

DESIGNING TRUST IN AUTONOMOUS AERIAL MOBILITY

Ricerca compilativa e sviluppo di un framework UX per aumentare la comprensione, di trasparenza e di sicurezza percepita nei Personal Aerial Vehicles

INDICE

Abstract	5	3.4 Sintesi Critica: condizioni per la fiducia nell'UAM	74
Introduzione	7		
Metodologia della ricerca	8		
Cap.1 Il trasporto aereo civile	10	Cap 4 Il Safety-Transparency Trust Framework nei PAV	76
1.1 Breve storia del trasporto aereo civile personale	11	4.1 L'utente come luogo in cui la fiducia si manifesta	77
1.3 Livelli di autonomia e ruolo dell'Intelligenza Artificiale negli eVTOL	18	4.2 Tipologie di utenti come configurazioni psicologiche di calibrazione della fiducia	78
1.4 La sfida dell'adozione: oltre la sicurezza tecnica, la sicurezza percepita	19	4.2.2 utente ansioso–ipervigilante (The vulnerable Passenger)	78
		4.2.3 Il supervisore in cerca di controllo (Control-seeking Supervisor)	80
		4.2.4 Utente fiducioso–passivo (the delegating Optimist)	82
Cap. 2 Meccanismi psicologici della fiducia	22	4.3 Mappatura dei pattern ricorrenti	83
2.1 Definizione di fiducia nei sistemi socio-tecnici	23	4.4 Il Safety- transparency trust Framework come strumento progettuale	86
2.1.1 Sviluppo del processo di fiducia	25	4.4.1 Obiettivo del framework	88
2.2 Dimensioni della fiducia	27	4.4.2 Linee guida strategiche per il design UAM	91
2.2.1 Fiducia Cognitiva	28	Per superare l'atteggiamento di attesa prudente (“wait-and-see”)	92
2.2.2 Fiducia Affettiva	30	4.4.3 Assi progettuali	93
2.2.3 Fiducia Operativa	31	4.4.5 Architettura Logica del framework Integrato	95
2.2.4 Fiducia percepita vs. fiducia effettiva	32	4.5 Applicazione del framework nelle cinque fasi operative del volo	96
2.3 I modelli di fiducia nell'automazione	35		
2.3.1 Il modello di Lee e See sulla Trust in Automation	37	Conclusioni	99
2.3.2 Il modello di Parasuraman sui livelli di automazione e la fiducia	38	Bibliografia	104
2.3.3 Il modello di Hoffman sulla fiducia, trasparenza e calibrazione	39		
2.4 Calibrazione della fiducia, dall'eccessiva fiducia alla sfiducia	41	Allegati	108
2.5 Bias cognitivi e percezione del rischio nei PAV/eVTOL	45		
Cap 3 Sicurezza percepita, comfort e percezione del rischio	50		
3.1 Assenza del pilota umano e perdita dell'utente agile	51		
3.1.1 Evidenze sperimentali: Agency e vulnerabilità	52		
3.2.1 Evidenze sperimentali: Rischio situazionale e affidamento	53		
3.1.3 Evidenze sperimentali: Pilota come segnale di sicurezza psicologica	56		
3.2 Trasparenza, Explainable AI (XAI) e controllo percepito	56		
3.2.1 Evidenze sperimentali: Trasparenza negoziabile e HAT	57		
3.2.2 Evidenze sperimentali: Explainable AI e controllo cognitivo	61		
3.2.3 Evidenze sperimentali: Controllo percepito e calibrazione	63		
3.3 UX come mediatore tra sicurezza tecnica in sicurezza percepita	68		
3.3.1 UX come traduttore	69		
3.3.2 UX come regolatore	71		
3.3.3 UX come narratore	72		

INDICE IMMAGINI

Figura 1.1 – Fulton Airphibian (1946)	14
Figura 1.2 – Volocopter VCI	15
Figura 1.3 – Archer Midnight a Dubai	19
Figura 1.4 – Uber Elevate	19
Figura 2.1 – Modello della fiducia di Mayer et al., 1995.	26
Figura 2.2 – Un quadro concettuale dello sviluppo della fiducia	38
Figura 2.3 – Modello di Parasuraman	41
Figura 2.3 – La relazione tra calibrazione, eccesso di fiducia e sfiducia	46
Figura 3.1 – Configurazione della simulazione di guida	56
Figura 3.2 – Modello iniziale dell'interazione HAT	61
Figura 3.3 – Integrazione dell'HAT nella cabina di pilotaggio	62
Figura 3.4 – Bias Cognitivi nell'interazione dell'utente con veicoli automatizzati	67
Figura 3.5 – Gli otto scenari di guida automatizzata	68
Figura 3.5 – Panoramica della percezione soggettiva del comportamento del veicolo	69
Figura 4.1 Mappa concettuale del Safety-transparency Framework.	103

INDICE TABELLE

Tabella 1 – Leader di mercato eVTOL	16
Tabella 2 – Processo di sviluppo della fiducia	28
Tabella 3 – Fiducia interpersonale e fiducia nei sistemi Uomo-Macchina	31
Tabella 4 – Confronto sintetico della fiducia Affettiva.	32
Tabella 5 – Confronto sintetico della fiducia Operativa.	34
Tabella 6 – Confronto sintetico della fiducia effettiva.	36
Tabella 7 – Confronto sintetico della fiducia percepita.	37
Tabella 8 – Livelli di automazione della decisione e della selezione dell'azione.	41
Tabella 9 – Sintesi comparativa delle quattro tipologie di utenti nei PAV/eVTOL	87
Tabella 10 – Dalla teoria alla decisione progettuale.	92
Tabella 11 – Massimizzazione del trust e Calibrazione del Trust.	93
Tabella 12 – Mappa longitudinale della Fiducia sulla gestione del tempo.	98

ABSTRACT

L'emergere della Urban Air Mobility (UAM) e dei velivoli elettrici a decollo e atterraggio verticale (eVTOL) rappresenta una trasformazione radicale dei sistemi di trasporto urbano. Nonostante l'elevato livello di affidabilità tecnica raggiunto dai sistemi autonomi, l'adozione su larga scala rimane vincolata a un divario significativo tra sicurezza tecnica e sicurezza percepita. In assenza di un pilota umano a bordo, la fiducia nell'automazione diventa una variabile determinante per l'accettazione sociale della mobilità aerea autonoma. La presente ricerca analizza come la progettazione dell'esperienza utente (User Experience) possa supportare la costruzione e la calibrazione della fiducia nei Personal Aerial Vehicles (PAV) e nei Personal Aerial Transport Systems (PATS). Attraverso un approccio compilativo interdisciplinare che integra contributi provenienti dal Trust in Automation, dalla Human–AI Interaction, dalla psicologia cognitiva e dalla percezione del rischio, il lavoro identifica i meccanismi cognitivi ed emotivi che influenzano il rapporto tra passeggero e sistema autonomo. L'esito della ricerca è il Safety–Transparency Trust Framework, un modello concettuale che descrive come la trasparenza dell'intelligenza artificiale possa tradurre la sicurezza tecnica in sicurezza percepita. Il framework propone tre livelli di trasparenza mediati dalla UX e definisce principi progettuali volti a prevenire fenomeni di eccessiva fiducia e sfiducia, promuovendo un affidamento appropriato. La tesi contribuisce a ridefinire la fiducia come variabile progettuale centrale nei sistemi autonomi safety-critical, offrendo un quadro interpretativo applicabile alla progettazione della mobilità aerea urbana del futuro.

Keywords: Trust in Automation, User Experience (UX), Aviation, Personal Transportation, Artificial Intelligent (AI).

ACRONOMI

ADS	Automated Driving Systems
AV	Autonomous Driving
EASA	European Union Aviation Safety Agency
EIS	Entry Into Service
EFIS	Synthetic Vision Electronic Flight Instrument Systems
eVTOL	Electric Vertical Take-Off and Landing
HAI	Human–AI Interaction
HAT	Human-Autonomy Teaming
NDRT	Non-Driving Related Tasks
PATS	Personal Aerial Transport Systems
PAV	Personal Air Vehicle
PeAV	Personal Electric Aerial Vehicle
SAT	Situation Awareness-based Agent Transparency
UAM	Urban Air Mobility
UAV	Urban Air Vehicle
UX	User Experience (Design dell'interazione)
VTOL	Vertical Take-Off and Landing
XAI	Explainable AI

INTRODUZIONE

Negli ultimi anni, il settore dell'Urban Air Mobility (UAM) si sta affermando come una delle evoluzioni più significative della mobilità contemporanea (Arafat et al.; 2023). L'introduzione di velivoli elettrici a decollo e atterraggio verticale (eVTOL) e dei Personal Aerial Vehicles/Transport Systems (PAV/PATS) promette di trasformare radicalmente il modo in cui le persone si sposteranno all'interno delle città, offrendo soluzioni di trasporto rapide, sostenibili e pienamente integrate con l'infrastruttura urbana (EASE, 2021; UAM, 2021; MyCoper; 2011). Parallelamente, i sistemi di guida autonoma stanno evolvendo rapidamente, consentendo ai velivoli di operare con un intervento umano sempre più ridotto, fino a raggiungere la possibilità di voli completamente autonomi di Livello 5 (Lovreglio, 2023; Yang e Kim, 2024). Questa transizione rappresenta una svolta nella relazione tra user e mezzo di trasporto: il passeggero potrebbe non interagire più con un pilota umano, ma con un sistema automatizzato basato su intelligenza artificiale, ridefinendo il concetto stesso di controllo ed esperienza di viaggio (Yang e Kim, 2024; Wells, 2021). Nonostante i progressi tecnologici e gli elevati standard di sicurezza tecnica (garantiti da sistemi ad alta integrità come la navigazione GNSS/IMU e la percezione basata su sensori LiDAR e Vision-Based), l'accettazione pubblica dei sistemi autonomi rimane un ostacolo centrale (Wang et al., 2022). La letteratura evidenzia un divario significativo tra sicurezza reale, determinata dalle performance tecniche del sistema, dell'affidabilità meccanica (Actual Trustworthiness) e sicurezza percepita, fortemente influenzata da fattori cognitivi, emotivi e culturali (Hammer et al., 2025; Lovreglio, 2023). Nel caso dei PAV autonomi, questo divario risulta ancora più marcato: l'assenza del pilota, l'opacità delle decisioni dell'AI, la difficoltà nell'interpretare le azioni del veicolo e la scarsa familiarità con l'esperienza del volo autonomo contribuiscono alla formazione di ansia e diffidenza (UAM; Allianz, 2025). A ciò si aggiungono preoccupazioni di natura socio culturale, come la paura dell'imprevisto, la sfiducia nei sistemi automatizzati e la percezione di perdita di controllo. In questo contesto, la costruzione della fiducia (Trust in Automation) assume un ruolo cruciale per l'adozione su larga

scala dei sistemi UAM. Il tema della fiducia rappresenta oggi un nodo centrale nel dibattito interdisciplinare tra Digital Interaction Design, Human-AI Interaction (HAI), psicologia cognitiva e ingegneria aeronautica. Con l'avvento della mobilità aerea autonoma, il design è chiamato a svolgere un ruolo determinante: rendere comprensibile e rassicurante la complessità tecnica dell'IA attraverso l'implementazione della Explainable AI (XAI). La rilevanza scientifica di questo lavoro risiede in due aspetti principali: Comprendere i meccanismi psicologici e i bias che regolano la formazione della fiducia nei sistemi di volo autonomo.

Identificare linee guida e strumenti progettuali (UX come traduttore, regolatore e narratore) che contribuiscono a colmare il divario tra sicurezza reale e percepita, \ L'obiettivo finale è promuovere una Appropriate Reliance (affidamento appropriato), in cui l'utente non si affida ciecamente alla macchina, ma sia in grado di calibrare il proprio intervento sulla base delle reali capacità e dei limiti operativi del sistema (Operational Design Domain).

METODOLOGIA

La ricerca adotta un approccio compilativo a orientamento progettuale (design-oriented conceptual research), finalizzato alla costruzione di un quadro teorico per il Design for Trust nei sistemi di Urban Air Mobility (UAM). La fiducia è interpretata non solo come costrutto psicologico o proprietà tecnica del sistema, ma come esito progettuale mediato dall'esperienza utente e dalla trasparenza dell'automazione.

Lo studio non prevede indagini empiriche originali, ma si fonda su una sintesi critica interdisciplinare della letteratura. I contributi analizzati provengono principalmente dagli ambiti del Trust in Automation, della psicologia cognitiva e percezione del rischio, della Human-AI Interaction, delle Explainable Artificial Intelligence e degli studi sull'aviazione e sulla mobilità autonoma terrestre. Le fonti sono state selezionate in base alla

loro rilevanza rispetto ai processi di formazione e calibrazione della fiducia, alla solidità teorico-empirica e alla trasferibilità al contesto UAM. Il percorso analitico si sviluppa in tre fasi: analisi del contesto socio-tecnico della UAM e dell'assenza del pilota umano; integrazione dei modelli teorici sulla fiducia e sulla sua calibrazione (eccessiva fiducia e scarsa fiducia); traduzione dei costrutti teorici in principi progettuali orientati alla User Experience.

L'esito è una sintesi concettuale prescrittiva, volta a supportare la progettazione di sistemi autonomi safety-critical, pur riconoscendo il limite dell'assenza di validazione empirica diretta.

IL TRASPORTO AEREO CIVILE

Breve storia del trasporto aereo civile

Il concetto di Urban Air Mobility (UAM) rappresenta l'evoluzione contemporanea di un paradigma che ha affascinato ricercatori e industria per diversi decenni, proponendosi come una possibile risposta ai problemi di congestione, inefficienza e saturazione dei sistemi di trasporto terrestre tradizionali (Clarke, 2018; Wells, 2021). L'idea delle macchine volanti ha da sempre alimentato l'immaginario collettivo, incarnando aspirazioni di progresso tecnologico, libertà di movimento ed esplorazione. Questo immaginario è stato ampiamente rafforzato dall'industria dell'intrattenimento, che ha rappresentato il volo personale in numerose opere di fantasia, da Chitty Chitty Bang Bang a The Jetsons, fino a Star Wars e Back to the Future, contribuendo a trasformare un'idea visionaria in un obiettivo tecnologico condiviso (Clarke, 2018). Il sogno del volo personale e degli air taxi affonda le sue radici nelle visioni del XX secolo, miranti a creare un sistema di trasporto aereo punto a punto che superasse i limiti delle infrastrutture stradali, rendendo il mezzo "volante" accessibile come un'auto (Wells, 2021) [Fig.1.1].

