to challenge risk perceptions & improve risk determinations**
- **5th: All interpretations are subjective, increasing the value of debates, "red teams" (Gen. Said, USAF, *drone attack in Afghanistan*) & *checks & balances***

Backup slides

25

25

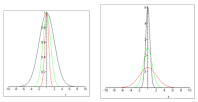
Questions & Future Research

- Is the greatest danger of risk perception -> emotional spasms (“group think,” fear)?
  - Is a major advantage of machines the skill to determine the emotional state of their human teammates? (Emotion identification ∝ trust; in Endsley, 2022, p. 38)
  - What actions can machines take to reduce the emotional states of human operators? (Red teams?)
- Do authoritarians & gangs control others by increasing independence among citizens to reduce factions?
- As the strength of free choice, does autonomy, when limited, drive $\partial I_{bistable} / \partial t$?
- Is it best to govern autonomous human-machine teams & systems with strict limits (Commanders Intent, rules of engagement, etc.)?
- How do emotional states melt boundaries (e.g., NATO’s closer realignment against Putin)?

26

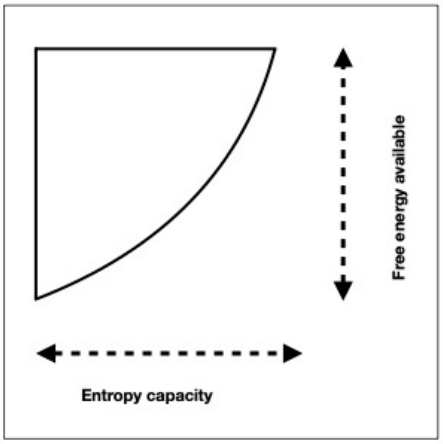
Open System Uncertainty, debate, structure, performance:

Color	Function	σ	Fourier Transform	σ	$\sigma\sigma$
Red	$f(t)=e^{-\frac{t^2}{2}}$	.33	$F(s)=\frac{\sqrt{2\pi}}{2}e^{-\frac{s^2}{2}}$	5.01	1.67
Green	$f(t)=e^{-\frac{t^2}{9}}$	.94	$F(s)=\sqrt{2\pi}e^{-\frac{s^2}{2}}$	2.51	2.36
Black	$f(t)=e^{-\frac{t^2}{18}}$	2.66	$F(s)=2\sqrt{2\pi}e^{-2s^2}$	1.25	3.33



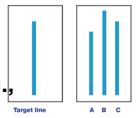
- **4th test: Metrics for complementarity**
  - **A new equation leads to several discoveries: $\mathbf{C \approx \Delta SEP * \Delta MEP}$**
    - **First, $\mathbf{tradeoffs}$: $\mathbf{\Delta SEP <-> \Delta MEP}$**
    - **Second, for maximum performance: $lim_{\Delta SEP->min} \Delta MEP \rightarrow max$**
    - **Third, $\Delta SEP_1 > \Delta SEP_2$ => a vulnerability in SEP_1 that decides competition**
      - **For example:**
        - **Enforced cooperation increases systemic vulnerability & the need to *steal* technology** (e.g., China, in Baker, 2015; Ratcliffe, 2020).
        - **Monopolies increase organizational vulnerability & the need to *steal* ideas from clients** (e.g., Amazon, in Mattioli, 2020, *WSJ*).
    - **Fourth, to enhance deception: Minimize $SEP: Team \rightarrow Unit \Rightarrow S = ln(Team) = 0$**
    - **Fifth, to counter Risk perceptions: Red Team challenges Blue Team decision: $\Delta K \rightarrow 0$**
      - Gen. Said's DoD report on the Aug. 29, 2021, ***drone attack in Afghanistan***; https://www.defense.gov/News/Transcripts/Transcript/Article/2832634/pentagon-press-secretary-john-f-kirby-and-air-force-lt-gen-sami-d-said-hold-a-p/

A notional diagram of ***free energy-entropy*** abstracted from Gibbs. We assume a team operates crudely like an "intelligent" fluid. This notional diagram underscores the importance of an organization's ability to collect and make available sufficient free energy "from an external source … [to maintain a] dissipative structure" (Prigogine, 1977) that offsets the waste and products produced by a team or organization.



