The contagion effect of trust in multi-UAV systems

BLAKE PIPER
Blake.Piper@leidos.com

JANINE MATOR
Janine.D.Mator@leidos.com

leidos

Background: Trust in automation

- Defining trust in automation
  - "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" (Lee & See, 2004)
- Impact of trust on automation use
  - Under-trust and disuse
  - Over-trust and misuse
- Implications for human performance
  - Under-trust → Disuse → Higher workload, less attentional resources
  - Over-trust → Misuse → Low situation awareness, poor takeover response
- Common measures of trust in automation
  - Self-report
  - Behavioral measures
  - Physiological measures

Problem Space

- Prevailing assumption: Trust is unbiased by errors in other system components
- Seminal study by Keller and Rice (2009)
  - Measured trust in system components with varying levels of reliability
  - Found that performance errors from unreliable system components led to lower trust in reliable aids
- The contagion effect
- Implications for multi-UAV systems
  - Increasing demand for uncrewed aerial vehicles (UAVs) and larger swarms
  - Unknown influence of the contagion effect on pilots' decision-making
- Research questions: Does the contagion effect exist in experienced fighter pilots? If so, how does heterogeneity of UAV swarms impact the contagion effect in this sample?

Current Study: Overview

- Data collection ongoing
- Sample of *n* = 16 experienced fighter pilots recruited at Wright Patterson Air Force Base (WPAFB)
- Experimental design: 2 (reliability: reliable vs unreliable) x 2 (heterogeneity: same vs different UAV platform) repeated-measures
- All dependent measures:
  - Behavioral measure of trust (which UAV the pilot chooses to attack a ground target)
  - Self-reported trust
  - Cognitive workload
  - Perfect automation schema (Merritt et al., 2015)
  - Self-reported error severity
  - Flight experience

Current Study: Experimental task

leidos

- Air-to-ground attack mission with 4 UAVs in immersive fighter cockpit simulator
- 6 trials/scenarios: 2 reliable, 4 unreliable
  - Reliable condition: No errors
  - Unreliable condition: 1 of 4 UAVs experiences an error (i.e., lagging in flight formation) before auto-correcting
  - Homogenous condition: Only “SNAKE” or “LION” UAVs
  - Heterogenous condition: 2 “SNAKE” and 2 “LION” UAVs
- Pilots must report the error within 20s; otherwise, an Air Battle Manager informs them via headset
- At the end of each scenario, pilots select <u>one</u> UAV to perform an air-to-ground strike on a high-value ground target

F-35 with loyal wingmen UCAVs (Source: US Air Force)
Photo Credit: (https://www.youtube.com/watch?v=HPZpp_Y6Er8)

Current Study: Pre-study measures

leidos

- Flight hours
  - Total
  - 4th-gen (e.g., F-16)
  - 5th-gen (e.g., F-35)
- Perfect Automation Schema (Merritt et al., 2015)
  - 6 items on 7-point Likert scale
  - Sample items:
    - "Automated systems have 100% perfect performance"
    - "Automated systems make more mistakes than people realize"
    - "If an automated system makes an error, then it is broken."

Perfect Automation Schema (Merritt et al., 2015)

Instructions: Use the scale provided to select the best response to the item. (Options range from: "strongly disagree to strongly agree = 7)

	Strongly Disagree	Disagree	Somewhat Disagree	Neither Disagree nor Agree	Somewhat Agree	Agree	Strongly Agree
Automated systems have 100% perfect performance.	1	2	3	4	5	6	7
Automated systems can always be counted on to make accurate decisions.	1	2	3	4	5	6	7
Automated systems make more mistakes than people realize.	1	2	3	4	5	6	7
If an automated system makes an error, then it is broken.	1	2	3	4	5	6	7
If an automated system makes a mistake, then it is completely useless.	1	2	3	4	5	6	7
Only faulty automated systems provide imperfect results.	1	2	3	4	5	6	7

Current Study: Post-trial measures

leidos

- Error identification (i.e., "Did any UAVs experience an error in this trial? If so, please indicate which")
- 7-point Likert scales:
  - Subjective error severity: "How severe was the error noted above?"
  - Trust (Reliance Intentions Scale by Lyons & Guznov, 2019)
  - Cognitive workload
    - "How mentally demanding was the task using [UAV]?"

Reliance Intentions Scale (Lyons & Guznov, 2019)

Instructions: In the following items, consider your attitudes toward the UAVs in the video. Use the scale provided to select the best response to the item. (Scale: strongly disagree = 1, to strongly agree = 7)

	Strongly Disagree	Disagree	Somewhat Disagree	Neither Disagree nor Agree	Somewhat Agree	Agree	Strongly Agree
I would rely on [UAV1] without hesitation.	1	2	3	4	5	6	7
I think using [UAV1] will lead to positive outcomes.	1	2	3	4	5	6	7
I would feel comfortable relying on [UAV1] in the future.	1	2	3	4	5	6	7

Current Study: Post-study measures

- Qualitative interview questions:
  - After the UAV experienced an error, how did that impact your view of subsequent errors?
  - Would you trust a UAV from the same platform in a similar future mission? Why or why not?
  - Would you trust a UAV from a different platform in a similar future mission? Why or why not?
  - What behaviors exhibited by the UAV would decrease your level of trust in UAVs in general?
  - Are there any other thoughts you would like to add about your experience during the experiment?

Current Study: Practical value

- Expanding the parameters of contagion effect research to an experienced (i.e., non-undergraduate) sample
  - If detected: Informs the need for potential interventions, such as pilot training on trust biases and related interface design decisions
  - If undetected: Informs disparity in trust research
- Impact of heterogeneity on the contagion effect
  - Informs the need to distinguish/differentiate system components for appropriate automation use

References

- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*(1), 50-80.
- Lyons, J. B., & Guznov, S. Y. (2019). Individual differences in human–machine trust: A multi-study look at the perfect automation schema. *Theoretical Issues in Ergonomics Science*, *20*(4), 440-458.
- Merritt, S. M., Unnerstall, J. L., Lee, D., & Huber, K. (2015). Measuring individual differences in the perfect automation schema. *Human Factors*, *57*(5), 740-753.

Additional notes

- Component-specific trust (CST) theory: Proposes that operators will form varying levels of trust across automated aids, based on each one’s individual reliability. (Ideal, since operator trust wouldn't be biased by unrelated system performance.)
- **System-wide trust (SWT) theory: Proposes that operator trust generalizes across the system so that errors made by one imperfect aid will impact trust in other aids, even if those other aids ARE reliable.** (Supported by findings on the contagion effect.)