A partire dalla seconda metà del Novecento, numerosi enti di ricerca aeronautica, tra cui la National Aeronautics and Space Administration (NASA), hanno iniziato a esplorare concetti di Personal Air Vehicle (PAV) e di trasporto aereo punto a punto, con l'obiettivo di superare le limitazioni operative dell'aviazione generale e ridurre la congestione dei sistemi di trasporto terrestre. In particolare, i programmi NASA dedicati al General Aviation Revitalization (GAR) e successivamente al Small Aircraft Transportation System (SATS) hanno individuato come principali barriere allo

sviluppo del volo personale la complessità del pilotaggio, la sicurezza operativa e l'integrazione nello spazio aereo controllato (NASA, 2005; NASA, 2014). Nel corso degli anni 2000, l'attenzione della ricerca si è progressivamente spostata verso l'automazione avanzata e l'assistenza al pilota come strumenti chiave per ampliare l'accessibilità del volo civile personale. In questo contesto, progetti europei finanziati dall'Unione Europea, come myCopter¹, hanno analizzato l'interazione uomo-macchina e dimostrato come l'elevato carico cognitivo richiesto dal

pilotaggio tradizionale rappresentasse uno dei principali ostacoli alla diffusione dei veicoli aerei personali presso utenti non professionisti (Jump et al., 2012).

Sotto il profilo sociologico, il passaggio dal XX al XXI secolo ha segnato una transizione fondamentale nella base della fiducia: dalla conoscenza personale del pilota o del fornitore alla conoscenza istituzionale basata su certificazioni e sistemi di regolamentazione rigorosi. McKnight e Chervany osservano che, storicamente, l'accettazione

Figura 1.1: Fulton Airphibian (1946), primo esempio certificato di veicolo stradale convertibile in aeromobile. Fonte: Archivio storico Fulton Airphibian.



Figura 1.2: Prototipo iniziale di velivolo eVTOL multicottero monoposto (Volocopter VCI), dimostratore tecnologico per Urban Air Mobility.

pubblica di tecnologie “sconosciute” come il volo personale è stata resa possibile solo grazie all'emergere di enti governativi che hanno garantito la competenza e l'integrità del sistema (UAM, 2022). Sebbene i primi tentativi di integrazione nell'area urbana siano stati dominati dagli elicotteri, questi hanno incontrato limitazioni dovute a costi proibitivi, inquinamento acustico e complessità operativa, favorendo la predominanza degli aerei di linea commerciali (Wells, 2021). Solo con l'avvento dei sistemi a propulsione elettrica (eVTOL) e delle tecnologie autonome, tra la fine del XX e l'inizio del XXI secolo, il paradigma si è orientato verso il concetto di Smart City, in cui il volo personale diventa una componente essenziale di una mobilità fluida e sostenibile. La storia del trasporto aereo personale riflette un continuo sforzo per superare la complessità operativa. Come sottolineato da Raivi e Moh, l'innovazione tecnologica ha agito da “ponte” per trasformare ciò che un tempo era considerato “fantascienza” in una realtà imminente per le città europee. Questo percorso ha trovato il suo primo vero punto di svolta commerciale con il volo del primo proof-of-concept di un eVTOL con equipaggio nel 2011 [Fig. 1.2], segnando l'inizio della transizione dai sogni del XX secolo alla mobilità aerea avanzata contemporanea.

1. myCopter è stato un progetto di ricerca europeo (finanziato dal 7° Programma Quadro dell'Unione Europea) conclusosi a dicembre 2014, il cui obiettivo principale era gettare le basi per un sistema di trasporto aereo personale rivolto al grande pubblico.

1.2 URBAN AIR MOBILITY: DEFINIZIONE, ATTORI, SCENARI APPLICATIVI

Urban Air Mobility (UAM) è un sistema di trasporto aereo urbano basato su velivoli elettrici a decollo e atterraggio verticale (eVTOL), destinati al trasporto di persone e merci, anche in modalità autonoma e su tratte dirette. L'UAM viene generalmente descritta come una rete dinamica di tipo "hub-and-spoke", nella quale rotte, nodi e frequenze operative si adattano in tempo reale alla domanda di mobilità nelle aree metropolitane (Clarke, 2018; UAM, 2022; NASA, 2018). Ciò risolve il problema delle infrastrutture terrestri, che soffrono di una "rigidità di percorso" e di costi di espansione proibitivi, che crescono proporzionalmente alla lunghezza del tragitto fisico costruito (Clarke, 2018).

Secondo la definizione adottata da NASA ed EASA, l'UAM è un sistema di operazioni di traffico aereo sicuro ed efficiente in ambiente urbano, che include velivoli con e senza pilota e richiede nuovi paradigmi di gestione dello spazio aereo (UAM, 2022; NASA, 2018). Questa definizione evidenzia come la complessità dell'UAM non risieda esclusivamente nel velivolo, ma nell'interazione tra tecnologia, regolazione e accettazione sociale.

Gli attori principali del sistema includono i produttori di velivoli eVTOL (electric Vertical Take-off and Landing), i gestori delle infrastrutture di terra (Vertiporti) e gli operatori di rete e le autorità di regolazione come EASA e FAA (Clarke, 2018; UAM, 2022). Nel 2025 il settore dell'Urban Air Mobility (UAM) si trova in una fase di transizione critica, collocata tra sviluppo tecnologico avanzato, processi di certificazione ancora in corso e l'avvio delle prime operazioni commerciali limitate. Negli Stati Uniti e in Europa, aziende come Joby Aviation, Archer Aviation [Fig. 2.1], Volocopter stanno seguendo percorsi certificati particolarmente rigorosi, guidati rispettivamente da FAA ed EASA, adottando strategie di introduzione graduale che prevedono inizialmente la presenza di un pilota a bordo o una supervisione umana, anche quando l'autonomia tecnica sarebbe già parzialmente

[Tabella 1: Leader di mercato eVTOL: sintesi delle specifiche tecniche e delle previsioni di ingresso in servizio \(EIS\) per i principali modelli a pilotaggio assistito e autonomo, stato 2025.](#)

Azienda	Modello	Paese	Configurazione	Posti passeggeri	Pilota a bordo	Livello di autonomia
Joby Aviation	Joby S4	USA	Lift+Cruise	4 pax	Si	Autonomia assistita (L3-L4)
Archer Aviation	Midnight	USA	Lift+Cruise	4 pax	Si	Autonomia assistita (L3-L4)
Volocopter	Volocity	Germania	Multirotono (Wingless)	1 pax	Si	Autonomia assistita (L3)
Wisk Aero	Gen 6	USA	Lift+Cruise	4 pax	No (Remote)	Autonomia completa (L4-L5)
Ehang	EH216-S	Cina	Multirotono (Wingless)	2 pax	No	Autonomia completa (L4)

disponibile. Questa scelta non è esclusivamente tecnica, ma risponde a esigenze di sicurezza percepita e di costruzione della fiducia pubblica, anticipando le problematiche che verranno approfondite nei paragrafi successivi. Negli Stati Uniti, Joby e Archer rappresentano i casi più avanzati, con piani di lancio commerciale tra il 2025 e il 2026, che puntano a servizi iniziali di collegamento aeroportuale, integrandosi con operatori aerei tradizionali. In Europa, Volocopter mira a diventare il primo operatore eVTOL certificato da EASA, mentre Lilium persegue un modello di mobilità aerea regionale a maggiore autonomia, accettando una tempistica di ingresso più lunga per garantire standard di sicurezza equivalenti all'aviazione commerciale. Parallelamente, Wisk Aero, sostenuta da Boeing, sta sviluppando un modello radicalmente diverso basato su operazioni completamente autonome, la cui entrata in servizio è prevista solo nel lungo periodo (Frackiewicz, 2025). In contrasto con l'approccio occidentale, la Cina ha già avviato operazioni commerciali limitate grazie a EHang, il cui eVTOL autonomo EH216-S ha ottenuto la certificazione completa dalla CAAC. Tuttavia, tale certificazione rimane valida esclusivamente a livello nazionale, evidenziando come l'UAM si stia sviluppando secondo modelli geopolitici eterogenei, con differenti equilibri tra rapidità di implementazione, rigore normativo e accettazione sociale (Frackiewicz, 2025; UAM, 2022).

Oltre ai produttori di hardware, il settore è fortemente influenzato da attori che definiscono i modelli operativi e l'infrastruttura di rete. In ambito network e ridesharing, Uber, attraverso l'iniziativa Elevate (la cui divisione operativa è stata acquisita da Joby Aviation), ha agito come catalizzatore per l'intero settore, definendo le specifiche per le "Skyports" e i requisiti di missione per gli OEM partner (Clarke, 2018) [Fig.2.2]. Per quanto riguarda la gestione delle infrastrutture, aziende come Skyports e Hillwood collaborano con i produttori per progettare i vertiporti, nodi critici per l'atterraggio, il decollo e la ricarica dei velivoli (Clarke, 2018). Infine, enti di ricerca e regolazione come NASA, FAA ed EASA, pur non essendo attori commerciali, risultano determinanti, poiché stabiliscono gli standard di sicurezza, navigazione e integrazione dello spazio

aereo attraverso programmi come l'UTM (UAS Traffic Management) e il Grand Challenge (Clarke, 2018). I velivoli eVTOL si dividono in tre archetipi tecnologici principali, comunemente riconosciuti nella letteratura. Nella configurazione Vectored Thrust (Spinta vettoriale), i motori ruotano per fornire spinta sia verticale sia orizzontale. Si tratta della soluzione tecnologicamente più complessa, a causa della

necessità di componenti mobili per il direzionamento della spinta, ma anche la più efficiente per i voli a lunga distanza (UAM, 2022). Questa architettura è adottata da aziende come Joby, Bell e Hyundai. La configurazione Lift + Cruise (Sollevamento e crociera separati) dispone di rotori dedicati al volo verticale e di motori separati per la spinta orizzontale, offrendo un compromesso tra semplicità costruttiva e prestazioni operative (UAM, 2022). Si ritiene tuttavia che questa architettura possa essere più agevole da certificare rispetto alla spinta vettoriale, grazie alla separazione netta dei sistemi propulsivi (UAM, 2022). È adottata da aziende come Wisk, EVE (Embraer) e BETA. La configurazione Wingless (Multicotteri) rappresenta la soluzione strutturalmente più semplice, priva di ali fisse e basata interamente sulla spinta costante dei rotori (UAM, 2022). L'assenza di parti mobili per la transizione di volo

garantisce un'elevata ridondanza e facilità di controllo, rendendo questi velivoli particolarmente adatti a brevi distanze urbane (UAM, 2022). È sviluppata da aziende come Airbus, Volocopter ed E-Hang.

L'evoluzione tecnologica dei velivoli eVTOL è strettamente interconnessa con gli scenari applicativi prioritari. Il primo scenario riguarda il trasporto passeggeri (Air Taxi), destinato principalmente ai collegamenti punto-punto all'interno delle aree metropolitane o tra centri urbani e aeroporti. In questo ambito, si prevede un utilizzo predominante di velivoli vectored thrust, in grado di garantire velocità e comfort, accompagnato da una strategia di adozione incrementale che prevede inizialmente la presenza di un pilota a bordo per favorire l'accettazione pubblica (UAM, 2022; Clarke, 2018). Il secondo scenario riguarda le consegne dell'ultimo miglio, focalizzate sul trasporto di merci tra magazzini e hub di distribuzione urbana. Per queste operazioni, si privilegiano velivoli lift + cruise o multicotteri, progettati per operare in modalità autonoma fin dalle prime fasi del servizio. Infine, i servizi di emergenza medica comprendono il trasporto rapido di personale sanitario, attrezzature, organi o pazienti e godono di una maggiore accettazione pubblica per il loro elevato valore sociale (NASA, 2018; EASA, 2022).

Figura 1.3: Archer Midnight a Dubai: esempio di tecnologia eVTOL "proof-of-concept" evoluta per il mercato commerciale.



FIGURA 1.3



Figura 1.4: Rendering del vertiporto "The Paw" progettato dallo studio Gannett Fleming per il programma Uber Elevate. Il concept illustra l'infrastruttura necessaria per supportare operazioni UAM (Urban Air Mobility) ad alta frequenza, con piazzole di ricarica elettrica integrate per velivoli eVTOL.

FIGURA 1.4

1.3 LIVELLI DI AUTONOMIA E RUOLO DELL'INTELLIGENZA ARTIFICIALE NEGLI EVTOL

Nell'ambito dell'Urban Air Mobility, il concetto di autonomia assume una rilevanza centrale, rappresentando il principale driver delle prestazioni tecnologiche dei nuovi velivoli. Per definire il grado di sofisticazione tecnologica implementato a bordo del velivolo, le norme SAE J3016 [Appendice A], sviluppate per l'automotive, non sono adatte all'aeronautica, in quanto non considerano requisiti di integrità della navigazione, di certificazione e di sicurezza richiesti per le operazioni civili (Gerbeth et al., 2023).

Il quadro normativo europeo per UAM ed eVTOL si fonda sulle regolamentazioni dell'European Union Aviation Safety Agency (EASA) e delle autorità nazionali, includendo certificazioni come il Type Certificate e il Certificate of Airworthiness, integrate da regole operative per gli UAS (Reg. UE 2019/947), normative di aeronavigabilità e di mantenimento (Reg. UE 2024/1107) e requisiti per operatori, piloti e remote pilot (Routaboul, 2025). Ad oggi, in ambito europeo, alcune aziende hanno ottenuto la Production Organization Approval (POA), che certifica l'organizzazione di produzione ma non il velivolo né le sue capacità operative.