Model-1: Facing uncertainty, humans seek consensus: 1. A minority can force conformity to a "rational" context (*even facing uncertainty or conflict*)**; or, 2. A majority** can debate the available information to build *K* and shared context.

1. Minority rules (consensus-seeking; socialism; ***easy to model***) **-> minority-control suppresses alternatives; + stability;** e.g., Asch's (1951) line study -> poor decisions **(e.g., destroying the Uighur culture in China)**



- *The requirement for consensus in the European Council often holds policy-making hostage to national interests in areas which Council could and should decide by a qualified majority.*" (WP, 2001, p. 29)
- e.g., the European Union's consensus-seeking in foreign affairs increases doubts => increased structural entropy production (SEP); (in Lawless, 2019b)
- **Monopolies** (e.g., Google, in *Forbes*, 2019); **Fraud** (e.g., VW emission scandal in 2015; in Boston, 2021, *WSJ*)

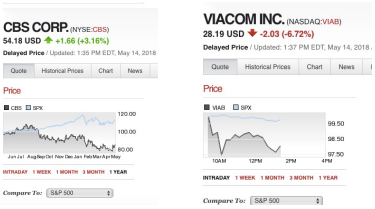
2. Majority rules (democratic debate builds a shared context; + instability; ***hard to model***); e.g., best legal decisions: "***informed assessment of competing interests***" (Justice Ginsburg, 2011)

- A U.S. DoD review of the August 29, 2021 Drone Strike in Afghanistan concluded that when facing uncertainty, a rigorous red teams test of that decision was imperative (Lawless & Sofge (2022), Risk determination versus risk perception: From misperceived drone attacks, hate speech and military nuclear wastes to human-machine autonomy. AAAI-2022).

Background: NSC-AI

- NSC-AI (2021). **Final Report:**
  - **… the rapidly improving ability of computer systems to solve problems … is world altering**
  - **The NSCAI Final Report recommends** [among several]:
    - Defend against emerging AI-enabled threats to America's free and open society. (p. 9)
      - Manage risks associated with AI-enabled and autonomous weapons. (p. 10)
      - [To] **Establish justified confidence in AI systems**
        - **Present a democratic model of AI use for national security** (p. 11).
          - ***As it happens, this is my area of research***

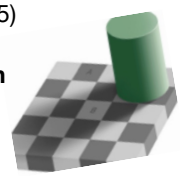
Endless fights are common, but who is right?

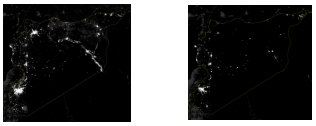


Long-running fights impact profits 2016-18

Non-factorable information: 2-sided factions & conflicts

Adelson's (2005) **illusion => misperception**





Endless civil-war disrupts social structures: Syria 2012 versus 2014

Non-factorable information

- **Perturbations of poorest teams generate too much Shannon information (e.g., divorce)**:
  - *H(x,y) ≥ H(x),H(y); however, interference from consequences & energy ignored*
  - Moreover, interdisciplinary science teams are the worst at team science (Cummings, 2015)
    - If $\sigma_{SEP}\sigma_{MEP} \sim C$, then as $\sigma_{SEP} \to \infty$, in the limit, $\sigma_{MEP} \to 0$
- **Perturbations of the best teams act as a unit => subadditivity ~ little information generated**:
  - **Von Neumann's Subadditivity: $S(\rho_{AB}) \leq S(\rho_A)+S(\rho_B)$** (e.g., in business, one train performing two functions; good marriage; the best teams)
    - If $\sigma_{SEP}\sigma_{MEP} \sim C$, then as $\sigma_{SEP} \to 0$, in the limit, $\sigma_{MEP} \to \infty$