Tra queste, Embention ha ottenuto la POA per autopiloti e avionica dedicati a UAS ed eVTOL sotto AESA, mentre Volocopter detiene una POA (oltre alla DOA) per la produzione del proprio eVTOL (VoloCity). I processi di Type Certification completi, necessari per consentire operazioni commerciali su larga scala, sono tuttora in corso e non è stato ancora rilasciato un Type Certificate europeo completo per un eVTOL (Embention, 2025; Kadel, 2025). A titolo di confronto, in Cina un eVTOL ha già ottenuto una certificazione locale dalla CAAC: il modello EH216-S di EHang è autorizzato per operazioni pubbliche, ma tale certificazione non è riconosciuta da EASA né dalla FAA, rimanendo un acceleratore regionale piuttosto che uno standard internazionale (Frąckiewicz, 2025). All'interno di questo quadro normativo e operativo, l'integrazione dell'Intelligenza Artificiale (IA) ha l'obiettivo di creare un sistema intrinsecamente più "sicuro" rispetto al trasporto terrestre, in grado di operare in scenari caratterizzati da elevata densità di traffico e da complessità decisionale. In questo contesto, l'IA non rappresenta una metafora del "cervello del sistema", ma uno strumento necessario per superare i limiti fisiologici e cognitivi dell'operatore umano (Clarke, 2018; Shizhuang, 2022). Attraverso machine learning, deep learning e computer vision, l'IA ottimizza il path planning,

la collision avoidance e la gestione del traffico aereo, riducendo il rischio di incidenti fino a livelli dell'ordine di 10^9 eventi per ora di volo, in linea con gli obiettivi di sicurezza aeronautica (Tech4Future, 2021). Un ambito critico di applicazione riguarda la sensor fusion, necessaria per identificare ostacoli, altri velivoli e aree di atterraggio sicure (verti porti), mediante reti neurali convoluzionali (CNN) e sensori ridondanti (GNSS, IMU, LiDAR). Un esempio applicativo è il progetto della start-up The Edge Company, in cui la co-fondatrice Francesca Dalla Venezia ha sviluppato BCMS Ventur, un sistema di rilevamento degli ostacoli basato su deep learning e reti neurali auto addestrate, progettato per mitigare fenomeni come il bird-strike. Considerata l'elevata densità di traffico prevista, l'IA sarà inoltre chiamata a gestire il traffico aereo urbano (UTM) in modo decentralizzato, consentendo una pianificazione precisa delle traiettorie tramite lo scambio di dati Vehicle-to-Everything (Shizhuang, 2022).

Infine, grazie alle Simplified Vehicle Operations (SVO), l'automazione consente di ridurre significativamente i requisiti di addestramento, gestendo le dinamiche di volo più complesse e consentendo agli operatori con competenze limitate di operare in sicurezza. Tuttavia, questo aumento dell'autonomia tecnica non si traduce automaticamente in maggiore fiducia pubblica, un tema che sarà approfondito in 1.4 con riferimento all'under-trust e alla percezione sociale (Zhengming & Renran, 2022; Hammer et al., 2025). Sebbene l'IA riduca statisticamente l'errore umano, il pubblico manifesta frequentemente un control bias e una diffidenza nei confronti del "pilota algoritmico". Per i Personal Air Vehicles di Livello 4, la principale barriera all'adozione rimane pertanto l'undertrust, causato dalla limitata trasparenza dei processi decisionali della macchina

1.4 LA SFIDA DELL'ADOZIONE: OLTRE LA SICUREZZA TECNICA, LA SICUREZZA PERCEPITA

L'accettazione pubblica della Mobilità Aerea Urbana (UAM) non è determinata esclusivamente da parametri ingegneristici, ma emerge da un ecosistema articolato di fattori sociali che includono dimensioni psicologiche, ambientali e culturali. Secondo

L'accettazione pubblica della Mobilità Aerea Urbana (UAM) non è determinata esclusivamente da parametri ingegneristici, ma emerge da un ecosistema articolato di fattori sociali che includono dimensioni psicologiche, ambientali e culturali. Secondo i dati raccolti dall'EASA e dalle analisi di settore, il successo dell'UAM dipende dalla capacità di colmare la dicotomia tra sicurezza tecnica e sicurezza percepita, superando barriere percettive profondamente radicate nella società civile (EASA, 2021). In questo contesto, John-Paul Clarke, facendo riferimento al rapporto del 2014 del National Research Council (NRC) intitolato "Autonomy Research for Civil Aviation—Toward a New Era of Flight", ha identificato "8 barriere tecnologiche, quattro barriere normative e di certificazione e due barriere legali e sociali all'aumento dell'autonomia"² che ostacolano l'integrazione dell'autonomia nell'aviazione civile e, di conseguenza, l'adozione dell'Urban Air Mobility (UAM). Tali barriere evidenziano come il problema non sia solo ingegneristico, ma profondamente relazionale tra uomo e tecnologia (NASA, 2018 ; UAM, 2022; Clarke, 2018). L'analisi delle barriere parte da criticità infrastrutturali quali la gestione dello spettro e delle comunicazioni, necessarie per garantire link di comando o dati di monitoraggio

ultra-affidabili (Barriera 1), e prosegue con le imperative necessità di una cybersicurezza "end-to-end" che protegga il velivolo da intrusioni malevole (Barriera 2). A queste si aggiungono le difficoltà di integrazione con il traffico aereo esistente, dove i sistemi autonomi devono convivere con mezzi pilotati da esseri umani in uno spazio aereo congestionato (Barriera 3), e la questione della resilienza del sistema, ovvero la capacità di mantenere la sicurezza anche in presenza di guasti parziali (Barriera 7). Clarke evidenzia inoltre la barriera della verifica e validazione (V&V) (Barriera 8), sottolineando come gli attuali standard di certificazione siano strutturalmente inadeguati per software complessi che operano in ambienti dinamici (Clarke, 2018). L'analisi delle barriere parte da criticità infrastrutturali quali la gestione dello spettro e delle comunicazioni, necessarie per garantire link di comando o di dati di monitoraggio ultra-affidabili (Barriera 1), e prosegue con le imperative necessità di una cybersicurezza "end-to-end" che protegga il velivolo da intrusioni malevole (Barriera 2). A queste si aggiungono le difficoltà di integrazione con il traffico aereo esistente, in cui i sistemi autonomi devono convivere con mezzi pilotati da esseri umani in uno spazio aereo congestionato (Barriera 3), e la questione della resilienza del sistema, ovvero la capacità

di mantenere la sicurezza anche in presenza di guasti parziali (Barriera 7). Clarke evidenzia inoltre la barriera della verifica e validazione (V&V) (Barriera 8), sottolineando come gli attuali standard di certificazione siano strutturalmente inadeguati per software complessi che operano in ambienti dinamici (Clarke, 2018). Tuttavia, il vero spartiacque tra successo commerciale e fallimento sociale dell'UAM risiede nel triangolo formato dalle barriere relative alla collaborazione uomo-macchina (Barriera 4), alla natura dei sistemi adattivi e non-deterministici (Barriera 5) e alla percezione e cognizione della macchina (Barriera 6). La sfida del Human-Machine Teaming non riguarda esclusivamente il controllo del mezzo, ma la costruzione di un linguaggio condiviso, affinché l'adozione avvenga: la macchina deve essere in grado di comunicare il proprio stato e le proprie intenzioni, evitando che l'utente si percepisca come un elemento estraneo al processo decisionale e prevenendo quel senso di impotenza che genera sfiducia. Questo problema è intrinsecamente legato alla natura dei sistemi adattivi e non-deterministici (Barriera 5); la capacità dell'intelligenza artificiale di apprendere ed evolvere rende infatti difficile per l'utente umano prevedere con certezza il comportamento del veicolo. Se il passeggero percepisce il funzionamento dell'IA come una "scatola nera" imperscrutabile, la disponibilità a delegare la propria sicurezza crolla drasticamente. A complicare ulteriormente il quadro interviene la barriera della percezione e della cognizione della macchina (Barriera 6): la capacità del veicolo di interpretare correttamente l'ambiente circostante, distinguendo tra un ostacolo reale e un falso allarme, ha riflessi psicologici immediati, poiché l'utente necessita di conferme visibili che il sistema "veda" e "comprenda" il mondo in modo pari o superiore a un pilota umano (Clarke, 2018). Quando queste barriere tecniche si intrecciano con le preoccupazioni sociali legate alla privacy (ovvero il timore di una sorveglianza costante tramite i sensori di bordo), alla sicurezza intesa come integrità fisica – sia come passeggero sia come pedone per il rischio di caduta del velivolo – emergono ulteriori resistenze. A queste si aggiungono le paure legate all'impatto acustico per i residenti e alla non familiarità del suono prodotto dagli eVTOL rispetto a mezzi tradizionali come autobus o motocicli. Sebbene la propulsione elettrica sia generalmente percepita come positiva

² **2. Committee on Science, Space, and Technology.** (2018, July 24). *Urban Air Mobility—Are Flying Cars Preparing for Takeoff?* (Hearing before the Committee on Science, Space, and Technology, House of Representatives, 115th Congress). U.S. Government Publishing Office.

MECCANISMI PSICOLOGICI DELLA FIDUCIA

Definizione di fiducia nei sistemi socio-tecnici personale

Numerosi studiosi appartenenti a diverse discipline hanno analizzato le cause, la natura e gli effetti della fiducia, evidenziando come il concetto di trust assuma significati differenti a seconda delle prospettive teoriche e dei contesti applicativi. Per approfondire ciò che la fiducia comporta all'interno dei sistemi interattivi, è utile partire dal modello di riferimento proposto da Mayer, Davis e Schoorman, considerato la matrice teorica archetipo in questo ambito per relazioni interpersonali. Gli autori definiscono la fiducia come. "The willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the truster, irrespective of the ability to monitor or control that other party." (Mayer et al., 1995). Tale definizione presuppone l'esistenza di un altro attore identificabile, rendendo la fiducia una relazione interpersonale tra trustee e truster. L'elemento centrale è la vulnerabilità, intesa come disponibilità ad assumere un rischio e quindi ad accettare la possibilità di una perdita. In questo senso, la fiducia non coincide con l'azione rischiosa in sé, ma con la disposizione ad esporsi al rischio, come sottolineano gli autori: "Trust is not taking risk per se, but rather it is a willingness to take risk, the stage to make oneself vulnerable." (Mayer et al., 1995)

3. Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. Academy of Management Review, 20(3), p. 712.

Per definire correttamente la fiducia è inoltre necessario distinguerla da concetti affini che possono offuscare la natura della definizione, quali cooperation, confidence e predictability. La cooperazione, secondo Mayer, non richiede necessariamente fiducia, poiché "you can cooperate with someone who you don't really trust. If there are external

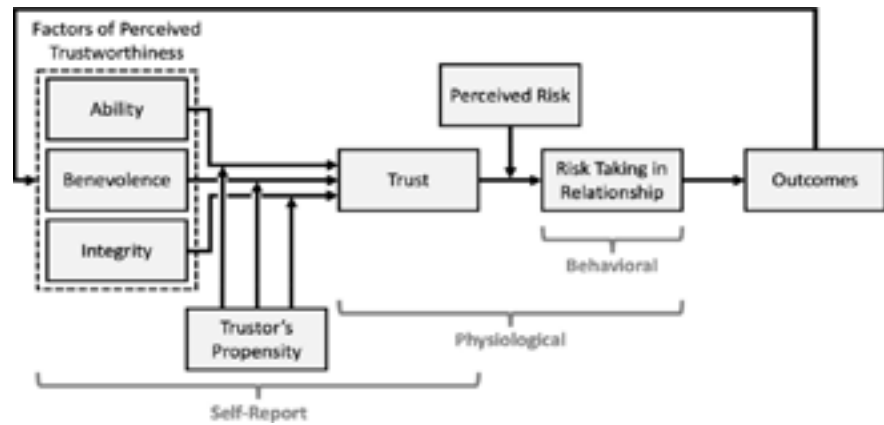


Figura 2.1: Modello della fiducia di Mayer et al., 1995.

control mechanisms that will punish the trustee for deceitful behavior” (Mayer et al., 1995): in presenza di meccanismi di controllo, infatti, la vulnerabilità è ridotta o assente. Anche la confidenza si distingue dalla fiducia: come afferma Luhmann, “... (Every morning you leave the house without a weapon!) you are in a situation of confidence...” (Mayer et al., 1995), poiché la confidenza non implica il riconoscimento consapevole del rischio né la valutazione delle alternative, elementi invece imprescindibili nella trust. Infine, la predictability, definita da Gabarro come “The extent to which one person can expect predictability in the other’s behavior in terms of what is ‘normally’ expected of a person acting in good faith”, contribuisce a ridurre l’incertezza e può favorire la cooperazione, ma non comporta la disponibilità ad assumersi un rischio né la scelta di esporsi alla vulnerabilità. Seguendo la proposta del modello della fiducia di

Mayer, Davis e Schoorman, andiamo a indagare le tre caratteristiche (appearing often in the literature) del fiduciario che costruiscono il terreno fertile per far sviluppare la fiducia. Queste caratteristiche sono importanti perché vanno a far capire ai ricercatori perché alcuni “parties” sono molto più affidabili di altri. Abbiamo “tre livelli generali che riassumono le basi della fiducia: competenza (ability), integrità (integrity) e benevolenza (benevolence)” (Mayer et al., 1995). La competenza è l’insieme di abilità, capacità e caratteristiche che consentono al fiduciario (trustee) di esercitare un’influenza efficace in uno specifico dominio. L’integrità è il grado in cui il fiduciario aderisce a un insieme di principi che il soggetto che accorda fiducia (trustor) considera accettabili. La benevolenza è la misura in cui le intenzioni e le motivazioni del fiduciario risultano allineate con quelle del soggetto che si affida a lui (Lee & See, 2004).

2.1.1 SVILUPPO DEL PROCESSO DI FIDUCIA

La relazione ideale, secondo il modello proposto da Mayer, Davis e Schoorman, è quella in cui la capacità, la benevolenza e l’integrità risultano allineate e percepite a un livello elevato, condizione in cui “the trustee would be deemed quite trustworthy” (Mayer et al., 1995). La fiducia, tuttavia, non deve essere concepita in termini dicotomici, come un fenomeno “bianco o nero”, ma piuttosto come un costrutto graduale, un continuum.

La presenza di un fattore non può compensare la mancanza di un altro: ciascuna dimensione svolge un ruolo specifico e insostituibile nel processo di valutazione del soggetto di fiducia. Ogni elemento, infatti, offre una prospettiva percettiva distinta attraverso cui chi si fida valuta l’affidabilità dell’altro. Se il soggetto di fiducia è percepito come elevato in tutti e tre i fattori, sarà considerato altamente affidabile. Al contrario, una persona può possedere un’elevata capacità (ad esempio essere un esperto) e una forte integrità (agire in modo coerente e onesto), ma se manca la benevolenza (ovvero la volontà di agire nell’interesse del trustor) la fiducia difficilmente si sviluppa. Analogamente, un individuo molto benevolo può non essere ritenuto degno di fiducia se non possiede le competenze necessarie per svolgere efficacemente il compito richiesto.

Come sottolinea Mayer, in alcune circostanze una fiducia significativa può emergere anche in presenza di livelli non pienamente elevati di uno o più fattori; tuttavia, tale dinamica dipende anche dalla propensione individuale a fidarsi, che modera il peso delle dimensioni di capacità, benevolenza e integrità. Un elemento centrale del modello è rappresentato dal ruolo del tempo e dallo sviluppo della relazione. Nelle fasi iniziali del rapporto, quando le interazioni dirette sono limitate, l’integrità assume un’importanza predominante, poiché può essere valutata attraverso osservazioni indirette, reputazione e informazioni provenienti da terzi. Con il progredire della relazione, invece, la benevolenza diventa sempre più rilevante, in quanto le interazioni dirette consentono di comprendere se l’altro agisce effettivamente con intenti positivi e orientati al benessere del

Dimensioni della fiducia

truster. Da ricordare che ogni volta che decidiamo di rischiare una relazione, l'esito (positivo o negativo) aggiorna le nostre percezioni (abilità, integrità e benevolenza), creando un processo di calibrazione continua (Mayer et al., 1995).

Tabella 2: Processo di sviluppo della fiducia tra due persone, suddiviso in 3 fasi progressive secondo il modello della Mayer et al., 1995.

Fase	Contesto	Determinanti principali della fiducia	Ruolo dei fattori	Note chiave
Fase 1 Inizio della relazione	Primo incontro; scarse informazioni dirette	Integrità percepita (coerenza, reputazione, fonti esterne) Capacità percepita (competenze, conoscenze) Propensione a fidarsi (caratteristica del trustor) Benevolenza percepita (limitata)	Integrità = fattore dominante Capacità = valutata indirettamente Benevolenza = poco osservabile	La fiducia iniziale si basa soprattutto sull'integrità e sulla propensione a fidarsi
Fase 2 Sviluppo della relazione	Aumento delle interazioni dirette e indirette	Benevolenza percepita (azioni concrete di supporto) Integrità (continua a influenzare) Capacità (risultati e competenza dimostrata) Propensione a fidarsi	Benevolenza = peso crescente Integrità = peso relativo minore Capacità = osservabile	La fiducia è dinamica e può aumentare o diminuire in base alle esperienze
Fase 3 Stabilizzazione della fiducia	Relazione consolidata	Capacità confermata Integrità confermata Benevolenza confermata	Tutti e tre i fattori risultano integrati e stabili	La fiducia è meno influenzata dalla propensione iniziale e più dalle evidenze accumulate

Analizziamo ora come i diversi tipi di fiducia influenzano le interazioni, distinguendo tra rapporti interpersonali e interazioni uomo-macchina, e come tali dinamiche variano in funzione del contesto. Come definisce il professor Bashir, "La formazione della Fiducia è un processo dinamico. Più le persone sono esposte a nuove informazioni e sentimenti, il loro livello di fiducia può cambiare drasticamente; questo può succedere sia a un livello conscio sia a un livello inconscio."⁴

Dopo aver spiegato i meccanismi di attivazione e lo sviluppo della fiducia nei rapporti, andiamo a classificare e a spiegare le diverse dimensioni della fiducia applicate ai rapporti interpersonali: abbiamo la Fiducia Cognitiva, che si basa sulla ragione e sulle competenze; la Fiducia Affettiva, basata sulle emozioni e sulla benevolenza; la Fiducia Percepita si basa sull'interpretazione soggettiva; e quella Appropriata, che si basa sull'allineamento tra percezione e realtà. La fiducia nella tecnologia, e in particolare nei sistemi automatizzati, richiede quindi processi diversi e più consapevoli rispetto alla fiducia interpersonale (McKnight e Chervany, 1998). Le quattro dimensioni della fiducia rappresentano componenti distinte che influenzano le interazioni sociali e tecnologiche. Nelle relazioni interpersonali prevalgono aspetti affettivi ed emotivi, mentre in contesti relazionali uomo-macchina (HMI/HCI) dominano quelli cognitivi e operativi.

4. Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3), 407–434

2.2.1 FIDUCIA COGNITIVA

La fiducia cognitiva si fonda su un processo di valutazione analitico e razionale dell'affidabilità dell'altro ed è il risultato di giudizi consapevoli basati su informazioni oggettive, esperienze pregresse e segnali osservabili relativi a competenza, coerenza e gestione del rischio (Lee & See, 2004; Bashir, 2023; Biermann et al., 2023). Il pensiero di Mayer, Davis e Schoorman (1995) identifica tre fattori che alimentano la percezione di affidabilità, due dei quali sono prettamente cognitivi. L'abilità (Ability) è la convinzione che l'individuo possieda le competenze e le caratteristiche necessarie per operare efficacemente in un dominio specifico. L'integrità (Integrity) si basa sulla percezione che il destinatario aderisca a un insieme di principi e regole che il fiduciante (truster) trova accettabili (Mayer et al., 1995). In modo coerente, McKnight e Chervany descrivono la fiducia cognitiva attraverso le "trusting beliefs", che sono le convinzioni che una persona ha sull'affidabilità dell'altro, valutando la competenza, la prevedibilità e l'integrità, rispondendo alla domanda "questa persona/ sistema ha le caratteristiche giuste per potermi fidare?" (McKnight e Chervany, 1998). A differenza della fiducia interpersonale, in cui gli esseri umani formulano giudizi

rapidi e inconsapevoli sull'affidabilità altrui attraverso meccanismi emotivi automatici (Engell et al., 2007)⁵. Questa forma di fiducia cognitiva emerge soprattutto nelle fasi iniziali delle relazioni, quando non sono ancora presenti legami emotivi significativi, e consente all'individuo di agire in condizioni di incertezza attraverso una riduzione della complessità, come descritto da Luhmann. La fiducia cognitiva può dunque essere intesa come una "scommessa razionale", nella quale il soggetto decide di affidarsi all'altro sulla base di prove tangibili di competenza e coerenza, piuttosto che su elementi affettivi (Mayer et al., 1995; Luhmann, 1998). La fiducia cognitiva presenta un carattere intrinsecamente specifico e contestuale. Un individuo può, ad esempio, riporre fiducia nelle capacità di un veicolo autonomo in uno scenario strutturato come l'autostrada, ma non estendere automaticamente tale fiducia a contesti urbani complessi e dinamici. Questa selettività anticipa il concetto di specificità funzionale e situazionale che verrà formalizzato nei modelli di Trust in Automation (Lee & See, 2004), evidenziando come la fiducia non sia un attributo globale del sistema, ma una valutazione legata a compiti e contesti definiti. Nel contesto dell'automazione e dei sistemi tecnologici, la fiducia cognitiva richiede un processo analitico che

5. Gli esseri umani formulano giudizi sull'affidabilità delle persone che incontrano in modo inconsapevole, sfruttando una notevole capacità di leggere segnali sottili. Secondo gli esperimenti di Engell, Haxby e Todorov (2007), l'amigdala, una regione del cervello situata nel lobo temporale medio, si attiva durante valutazioni rapide dell'affidabilità dei volti umani. La risposta dell'amigdala aumenta quando si è esposti a volti percepiti come meno affidabili, e si è osservato che una persona è in grado di formare giudizi iniziali sull'affidabilità in circa 100 millisecondi. (Bashir, 2023) Tale attivazione non si applica invece per i sistemi tecnologici. (Dal libro di Hoff Bashir, 2023)

utilizza risorse cognitive per valutare tre dimensioni principali: prestazioni (performance), ossia ciò che il sistema fa e quanto i suoi risultati siano affidabili e validi; processi (process), cioè come il sistema opera, comprendendo algoritmi e logica interna; e scopo (purpose), ossia perché il sistema è stato progettato e se agisce entro i limiti previsti (Lee & See, 2004). Gli utenti trattano la tecnologia come un "rischio prevedibile", valutando oggettivamente la competenza e la coerenza del sistema piuttosto che affidandosi alle emozioni (Biermann et al., 2023). La fiducia cognitiva nell'automazione è dinamica e fragile, poiché si aggiorna continuamente in base al feedback sulle prestazioni reali: se il sistema fallisce anche in compiti percepiti come semplici, la fiducia può diminuire rapidamente, mentre successi coerenti rafforzano la percezione di affidabilità e competenza.

Portando un esempio di rapporto interpersonale, quando un cliente che decide di rivolgersi a un avvocato o a un medico che non conosce personalmente esercita una forma di fiducia cognitiva basata su prove tangibili di competenza (McKnight e Chervany, 1998). Invece, in un contesto di alta tecnologia, come l'UAM, la fiducia di un operatore verso un sistema di navigazione si fonda sulla comprensione tecnica della sua robustezza (Zhan et al., 2022; Daniel Gerbeth, Chen et al., 2023).

Figura 3: Confronto sintetico della fiducia Cognitiva.

Caratteristiche	Rapporto interpersonale	Rapporto Uomo-Macchina
Pillar centrale	Abilità e Integrità	Performance e Processo
Origine	Reputazione, titoli professionali, osservazione della condotta passata	Affidabilità tecnica, trasparenza dell'interfaccia e specifiche di sistema
Natura del rischio	Incompetenza nel compito, violazione dei principi etici concordati	Rischio prevedibile di probabilità di guasto
Meccanismo decisionale	Scommessa razionale	Giudizio basato sui dati

MENTRE NEI RAPPORTI UMANI CERCHIAMO UN'ANIMA CHE NON CI TRADISCA, NEL RAPPORTO CON LA MACCHINA CERCHIAMO SEGNALI SOCIALI CHE TRASFORMINO UN ALGORITMO COMPLESSO IN UN COMPAGNO DIVIAGGIO DI CUI POTER "SENTIRE" LA SICUREZZA.

Lee e See, 2004

2.2.2 FIDUCIA AFFETTIVA

La fiducia affettiva si fonda su un processo di valutazione analogico e affettivo per gestire la complessità e l'incertezza in momenti di vulnerabilità in una situazione di rischio (Bashir 2015; Biermann 2023), portando l'individuo a sentirsi sicuro, assicurato e a proprio agio (Felt Security). Le emozioni sono spesso il determinante primario (Bashir 2015). Secondo Mayer, Davis e Schoorman (1995), questo tipo di fiducia è legata alla credenza che l'altro provi "Benevolenza" nei nostri confronti, desiderando il nostro bene indipendentemente da un tornaconto personale (Mayer et al., 1995; McKnight e Chervany, 1998). È definito come un "istinto viscerale" o un "gut

feeling" (Biermann, 2023). Accade che molte persone non si fidano di qualcuno semplicemente a causa di una sensazione di disagio che non sappiamo tradurre a parole (Bashir, 2015). Lee e See osservano che l'affetto agisce come un segnale euristico decisionale quando le risorse cognitive sono limitate: le emozioni ci permettono di agire senza dover analizzare ogni singola intenzione dell'altro (Lee e See, 2004). Per capire meglio le dinamiche, prendiamo per esempio il rapporto tra un mentore e un protetto. Il protetto si fida affettivamente del mentore non solo perché è competente (abilità cognitiva), ma perché percepisce un legame di cura e lealtà (Bashir, 2015; McKnight e Chervany, 1998). Nel contesto dell'automazione e dei

Tabella 4: Confronto sintetico della fiducia Affettiva.

Caratteristiche	Rapporto interpersonale	Rapporto Uomo-Macchina
Pillar centrale	Abilità e Integrità	Performance e Processo
Origine	Reputazione, titoli professionali, osservazione della condotta passata	Affidabilità tecnica, trasparenza dell'interfaccia e specifiche di sistema
Natura del rischio	Incompetenza nel compito, violazione dei principi etici concordati	Rischio prevedibile di probabilità di guasto
Meccanismo decisionale	Scommessa razionale	Giudizio basato sui dati

sistemi tecnologici, la fiducia tende a essere più orientata alla dimensione cognitiva (prestazione e processo), ma la ricerca dimostra che gli esseri umani rispondono alle macchine in modo sociale, applicando sentimenti e norme di "etiquette" (Biermann, 2023; Lee e See, 2004). La fiducia affettiva qui emerge attraverso "l'Antropomorfismo": dare un nome, un genere o una voce umana a un veicolo autonomo (come nel caso del setup sperimentale "Iris") aumenta la percezione che la macchina abbia una "mente" capace di sentire, rendendo l'utente più rilassato anche in caso di incidente (A. Waytz et al., 2014). Una macchina che genera un legame affettivo beneficia di una maggiore tolleranza agli errori. Gli utenti tendono a colpevolizzare meno un sistema "umanizzato" in caso di malfunzionamento, poiché gli attribuiscono una "buona intenzione" (Biermann, 2023).

2.2.3 FIDUCIA OPERATIVA

La fiducia operativa si fonda su un processo di valutazione legato alla dimensione dell'abilità, ai risultati tangibili, alla prevedibilità e all'affidabilità delle azioni di un soggetto o di un sistema nel portare a termine con successo un compito specifico (Bashir, 2015; Biermann, 2023). Essa rappresenta il passaggio dall'atteggiamento fiduciario all'atto concreto di affidamento, nel quale il soggetto accetta la vulnerabilità assumendo consapevolmente un rischio, già definita nei modelli interpersonali (Mayer et al., 1995). In questo senso, la fiducia operativa rappresenta il passaggio dal "pensare di potersi fidare" all'"agire sulla base di tale convinzione" (Biermann, 2023; Mayer et al., 1995) e inizia precisamente nel momento in cui termina il controllo diretto. Per comprendere meglio questa dinamica, si può considerare il rapporto tra un manager e un collaboratore a cui viene delegato un progetto cruciale. Il manager non si limita ad avere fiducia come atteggiamento cognitivo, ma smette di monitorare costantemente l'operato del collaboratore, esercitando una forma di fiducia operativa. In questo modo accetta il rischio che un eventuale errore possa avere conseguenze negative, ad esempio sulla propria reputazione professionale (Mayer et al., 1995). Nell'ambito dell'automazione e dei sistemi tecnologici, la

fiducia operativa si concentra prevalentemente sulla capacità tecnica del sistema di eseguire compiti specifici in modo accurato e prevedibile (Lee & See, 2004; Zhang et al., 2021; Tian & Duff). Secondo il modello di Lee e See, essa può essere analizzata attraverso due dimensioni principali. La performance riguarda l'affidabilità (reliability) del sistema e descrive ciò che la macchina è effettivamente in grado di fare, ovvero la capacità di raggiungere gli obiettivi dell'utente in modo corretto e senza errori (Bashir, 2015). Il processo, invece, si riferisce al modo in cui il sistema opera: quando l'utente comprende la logica o l'algoritmo sottostante, ovvero percepisce un adeguato livello di trasparenza, è più incline a fare affidamento sul sistema anche in situazioni nuove, potendone prevedere il comportamento. Nei sistemi complessi, come gli aerotaxi o i veicoli autonomi, la fiducia operativa è strettamente connessa al concetto di integrità, intesa non

solo come valore morale, ma come misura quantitativa della sicurezza della navigazione (Zhan, Zhai & Zheng, 2022). In questo contesto, l'integrità indica il grado di affidabilità attribuibile alla correttezza dell'output di navigazione e alla capacità del sistema di operare entro margini di sicurezza accettabili (Zhan, Zhai & Zheng, 2022). Per fare un esempio concreto, la fiducia operativa è come un pilota che inserisce il cruise control in autostrada. Questa scelta non deriva da un legame affettivo con il veicolo, ma perché ha verificato che il sistema mantiene la velocità impostata (performance) e sa che i sensori radar frenano se un'auto taglia la strada (processo) (Lee e See, 2004). Qualora il sistema accelerasse senza motivo anche solo una volta, il pilota prenderebbe il controllo manuale e la sua fiducia operativa svanirebbe, poiché la macchina ha dimostrato di non avere più l'abilità "necessaria" per quel compito. (Lee e See, 2004).

Tabella 5: Confronto sintetico della fiducia Operativa.

Caratteristiche	Rapporto interpersonale	Rapporto Uomo-Macchina
Pillar centrale	Abilità e reputazione tecnica	Performance e Processo
Origine	Evidenze di competenze	Integrità e affidabilità
Natura del rischio	Incompetenza specifica	Guasto o errore di sistema
Meccanismo decisionale	Valutazione razionale	Calibrazione basata sui dati

2.2.3 FIDUCIA PERCEPITA VS. FIDUCIA EFFETTIVA

La fiducia percepita è l'atteggiamento mentale e la valutazione soggettiva dell'individuo; essa si basa su una valutazione razionale delle qualità del trustee (Mayer et al., 1995; Bashir, 2015). All'inizio di una relazione, la percezione della fiducia si fonda spesso sulla reputazione o sull'integrità osservata (Mayer et al., 1995). La fiducia effettiva, detta anche fiducia calibrata (che verrà approfondita nel capitolo 2.4), rappresenta invece la misura della corrispondenza oggettiva tra il livello di fiducia del trustor e le reali capacità o l'affidabilità del trustee (Walker et al., 2023). Essa si manifesta attraverso il Risk Taking in Relationship (RTR), ovvero l'assunzione concreta di rischio all'interno della relazione (Mayer et al., 1995): solo quando l'individuo si rende vulnerabile, la fiducia percepita si traduce in un comportamento effettivo. Nel tempo, la percezione viene aggiornata sulla base dei risultati delle azioni passate attraverso un feedback loop, allineando gradualmente le valutazioni soggettive alla realtà (Mayer et al., 1995). Per esempio, la fiducia percepita è paragonabile all'assunzione di un dipendente sulla base di un ottimo curriculum e di un atteggiamento cordiale (reputazione/affetto); la fiducia effettiva consiste invece nel sapere, dopo un anno di lavoro, quali compiti assegnare con precisione, conoscendone i limiti e i punti di forza reali (calibrazione). Basarsi esclusivamente sulla percezione può portare ad affidargli incarichi eccessivamente complessi (overtrust/misuse) oppure, al contrario, a non delegare nulla per timore (undertrust/disuse).

Nel contesto dell'automazione e dei sistemi tecnologici, il divario tra percezione e realtà risulta spesso più marcato a causa dei bias cognitivi. In questi casi, la fiducia percepita è frequentemente influenzata da fattori superficiali quali l'antropomorfismo, l'estetica dell'interfaccia o la reputazione del marchio (Lee & See, 2004). L'utente può quindi sviluppare un'elevata fiducia semplicemente perché la macchina appare "intelligente" (Waytz et al., 2014). La fiducia effettiva, invece, dipende dalla comprensione dei parametri di accuratezza, risoluzione e specificità del sistema e diventa pienamente calibrata solo quando l'utente conosce con precisione le situazioni

Negli ultimi anni abbiamo assistito allo sviluppo di alcune tecnologie molto significative, ad esempio i droni impiegati nella logistica, nel monitoraggio civile e nella sensoristica avanzata e i sistemi di guida autonoma basati sull'intelligenza artificiale e sulla computer vision applicata al settore automotive (Romero, 2025). Contemporaneamente stiamo assistendo a una crescita

in cui la macchina può fallire o, al contrario, in cui è altamente affidabile (Walker et al., 2023). Mentre nelle relazioni interpersonali la fiducia tende a crescere lentamente, nei confronti delle macchine si osserva spesso un'elevata fiducia percepita iniziale (bias di perfezione), che può crollare drasticamente al primo errore, rendendo la calibrazione particolarmente fragile (Bashir, 2015). Queste due forme di fiducia sono complementari e, quando non coincidono, possono generare dinamiche di eccesso di fiducia (overtrust) o di diffidenza (distrust).

La fiducia percepita verso un capitano umano nasce dal pensare: “È una brava persona che non ci metterebbe mai in pericolo” (Benevolenza/Integrità); la fiducia percepita verso un pilota automatico nasce dal pensare: “Il software è stato programmato da

esperti per gestire ogni turbolenza” (Processo/Scopo). In entrambi i casi, ci sediamo sull'aereo basandoci su un'idea di sicurezza prima ancora che il volo sia decollato.

La fiducia effettiva verso un collega è sapere che, se gli affidi un compito difficile, la sua integrità lo porterà a chiederti aiuto se non ce la fa (esito del rischio); la fiducia effettiva verso un veicolo autonomo è sapere che i suoi sensori sono perfetti in autostrada ma fallibili nei centri storici affollati (specificità funzionale e situazionale). In entrambi i casi, la fiducia è “buona” non perché è massima, ma perché è esattamente proporzionata alla realtà.

Tabella 6: Confronto sintetico della fiducia Effettiva.

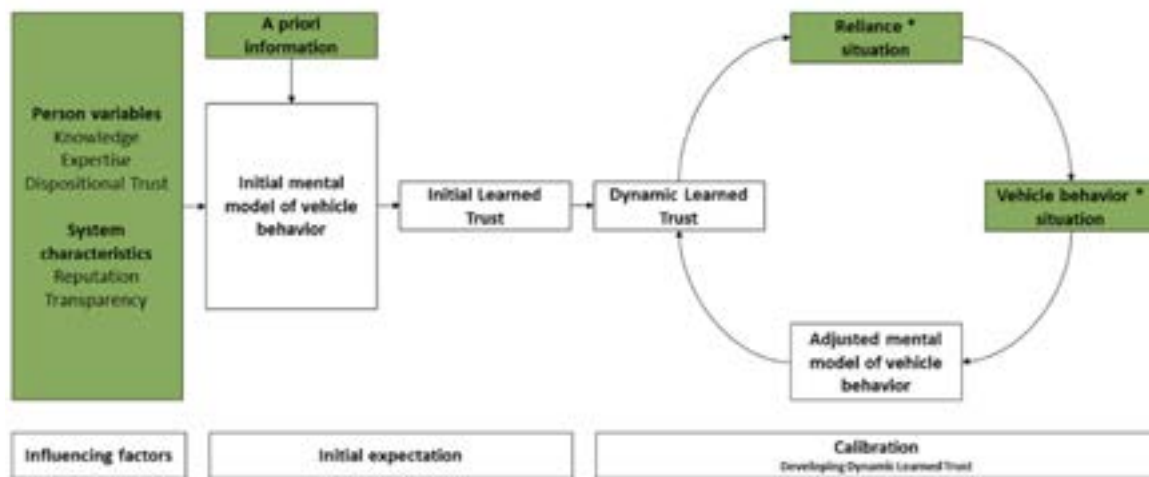
Caratteristiche	Rapporto interpersonale	Rapporto Uomo-Macchina
Pillar centrale	Aspettativa morale e collaborativa	Valutazione soggettiva della capacità
Origine	Attribuzioni caratteriali: Abilità, Integrità, Benevolenza	Dimensioni tecniche e sociali: Prestazioni, Processo, Scopo
Natura del rischio	Vulnerabilità Sociale	Incertezza tecnica e sistematica
Meccanismo decisionale	Prevalentemente affettiva	Euristica cognitiva

2.3 I MODELLI DI FIDUCIA NELL'AUTOMAZIONE

Gli studi condotti dai ricercatori sulle relazioni tra individui e agenti automatizzati utilizzano concetti derivati dagli studi sulla fiducia interpersonale (Walker et al., 2018; Kraus, 2020). La definizione più ampiamente adottata proviene da Lee e See (2004), che definiscono la fiducia come “l'atteggiamento secondo cui un agente aiuterà a raggiungere gli obiettivi di un individuo in una situazione caratterizzata da incertezza e vulnerabilità” (p.51). “..Mentre la fiducia interpersonale può basarsi sull'abilità, l'integrità o la benevolenza di un fiduciario (Mayer et al., 1995), la fiducia uomo-automazione dipende dalle prestazioni, dal processo o dallo scopo di un sistema automatizzato (Lee & Moray, 1992). Nel contesto delle relazioni tra individui, la fiducia tende a svilupparsi in modo graduale: nelle fasi iniziali essa si fonda prevalentemente sulla prevedibilità del comportamento dell'altro, riflettendo un approccio prudente verso soggetti sconosciuti (Rempel, Holmes, & Zanna, 1985). Con il progredire della relazione, l'attenzione si sposta sull'attendibilità e sull'integrità del fiduciario, fino a raggiungere, nelle relazioni mature, una fiducia basata sulla benevolenza e sulla fede reciproca (Lee & See, 2004). Al contrario, la fiducia uomo-automazione sembra seguire un percorso inverso. Studi empirici indicano che gli utenti manifestano spesso un pregiudizio di positività (positivity bias) nei confronti dei nuovi sistemi automatizzati, attribuendo loro un'elevata affidabilità fin dalle prime interazioni (Dzindolet et

Tabella 7: Confronto sintetico della fiducia Percepita.

Caratteristiche	Rapporto interpersonale	Rapporto Uomo-Macchina
Pillar centrale	Coerenza caratteriale	Calibrazione tecnica
Origine	Esito del rischio (RTR)	Esperienza dei Limiti
Natura del rischio	Tradimento o Delusione Morale	Overtrust e Undertrust
Meccanismo decisionale	Rivalutazione delle Attribuzioni	Risoluzione e Specificità



al., 2003). In queste fasi iniziali, la fiducia è frequentemente basata su una fede implicita nella perfezione tecnologica. Tuttavia, tale fiducia risulta particolarmente fragile e tende a diminuire rapidamente in seguito al verificarsi di errori. Con l'aumentare dell'esperienza d'uso, la fede viene progressivamente sostituita da valutazioni più realistiche, "fondate sull'attendibilità e sulla prevedibilità del comportamento del sistema" (Hoffman e Bashir, 2015).

La fiducia nei sistemi automatizzati riflette in parte i meccanismi interpersonali: gli utenti valutano l'affidabilità in funzione degli obiettivi da raggiungere e della presenza di incertezza o rischio. Diversi studi negli anni '90 hanno dimostrato che le persone applicano regole apprese socialmente, come la gentilezza, alle interazioni con le macchine (Hoffman e Bashir, 2015). Studi scientifici dimostrano che quando decidiamo se fidarci di un sistema tecnologico

(es. un'offerta su eBay), attiviamo le stesse aree del cervello che usiamo nei "giochi di fiducia" tra esseri umani. "Una possibile ragione di queste somiglianze risiede nel fatto che, in un certo grado, la fiducia delle persone nei sistemi tecnologici rappresenta la fiducia riposta nei progettisti di tali sistemi (Parasuraman & Riley, 1997). In questo senso, la fiducia uomo-automazione può essere considerata come un tipo specifico di fiducia interpersonale in cui il fiduciario (trustee) è separato dal fiduciante (truster) da un ulteriore livello di mediazione." È importante sottolineare che le dinamiche descritte risultano particolarmente rilevanti nelle interazioni con fiduciari e sistemi automatizzati non familiari. Nel contesto reale, la formazione della fiducia è invece un processo complesso e multidimensionale, influenzato da fattori legati all'operatore, all'ambiente operativo e alle caratteristiche del sistema automatizzato. I principali modelli di

Figura 2.2: Un quadro concettuale dello sviluppo della fiducia verso uno specifico sistema di guida automatizzata (basato su Lee e See, 2004; Ghazalehet a 2012, Hoff e Bashir, 2015, Kraus, 2020, Walker, 2021). By Trust in automated vehicles constructs, psychological processes, and assessment, 2023.

9. Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3), 407–434.

Trust in Automation, analizzati nelle sezioni seguenti, sono stati sviluppati e validati in contesti ad alta criticità, in particolare in quello aeronautico, che rappresenta uno dei primi e più complessi esempi di interazione uomo-automazione.

2.3.1 IL MODELLO DI LEE E SEE SULLA TRUST IN AUTOMATION

Il contributo di John D. Lee e Katrina A. See rappresenta una pietra miliare nello studio della fiducia nei sistemi complessi, definendola come l'attitudine mentale secondo cui un agente, umano o tecnologico, aiuterà un individuo a raggiungere i propri obiettivi in condizioni di incertezza e vulnerabilità, introducendo il concetto di appropriate reliance (Lee & See, 2004; Zhang et al., 2021). Il modello descrive la fiducia come un processo dinamico, influenzato dal contesto, dalle caratteristiche dell'automazione e dai processi cognitivi, con feedback continuo che ne permette la calibrazione per evitare overtrust o undertrust (Lee & See, 2004).

Lee e See hanno identificato tre livelli di astrazione attraverso i quali gli esseri umani valutano la tecnologia. Queste dimensioni permettono ai progettisti di capire quali informazioni mostrare all'utente per supportare la fiducia. Abbiamo le prestazioni (performance-based), simili alla fiducia basata sull'abilità di Mayer et al., che varia a seconda di quanto efficacemente un sistema automatizzato esegue un compito. Risponde alla domanda: "Come fa?" La fiducia basata sul processo (process-based), analoga alla fiducia basata sull'integrità tra esseri umani, fluttua in base alla comprensione che l'operatore ha dei metodi utilizzati dal sistema per svolgere i compiti e risponde alla domanda: "Come lo fa?" Infine, la fiducia basata sullo scopo (purpose-based) è legata all'uso previsto dal progettista per il sistema automatizzato e risponde alla domanda: "Perché lo fa?" (Hoffman e Bashir, 2015). Per valutare che il rapporto uomo-macchina sia "sano", il modello introduce tre parametri fondamentali: la calibrazione, che misura l'allineamento tra ciò che si crede che il sistema sappia fare e ciò che sa realmente fare ("Mi fido nel modo giusto?"); la risoluzione, che indica quanto finemente la fiducia reagisce ai cambiamenti di

performance (“Mi accorgo quando le cose cambiano?”); e la specificità, che permette di fidarsi selettivamente di funzioni o contesti specifici (“So di cosa fidarmi e di cosa no?”). Infine, Lee e See sottolineano che la fiducia evolve secondo un modello a ciclo chiuso (closed-loop), caratterizzato da un’inerzia temporale: errori o miglioramenti delle prestazioni influenzano l’atteggiamento fiduciario in modo graduale (Lee e See, 2004).

Questo processo è mediato da diversi meccanismi cognitivi (analitici, analogici e affettivi) che operano simultaneamente. In situazioni di emergenza, ad esempio, una risposta emotiva di paura può temporaneamente prevalere su valutazioni razionali delle prestazioni del sistema. Nel complesso, il modello di Lee e See fornisce prove di concetti chiave per progettare sistemi automatizzati che supportino una fiducia consapevole, differenziata e correttamente calibrata (Lee e See, 2004).

2.3.2 IL MODELLO DI PARASURAMAN SUI LIVELLI DI AUTOMAZIONE E LA FIDUCIA

Il contributo di Parasuraman et al. risiede nella consapevolezza che l’automazione non è un’entità

monolitica, ma un processo che trasforma la natura del lavoro umano. Affermano che “l’automazione non si limita a sostituire, ma cambia l’attività umana e può imporre nuove richieste di coordinamento all’operatore”.¹¹ Nasce l’esigenza di distinguere quattro tipi di automazione, basati sulle funzioni cognitive umane (Parasuraman et al., 2000). Abbiamo l’acquisizione delle informazioni (Information Acquisition); l’automazione va a supportare i sensi, ad esempio i radar che scansionano il traffico aereo o sensori che evidenziano ostacoli critici. L’analisi delle informazioni (Information Analysis) è l’automazione che supporta i processi cognitivi come la memoria di lavoro e l’inferenza; un esempio è il display predittivo che mostra la rotta futura di un aereo o integra più variabili in un unico indicatore. Selezione delle decisioni (Decision Selection); la macchina suggerisce o sceglie tra diverse opzioni. Ad esempio, un sistema esperto che propone una rotta alternativa per evitare il maltempo. Infine, implementazione dell’azione (Action Implementation); la macchina esegue fisicamente il compito, sostituendo la mano o la voce dell’umano. Esempi classici sono l’autopilota che esegue una manovra o il caricamento automatico di un piano di volo (Parasuraman et al., 2000). All’interno di ogni stadio, l’automazione può operare su una

Tabella 8: Livelli di automazione della decisione e della selezione dell’azione.

HIGH	10. The computer decides everything, acts autonomously, ignoring the human.
	9. informs the human only if it, the computer, decides to
	8. informs the human only if asked, or
	7. executes automatically, then necessarily informs the human, and
	6. allows the human a restricted time to veto before automatic execution, or
	5. executes that suggestion if the human approves, or
	4. suggests one alternative
	3. narrows the selection down to a few, or
	2. The computer offers a complete set of decision/action alternatives, or
LOW	1. The computer offers no assistance: human must take all decisions and actions.

Figura 2.3: Modello semplice a quattro fasi dell’elaborazione delle informazioni umane, by Parasuraman, 2020.



scala continua che va da 1 (è manuale) a 10 (è completamente autonomo), la quale deve basarsi su test di usabilità e affidabilità. [Fig.2.3] Sottolineano che “un’automazione affidabile può essere peggiore dell’assenza di automazione”¹². Se il sistema sbaglia anche raramente, l’operatore può sviluppare una “compiacenza da automazione”, riducendo il monitoraggio e fallendo nel rilevare gli errori quando si verificano. Pertanto, il design deve puntare a mantenere l’utente coinvolto nel processo cognitivo (situational awareness) per garantire la sicurezza complessiva del sistema.

Lui vede la fiducia come mediatore tra i livelli di automazione e il comportamento dell’operatore non è sempre necessario puntare sul livello più alto di automazione per avere la massima resa. Porta come esempio la strategia di “Error Trapping”, ovvero la cattura dell’errore, dove; in situazioni ad alto rischio, se la fase di selezione della decisione è automatizzata a livelli alti, è prudente mantenere l’automazione dell’azione a un livello basso. Per esempio: Invece di far eseguire all’aereo una manovra correttiva automaticamente, il sistema dovrebbe richiedere al pilota di premere fisicamente un pulsante per “accettare”

¹¹. A Model for Types and Levels of Human Interaction with Automation” di Parasuraman, Sheridan e Wickens (2000).

¹². A Model for Types and Levels of Human Interaction with Automation” di Parasuraman, Sheridan e Wickens (2000).

la decisione caricata nell’FMS. Questo obbliga l’umano a un atto consapevole di revisione, riducendo l’acquiescenza acritica. Per una buona gestione dei carichi di lavoro (Mental Workload): L’automazione della prima fase (Acquisizione) e della seconda (Analisi) è generalmente raccomandata anche a livelli alti, poiché riduce il carico di lavoro senza escludere l’umano dal ciclo decisionale. Il livello di automazione adattiva non dovrebbe essere fisso, ma varia in base alle richieste situazionali. Per esempio: In un decollo, se avviene un guasto motore vicino alla velocità critica VI, l’automazione può passare istantaneamente a un livello alto (decisione automatica di abortire o continuare) perché l’umano non avrebbe tempo fisico per reagire in sicurezza (Parasuraman et al., 2000). Per lui, l’uomo assume il ruolo di un operatore d’orchestra.

2.3.3 IL MODELLO DI HOFFMAN SULLA FIDUCIA, TRASPARENZA E CALIBRAZIONE

Il modello di Kevin Anthony Hoff e Masooda Bashir va a sintetizzare la variabilità della fiducia come un costruito a tre strati; loro vanno a “proporre raccomandazioni di progettazione per la creazione ambientale che possono influenzare

la forza della relazione tra fiducia e affidamento”¹³, cioè va a vedere come l’ambiente circostante influenzi le dinamiche di fiducia sulle persone. Questo modello è fondamentale per comprendere come la trasparenza del design porti a una calibrazione efficace, evitando i rischi di abuso (misuse) o abbandono (disuse) della tecnologia (Hoff e Bashir 2015). Il primo strato è Fiducia Disposizionale (Dispositional Trust), che rappresenta la tendenza duratura e stabile di un individuo a fidarsi dell’automazione, indipendentemente dal contesto o dal sistema specifico. È influenzata da fattori intrinseci della persona come la cultura, l’età, il genere e la personalità. Per esempio, una ricerca indica che le persone estroverse o con una personalità “intuitiva” tendono ad avere una predisposizione maggiore a fidarsi dei sistemi intelligenti rispetto agli introversi. In aviazione, un pilota con alta fiducia disposizionale potrebbe essere inizialmente più aperto a testare un nuovo sistema UAM rispetto a un collega più scettico. Il secondo strato è la Fiducia Situazionale (Situational Trust), che dipende dal contesto specifico e transitorio dell’interazione. È influenzata da fattori esterni (la complessità del sistema, difficoltà del compito e il livello di rischio ambientale) e da fattori interne (l’umore e l’autoefficacia) che variano a seconda del contesto. Per esempio, un operatore potrebbe fidarsi molto

di un sensore di navigazione in condizioni di bel tempo (basso rischio), ma ridurre drasticamente tale fiducia in presenza di nebbia fitta o in un “canyon urbano” dove il segnale GNSS è instabile (alto rischio).

Infine, il terzo livello è la fiducia appresa o operativa (learned trust), che si basa sulle esperienze passate e sulla valutazione corrente di un sistema specifico. Essa comprende una fiducia iniziale, fondata sulla reputazione del produttore o sulla documentazione, e una fiducia dinamica, che evolve durante l’uso in base alle prestazioni reali della macchina. Ad esempio, se un velivolo autonomo commette un errore grave nelle fasi iniziali di utilizzo, la fiducia dell’utente diminuirà molto più rapidamente rispetto a un errore che si verifica dopo un lungo periodo di funzionamento corretto. La fiducia appropriata (calibrata), si ottiene quando questi tre livelli risultano allineati con le reali capacità del sistema (Hoff e Bashir 2015). Un elemento che favorisce la fiducia, secondo Hoff e Bashir, è la trasparenza, intesa come il grado in cui la logica interna del sistema risulta comprensibile all’essere umano. Essa può essere supportata attraverso feedback chiari e in tempo reale, l’integrazione controllata dell’antropomorfismo, che aumenta la resilienza della fiducia, e l’adozione di uno stile di comunicazione non interrottivo e paziente (etiquette).

2.4 CALIBRAZIONE DELLA FIDUCIA, DALL’ECESSIVA FIDUCIA ALLA SFIDUCIA

In settori ad alta automazione (come aviazione, automotive e controllo dei processi) la calibrazione della fiducia è cruciale per garantire sicurezza e performance ottimali dei sistemi socio-tecnici. In questo contesto, l’eccessiva fiducia (descritto anche come compiacenza; Muir, 1987; Parasuraman et al., 1993; Parasuraman e Riley, 1997; Lee e See, 2004) e sfiducia (disuse; Muir, 1987; Lee e See, 2004; Sheridan et al., 2005; Hoff e Bashir, 2015; Wintersberger et al., 2018; Lee e Ji, 2023) rappresentano due forme di calibrazione errata della fiducia che possono condurre a conseguenze catastrofiche, compromettendo la sicurezza e l’efficienza della cooperazione uomo-automazione.

¹³. Trust In Automation: Integrating Empirical Evidence on Factors that Influence Trust, Hoff e Masooda Bashir, 2015.

Questa distinzione concettuale costituisce una sintesi delle teorie fondamentali formulate originariamente da Parasuraman e Riley (1997) e successivamente approfondite da Lee e See (2004). L'eccessiva fiducia (misuse) "si riferisce ai fallimenti che si verificano quando le persone violano inavvertitamente presupposti critici e si affidano all'automazione in modo inappropriato" (Lee e See, 2004).

Un eccesso di fiducia nei sistemi automatizzati, come evidenziato da Bainbridge attraverso le cosiddette "ironie dell'automazione", può determinare una perdita di competenze (skill loss) e una riduzione della vigilanza, conducendo a risposte lente o inadeguate in presenza di guasti imprevisti (Lee e See, 2004). Organizzazioni come la NHTSA (2022) e il Dutch Safety Board (OVV, 2019) hanno riportato e analizzato incidenti e sinistri mortali direttamente riconducibili all'uso e all'eccessiva fiducia di tali tecnologie. Un esempio emblematico di questo pattern si osserva quando i piloti, confidando eccessivamente nelle capacità del pilota automatico, non intervengono assumendo il controllo manuale anche in situazioni critiche, come nel caso dell'incidente dell'Airbus A320 (Sparaco, 1995; Lee e See, 2004; Walker, 2023). D'altra parte, gli operatori non sono sempre disposti a riportare un livello

adeguato di fiducia nell'automazione. In questi casi si manifesta sfiducia (disuse), che "signifies failures that occur when people reject the capabilities of automation" (Lee e See, 2004). Il rifiuto di utilizzare l'automazione comporta la rinuncia ai potenziali benefici in termini di sicurezza e precisione, aumentando il carico di lavoro umano e la probabilità di errori manuali in situazioni complesse. Per esempio, "the Costa Concordia cruise ship disaster that killed 32 passengers in January 2012 may have been the result of the captain's under-trusting of the ship's navigation system in favor of manual control. Ensuing investigations discovered that the captain diverged from the ship's computer-programmed route before hitting the shallow reef that caused the sinking" (Hoff and Bashir, 2015).

Queste forme di calibrazione errata creano "partnership difettose" (Lee and See, 2004). Se l'uomo non capisce quando intervenire (overtrust) o quando lasciar agire la macchina (undertrust), la sinergia del sistema socio-tecnico fallisce (Lee e See, 2004; Walker, 2023). In settori come l'aviazione e la navigazione, dove l'incertezza e la vulnerabilità sono costanti, la fiducia agisce come una scorciatoia decisionale: se è calibrata male, l'intero sistema diventa instabile e pronto a incidenti catastrofici. In teoria, un sistema perfettamente

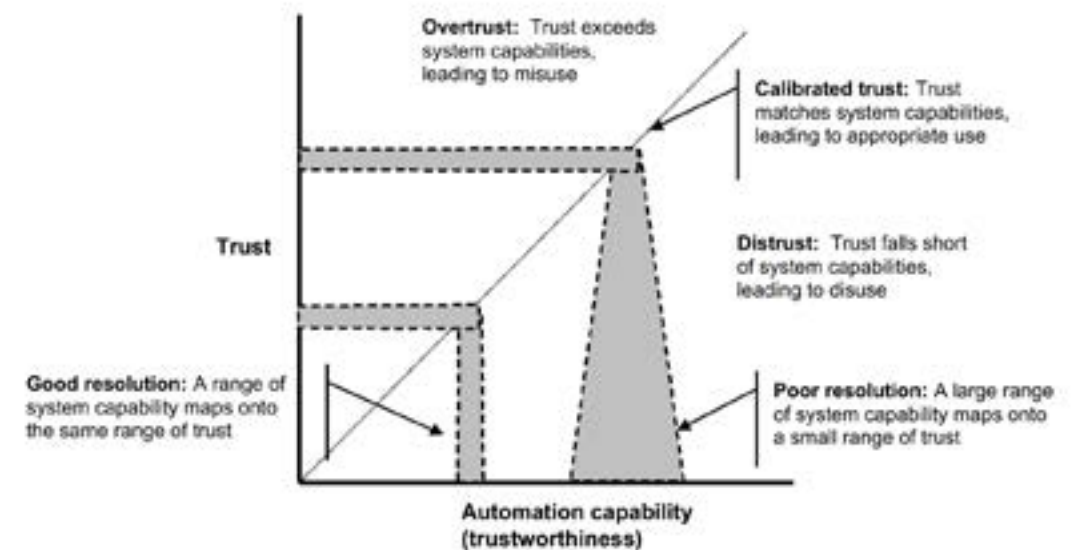


Figura 2.3: La relazione tra calibrazione, risoluzione e capacità dell'automazione nel definire un'appropriata fiducia nell'automazione. Un eccesso di fiducia (overtrust) può portare all'uso improprio (misuse), mentre la sfiducia (distrust) può portare al mancato utilizzo (disuse).

affidabile dovrebbe essere sempre considerato degno di fiducia; tuttavia, il conseguimento di un'affidabilità totale nei sistemi automatizzati complessi è irrealistico. Inoltre, la fiducia, in quanto costruito psicologico, è profondamente influenzata da fattori soggettivi che non sempre rispecchiano l'affidabilità oggettiva del sistema (Lee e See, 2004; Hergeth et al., 2016). Di conseguenza, l'obiettivo della progettazione dei sistemi automatizzati e dei programmi di addestramento non è l'incremento indiscriminato della fiducia, bensì la promozione di una fiducia calibrata. Questa "Calibrazione" fa riferimento "alla corrispondenza tra la fiducia di una persona nell'automazione e le capacità dell'automazione" (Lee e See, 2004). Viene considerata ottima quando il livello di fiducia percepita dall'utente è perfettamente allineato alla fiducia effettiva della tecnologia in una determinata situazione. La calibrazione è un processo dinamico che si evolve attraverso cicli di feedback durante l'interazione. Essa dipende anche dalla risoluzione, che indica quanto la fiducia cambia quando cambia la capacità dell'automazione, e dalla specificità funzionale e/o temporale, che riguarda la capacità di fidarsi di singole funzioni o modalità piuttosto che dell'intero veicolo in blocco. (Lee e See, 2004). La relazione ideale, secondo il modello proposto da Mayer,

Davis e Schoorman, è quella in cui la capacità, la benevolenza e l'integrità risultano allineate e percepite a un livello elevato, condizione in cui "the trustee would be deemed quite trustworthy" (Mayer et al., 1995). La fiducia, tuttavia, non deve essere concepita in termini dicotomici, come un fenomeno "bianco o nero", ma piuttosto come un costrutto graduale, un continuum. La presenza di un fattore non può compensare la mancanza di un altro: ciascuna dimensione svolge un ruolo specifico e insostituibile nel processo di valutazione del soggetto di fiducia. Ogni elemento, infatti, offre una prospettiva percettiva distinta attraverso cui chi si fida valuta l'affidabilità dell'altro. Se il soggetto di fiducia è percepito come elevato in tutti e tre i fattori, sarà considerato altamente affidabile. Al contrario, una persona può possedere un'elevata capacità (ad esempio essere un esperto) e una forte integrità (agire in modo coerente e onesto), ma se manca la benevolenza (ovvero la volontà di agire nell'interesse del truster) la fiducia difficilmente si sviluppa. Analogamente, un individuo molto benevolo può non essere ritenuto degno di fiducia se non possiede le competenze necessarie per svolgere efficacemente il compito richiesto.

Come sottolinea Mayer, in alcune circostanze una fiducia significativa

può emergere anche in presenza di livelli non pienamente elevati di uno o più fattori; tuttavia, tale dinamica dipende anche dalla propensione individuale a fidarsi, che modera il peso delle dimensioni di capacità, benevolenza e integrità. Un elemento centrale del modello è rappresentato dal ruolo La relazione ideale, secondo il modello proposto da Mayer, Davis e Schoorman, è quella in cui la capacità, la benevolenza e l'integrità risultano allineate e percepite a un livello elevato, condizione in cui "the trustee would be deemed quite trustworthy" (Mayer et al., 1995). La fiducia, tuttavia, non deve essere concepita in termini dicotomici, come un fenomeno "bianco o nero", ma piuttosto come un costrutto graduale, un continuum. La presenza di un fattore non può compensare la mancanza di un altro: ciascuna dimensione svolge un ruolo specifico e insostituibile nel processo di valutazione del soggetto di fiducia. Ogni elemento, infatti, offre una prospettiva percettiva distinta attraverso cui chi si fida valuta l'affidabilità dell'altro. Se il soggetto di fiducia è percepito come elevato in tutti e tre i fattori, sarà considerato altamente affidabile. Al contrario, una persona può possedere un'elevata capacità (ad esempio essere un esperto) e una forte integrità (agire in modo coerente e onesto), ma se manca la benevolenza (ovvero la volontà di agire nell'interesse del

truster) la fiducia difficilmente si sviluppa. Analogamente, un individuo molto benevolo può non essere ritenuto degno di fiducia se non possiede le competenze necessarie per svolgere efficacemente il compito richiesto.

Come sottolinea Mayer, in alcune circostanze una fiducia significativa può emergere anche in presenza di livelli non pienamente elevati di uno o più fattori; tuttavia, tale dinamica dipende anche dalla propensione individuale a fidarsi, che modera il peso delle dimensioni di capacità, benevolenza e integrità. Un elemento centrale del modello è rappresentato dal ruolo

2.5 BIAS COGNITIVI E PERCEZIONE DEL RISCHIO NEI PAV/EVTOL

La percezione del rischio associata ai Personal Air Vehicles (PAV) e agli electric Vertical Take-Off and Landing aircraft (eVTOL) non dipende esclusivamente da valutazioni razionali e oggettive di sicurezza, ma è fortemente influenzata da bias cognitivi che intervengono nei processi decisionali in condizioni di incertezza e di novità tecnologica (Saletti; De Backer, 2026). In particolare, l'uso di scorciatoie mentali (heuristics) consente agli individui di semplificare decisioni complesse, ma può portare a distorsioni sistematiche nella valutazione del rischio percepito, soprattutto quando la tecnologia risulta poco familiare o altamente innovativa (Saletti; Manger, 2025).

Un bias cognitivo può essere definito come un "filtro" involontario che deforma la percezione della realtà per consentire una risposta rapida, spesso a scapito della precisione logica; sebbene tali meccanismi siano frequentemente funzionali, possono indurre distorsioni percettive, giudizi imprecisi o interpretazioni illogiche (Manger, 2025). Non esiste un numero univoco di bias cognitivi: le stime variano da oltre 100 a oltre 200 distorsioni documentate, organizzate in diverse categorie (Appendice B), e nuovi bias continuano a emergere in relazione all'introduzione dell'intelligenza artificiale e dei big data (Buster Benson, 2016; Saletti; Manger, 2025).

I bias cognitivi sui quali ci focalizzeremo sono pertanto quelli più direttamente connessi ai PAV e agli eVTOL. Definite da studi condotti da Airbus, dal Politecnico di Milano e dalla NASA, che hanno analizzato il comportamento di potenziali passeggeri attraverso survey, modellazione comportamentale e simulatori in realtà virtuale, evidenziando come tali distorsioni cognitive incidono in modo significativo sull'accettazione della tecnologia e sulla valutazione del rischio percepito (De Fabiis, 2025; Yedavalli & Mooberry; NASA, 2019).

AUTOMATION BIAS

L'Automation Bias si verifica quando l'utente si affida eccessivamente ai suggerimenti del sistema, ignorando o svalutando input manuali o segnali ambientali corretti (Manger, 2025). È la manifestazione pratica e l'effetto più pericoloso dell'Overtrust (fiducia eccessiva) che abbiamo trattato nel paragrafo precedente. Si può verificare durante una manovra di deconflitto: un operatore remoto approva una rotta suggerita dall'IA che lo porta in uno spazio aereo interdetto (No-Drone Zone), fallendo nel verificare le limitazioni geografiche reali (Manger, 2025).

ILLUSORY CONTROL

L'illusory control è un fenomeno in cui gli individui hanno la tendenza universale a sovrastimare la propria

capacità di influenzare eventi che sono in realtà determinati dal caso o dagli algoritmi autonomi (Manger, 2025). “Essa conduce alla convinzione soggettiva di esercitare un controllo personale su un risultato, traducendosi spesso in un'autovalutazione distorta e amplificata della propria influenza effettiva sulla situazione.” In uno studio condotto da Koo et al., se si fornisce al conducente un ampio numero di informazioni dettagliate durante la guida autonoma del veicolo, il senso di controllo dell'individuo aumenta, anche se non ha possibilità di intervento sulla situazione (Manger, 2025).

CONFIRMATION BIAS

Il Confirmation Bias descrive la tendenza intrinseca dell'individuo a selezionare, interpretare e favorire informazioni che confermano le proprie credenze preesistenti, ignorando sistematicamente le prove contrarie (Manger, 2025; Backer, 2024). Nel contesto dell'Urban Air Mobility (UAM), questo bias agisce come un 'congelatore' del modello mentale iniziale: se un utente nutre una sfiducia a priori verso l'intelligenza artificiale, tenderà a interpretare ogni minima anomalia del sistema come una prova definitiva della sua inaffidabilità. Tale fenomeno è strettamente legato all'impiego di “conoscenze scudo”, ovvero argomentazioni e razionalizzazioni

che piloti o passeggeri utilizzano per preservare le proprie comprensioni riduttive o errate del sistema, proteggendole dal confronto con la realtà tecnica (Hoffman & Bashir, 2015; Dzindolet et al.). Questo meccanismo è particolarmente evidente in soggetti con paura del volo, i quali tendono a privilegiare e ricordare selettivamente informazioni che confermano i propri timori, riducendo drasticamente la propensione all'uso della tecnologia (de Fabiis, 2025).

AVAILABILITY HEURISTIC

È un meccanismo cognitivo che domina la percezione del rischio. Porta a valutare la probabilità di un evento in base alla facilità con cui esempi pertinenti vengono richiamati alla memoria (Manger, 2025; de Backer, 2025). In un ambiente aeronautico, questo bias è alimentato dal fenomeno “dell'imaging th enumerator”. Per esempio, l'immagine vivida e spaventosa di un singolo incidente aereo (il numeratore) domina la percezione dell'utente, oscurando il dato statistico dell'enorme volume di voli sicuri (il denominatore) (Slovic et al., 2022). Questo bias è una componente fondamentale del “rischio come sentimento”, dove l'intuizione veloce sostituisce il calcolo lento e faticoso del rischio probabilistico (Slovic et al., 2022). Per un progettista di interfacce PAV, questo significa che la sicurezza

percepita non dipende dalla reale affidabilità tecnica, ma dalla gestione mediatica e visiva dei fallimenti, poiché i ricordi recenti o emotivamente carichi pesano più dei dati analitici (Slovic et al., 2022; de Backer, 2025).

ACTION BIAS

L'Action Bias rappresenta la propensione umana a preferire l'azione all'inazione, specialmente in situazioni di alto stress o incertezza, anche quando intervenire potrebbe peggiorare l'esito. Negli eVTOL di livello 3 e 4, questo bias si manifesta nel desiderio compulsivo del passeggero di riprendere il controllo manuale (takeover) durante manovre automatizzate che percepisce come rischiose, sebbene il velivolo stia operando entro i limiti di sicurezza (Manger, 2025; Omeiza et al., 2021). La ricerca indica che fornire spiegazioni in tempo reale sulle intenzioni della macchina può mitigare questo impulso, riducendo l'ansia e la necessità soggettiva di intervento fisico (Manger, 2025).

FRAMING EFFECT

L'effetto cornice dimostra come le decisioni degli individui siano influenzate dal modo in cui le informazioni vengono presentate, piuttosto che dal loro contenuto oggettivo. Ad esempio, descrivere una manovra di decollo di un eVTOL come avente una “probabilità di successo del 98%” genera molta più

fiducia rispetto a indicare un “rischio di fallimento del 2%”, nonostante i dati siano identici (Slovic et al., 2022; de Backer, 2025). Nella progettazione delle HMI (Human-Machine Interface), il framing è cruciale per la comunicazione dell’incertezza: la trasparenza non deve solo fornire dati, ma deve “incorniciarli” in modo che l’utente possa calibrare la propria fiducia senza cadere nel panico (Hoffman & Bashir, 2015; Slovic et al., 2022).

STATUS QUO BIAS

Il bias dello Status Quo emerge come la preferenza per la condizione attuale e la resistenza a decisioni irreversibili o a innovazioni radicali. Nel settore UAM, questo si traduce in una “rigidità di percorso” in cui gli utenti preferiscono i trasporti terrestri tradizionali perché percepiti come lo standard “normale”, nonostante la saturazione e l’inefficienza delle infrastrutture bidimensionali (Ararat et al., 2023; Clarke, 2019; de Backer, 2025). Gli utenti tendono a scegliere i comandi manuali più rischiosi semplicemente perché è la modalità con cui hanno più familiarità (Dzindolet et al.). Per superare questa barriera, l’industria adotta l’approccio incrementale del “rana nell’acqua bollente”, introducendo gradualmente l’automazione per evitare il rifiuto istintivo del nuovo (de Fabiis, 2025; Hu et al., 2023).

SALIENCE BIAS

Il bias di salienza descrive la propensione degli individui a dare maggior peso alle informazioni che risultano più evidenti, bizzarre o visivamente marcate, trascurando spesso dati che, seppur critici, appaiono come “normali” o attesi. In ambito aeronautico, la salienza funge da meccanismo di orientamento che determina quali porzioni dell’ambiente operativo vengono processate prioritariamente (de Backer, 2025; Endsley, 1995). Si osserva che la salienza non è determinata solo da caratteristiche fisiche (come la luminosità o il colore di un’allerta), ma anche dal valore affettivo dello stimolo (Slovic et al., 2022). Si rileva che i guasti “salienti”, come i falsi allarmi, hanno un impatto negativo sulla fiducia molto più profondo rispetto agli errori meno evidenti, come i “mancati rilevamenti” (misses). Un falso allarme richiede un’investigazione attiva e faticosa da parte dell’utente, rendendo l’errore della macchina estremamente memorabile e dannoso per il legame di fiducia (Hoffman & Bashir, 2015; Dzindolet et al; Parasuraman et al., 1999).

SICUREZZA PERCEPITA, COMFORT E PERCEZIONE DEL RISCHIO

I servizi eVTOL sono ancora in fase di sperimentazione o di introduzione limitata, l'analisi di queste dinamiche si basa prevalentemente su studi sperimentali, simulativi e di percezione sociale, che consentono di esplorare in modo sistematico le criticità introdotte.

DOPO AVER ANALIZZATO I MECCANISMI PSICOLOGICI DELLA FIDUCIA NEL CAP. 2, QUESTO CAPITOLO ESAMINA COME TALI MECCANISMI SI MANIFESTINO NELL'ESPERIENZA CONCRETA DEL PASSEGGERO ALL'INTERNO DEI PAV, PORTANDO ATTENZIONE AL RUOLO DEL CO-PILOTA AUTOMATICO, DEL PILOTA REMOTO, DELLA TRASPARENZA, DELLA UX COME MEDIATORE TRA SICUREZZA TECNICA E SICUREZZA PERCEPITA.

3.1 ASSENZA DEL PILOTA UMANO E PERDITA DELL'UTENTE AGILE

Come discusso nei Capitoli 1 e 2, l'introduzione del co-pilota automatizzato, del pilota remoto o di sistemi di volo completamente autonomi nel settore aeronautico offre indubbi vantaggi in termini di gestione della complessità operativa, riduzione del carico cognitivo e ampliamento dell'accessibilità al volo. Questi sviluppi riavvicinano l'aviazione personale all'obiettivo originario dei Personal Air Vehicles (PAV): rendere il volo accessibile a utenti non professionisti. Tuttavia, i benefici tecnologici sono accompagnati da nuove barriere di natura psicologica, sociale e culturale, in larga parte riconducibili al tema dell'agency. L'assenza del pilota umano a bordo riduce l'agency percepita dal passeggero, trasferendo il controllo decisionale a un agente algoritmico non direttamente osservabile, interrogabile o interpretabile. Questo passaggio modifica radicalmente l'esperienza utente, spostando il focus dalla conduzione tecnica del mezzo alla gestione psicologica della vulnerabilità. Quando la sottrazione di agency non è adeguatamente mediata, essa incide negativamente sul comfort percepito, aumentando l'ansia, riducendo il controllo soggettivo e amplificando la richiesta di spiegazioni e trasparenza algoritmica. La transizione verso sistemi UAM implica quindi una ridefinizione profonda dei concetti di controllo, responsabilità e affidamento, con effetti diretti sulla fiducia di passeggeri e operatori (ACI, 2022).

Dal punto di vista dei modelli di interazione uomo–automazione, l'allontanamento dell'essere umano dal ciclo di controllo attivo introduce il noto fenomeno dell'“out-of-the-loop”, associato a una riduzione della consapevolezza situazionale e della capacità di intervento in caso di anomalie (Parasuraman et al., 2001). I veicoli PAV pongono una sfida UX specifica rispetto ai trasporti terrestri. L'ambiente di volo è percepito come un “loop chiuso”, dal quale il passeggero non può allontanarsi in caso di disagio o pericolo, a differenza di autobus o treni, non è possibile spostarsi in un'altra sezione o interrompere il viaggio, accentuando il senso di impotenza e vulnerabilità (Wells, 2021).

3.1.1 EVIDENZE SPERIMENTALI: AGENCY E VULNERABILITÀ

Numerosi studi empirici confermano che l'assenza del pilota fisico rappresenta una delle principali barriere all'adozione dei sistemi UAM. La ricerca di Omeiza et al. 2021, ha testato l'impatto di spiegazioni “astratte” rispetto a quelle “specifiche”. Sebbene entrambe migliorino la percezione di sicurezza, le spiegazioni specifiche aumentano paradossalmente il desiderio di intervenire, poiché consentono all'utente di giudicare il processo

decisionale dell'IA e confrontarlo con il proprio stile di guida. È stato inoltre riscontrato che spiegazioni ricche di dettagli tecnici, anche se parzialmente irrilevanti, aumentano drasticamente la credibilità percepita del sistema sfruttando il “Truthiness Effect”: gli utenti tendono a fidarsi di ciò che appare intuitivamente vero, rendendo il design grafico e il formato delle spiegazioni un fattore di persuasione più potente della precisione algoritmica stessa. Lo studio riportato nell'articolo di Rice et al., del 2020, “A prediction model of consumer's willingness to fly in autonomous air taxis, pubblicato sul Journal of Air Transport Management”. Un'ulteriore evidenza proviene da uno studio di mercato condotto da Booz Allen Hamilton per conto della NASA (2018), che ha coinvolto circa 1.700 potenziali utenti UAM, con poche familiarità del settore (solo il 23% circa dichiarava di conoscere i progetti UAM attuali), in cinque aree metropolitane statunitensi (Los Angeles, Houston, New York, San Francisco e Washington, D.C.). Booz Allen ha mostrato uno spot commerciale di 90 secondi privo di marchio, realizzato per la Uber Elevate, lo spot mette lo spettatore nei panni di un passeggero UAM che utilizzava un servizio di ride-sharing volante per tornare a casa dal lavoro. Ai partecipanti è stato chiesto di valutare automazione, sicurezza e disponibilità all'uso dei veicoli UAM.

Analogamente a quanto rilevato da Rice e Winter, lo studio di Booz Allen Hamilton, “il 77% degli utenti, ha dichiarato che sarebbe disposto a volare su un UAM se a bordo fosse presente un pilota (o un assistente di volo), con un significativo aumento della propensione all'adozione” (Wells, 2021). Nel rapporto leggiamo che i partecipanti «hanno espresso preoccupazione per il fatto che i passeggeri a bordo possano arrecare danno agli altri» e sono emerse anche «preoccupazioni legate alle aggressioni sessuali, in particolare in uno scenario automatizzato senza equipaggio a bordo», afferma il rapporto. «È interessante notare che molti partecipanti ai focus group hanno dichiarato di non essere disposti a considerare l'uso di alcuna forma di mobilità automatizzata (ad esempio veicoli condivisi autonomi) proprio per questo motivo». Questo disagio non deriva solo dal timore di un guasto tecnico, ma anche dalla mancanza di un “garante uomo” capace di gestire scenari imprevisti o minacce da parte di altri passeggeri in uno spazio ristretto e inaccessibile (Wells, 2021).

3.2.1 EVIDENZE SPERIMENTALI: RISCHIO SITUAZIONALE E AFFIDAMENTO

Un'altra ricerca sperimentale di Hoesterey e Onnasch (2022) ha indagato come il rischio reale (la sensazione di vulnerabilità) influenzi i comportamenti di affidamento verso l'automazione sotto pressione temporale, utilizzando il I Virtual Reality Testbed for Risk and Automation Studies (ViRTS). Il rischio è stato manipolato attraverso due variabili indipendenti: Il grado di automazione, dove i soggetti sono stati suddivisi tra un gruppo con Information Automation (IA), in cui il sistema supporta solo l'analisi preliminare dei dati (Stadio 2 di Parasuraman), e un gruppo con Decision Automation (DA), in cui l'algoritmo forniva direttamente la diagnosi e suggeriva l'azione correttiva (Stadio 3) e Il Rischio Situazionale: Manipolato attraverso l'altitudine della postazione di lavoro.

Nella condizione di basso rischio, la capsula operava a 0,5 metri dal suolo; nella condizione di alto rischio, la capsula era sospesa a 70 metri di altezza, con la minaccia di una caduta virtuale (visiva e uditiva) in caso di errore diagnostico. L'analisi dei dati

ha confermato che l'automazione decisionale (DA) garantisce una performance superiore in termini di velocità e accuratezza, ma introduce un rischio latente di compiacenza. In condizioni di basso rischio, infatti, i partecipanti supportati da DA hanno mostrato un marcato Automation Bias, riducendo drasticamente il "campionamento delle informazioni" (information sampling), ovvero la verifica manuale dei parametri grezzi prima di agire. L'elemento di maggiore interesse per l'ambito della Urban Air Mobility risiede tuttavia

nell'interazione tra automazione e rischio elevato. I risultati indicano che, sotto la pressione di un rischio reale (l'altezza di 70 metri), il gruppo DA ha abbandonato il comportamento di affidamento cieco, aumentando significativamente l'attività di monitoraggio e verifica dei dati.

Mentre la fiducia soggettiva riportata nei questionari rimaneva alta in entrambe le condizioni (segno che gli utenti riconoscevano l'affidabilità del sistema), il comportamento pratico diventava più prudente e

vigile non appena aumenta la vulnerabilità percepita, presentando una discrepanza tra atteggiamento e comportamento. Questo studio ci permette di dimostrare che l'automazione avanzata (AA) migliora la performance, ma il rischio situazionale agisce come un "freno" che aumenta i tempi di reazione, probabilmente perché l'utente è più attento (Hoesterey, 2022). Questo suggerisce che la perdita di agency è accettata più facilmente quando le conseguenze del fallimento sono minime, ma viene attivamente contrastata attraverso il monitoraggio quando la sicurezza personale è in gioco (Hoesterey, 2022). Questa dipendenza non era mediata da un reale aumento della fiducia nel sistema, ma da una necessità operativa. In assenza di tempo per esercitare la propria agency, l'operatore umano si affida passivamente al sistema, aumentando il rischio di errori nel caso in cui l'automazione sia scarsamente affidabile (Rice et al., 2009). Nonostante l'automazione miri a sollevare l'utente da ogni compito, il desiderio di supervisione rimane una barriera critica. Circa il 64% degli utenti europei dichiara che si sentirebbe a proprio agio solo se avesse la possibilità di riprendere il controllo manuale (takeover) in qualsiasi momento. Questo dato sale all'83% se si considera l'importanza percepita di poter intervenire durante il volo (Allianz Motor, 2025).

La ricerca di Omeiza et al., ha testato l'impatto di spiegazioni "astratte" (es. "Rallento per un segnale stradale") rispetto a quelle "specifiche" (es. "Rallento perché il semaforo è rosso"). Entrambe migliorano la percezione di sicurezza, ma le spiegazioni specifiche aumentano paradossalmente il desiderio di intervenire, poiché permettono all'utente di giudicare il processo decisionale dell'IA e confrontarlo con il proprio stile di guida. Si è riscontrato che spiegazioni ricche di dettagli tecnici o dati statistici (anche se potenzialmente fuorvianti o irrilevanti) aumentano drasticamente la credibilità percepita del sistema sfruttando il "Truthiness Effect". Gli utenti tendono a fidarsi di ciò che appare intuitivamente vero, rendendo il design grafico delle spiegazioni un fattore di persuasione più potente della precisione algoritmica stessa (Omeiza et al., 2021).



Figura 3.1: Configurazione della simulazione di guida per lo studio. L'allestimento comprende un visore VR, un volante, i pedali del freno e dell'acceleratore, uno schermo e un sedile da simulazione (arcade seat). Lo schermo mostra un pedone che attraversa sulle strisce pedonali, fotografia dell'esperimento condotto da Omeiza et al., 2021.

3.1.3 EVIDENZE SPERIMENTALI: PILOTA COME SEGNALE DI SICUREZZA PSICOLOGICA

Le indagini sulla percezione sociale condotte dall' Agenzia dell'Unione Europea per la sicurezza aerea (EASA) e Airbus (2021) su circa 4.000 cittadini, definiscono che il 70% degli intervistati accetta il volo di air taxi se guidati da un pilota umano, mentre questa percentuale crolla al 44% per i voli senza pilota. La presenza del pilota non è solo una necessità tecnica, ma un segnale psicologico che garantisce la possibilità di un intervento immediato. La “perdita di controllo” percepita nei sistemi autonomi è identificata come la principale barriera all'accettazione, superando persino le preoccupazioni sul rumore (UAM, 2021). Gli studi qualitativi suggeriscono che persino la terminologia influenza la percezione dell'agency. È stato suggerito di utilizzare il termine “self-piloted” (auto-pilotato) invece di “autonomous” (autonomo), poiché il primo implica una pre-determinazione delle rotte, mentre il secondo evoca l'idea di una macchina che prende decisioni indipendenti e potenzialmente imprevedibili (UAM, 2021).

In sintesi, l'assenza del pilota e la perdita di agency nell'UAM è il principale ostacolo alla calibrazione

della fiducia (Alonso, 2018; Lee & See, 2004). Come evidenziato dai modelli di Luhmann (1979), la fiducia serve a ridurre la complessità sociale e sistemica, ma può funzionare solo se supportata da una struttura di comunicazione efficace che permetta all'uomo di sentirsi ancora il “centro” decisionale, anche se l'esecuzione tecnica è delegata alla macchina (Luhmann, 1979; Parasuraman, 2010; Walker, 2023). Come evidenziato nei test di Supervised autonomy condotti dalla Wisk Aero¹⁴, la trasparenza non deve limitarsi a giustificare l'azione presente, ma deve educare l'utente ai limiti del dominio operativo del velivolo ODD (Operational Design Domain), prevenendo l'effetto “Out-of-the-Loop” che rende impossibile un intervento umano efficace in caso di guasto improvviso (Hoesterey, 2022).

3.2 TRASPARENZA, EXPLAINABLE AI (XAI) E CONTROLLO PERCEPITO

La trasparenza del sistema e l'implementazione delle Explainable AI (XAI) risultano determinanti non solo per incrementare la fiducia, ma soprattutto per garantire una corretta calibrazione nei velivoli automatizzati. Secondo Chen et al., la trasparenza nei veicoli automatizzati è definita come “la qualità descrittiva di un'interfaccia riguardo alla sua

capacità di favorire la comprensione, da parte dell'operatore, degli intenti, delle prestazioni, dei piani e dei processi di ragionamento di un agente intelligente.”¹⁵ In questa prospettiva, la trasparenza consente all'utente di comprendere cosa il sistema stia facendo, perché lo stia facendo e cosa aspettarsi nel passo successivo (Chen et al., 2014).

Il modello Situation Awareness Transparency (SAT), proposto da Chen et al. (2014) sulla base della teoria della Situation Awareness di Endsley (1995), articola la trasparenza in tre livelli, su cui anche la XAI trova una solida base concettuale per agire, e sono: Il primo livello supporta la percezione dello stato corrente del sistema (“cosa sta succedendo”), includendo informazioni su obiettivi, stato e progressi. L'assenza di questo livello è associata a una ridotta consapevolezza situazionale e può favorire fenomeni di eccesso di fiducia e sfiducia. Il secondo livello, supporta la comprensione del ragionamento e dei limiti del sistema (“perché sta succedendo”), mentre il terzo livello abilita la proiezione, consentendo all'utente di anticipare esiti futuri e probabilità di successo o fallimento (“cosa succederà dopo”). L'integrazione dei livelli 2 e 3 favorisce una fiducia calibrata riducendo il rischio di compiacenza, già affrontata negli studi sperimentali di Hoesterey e Onnasch su automazione decisionale (Chen et al., 2014; Lee &

3.2.1 EVIDENZE SPERIMENTALI: TRASPARENZA NEGOZIABILE E HAT

I programmi di ricerca NASA hanno approfondito il ruolo della trasparenza nel migliorare fiducia e prestazioni in scenari aeronautici complessi, nel paradigma del Human–Automation Teaming (HAT) (Battiste et al., 2018).

Negli esperimenti 5 e 6, piloti e controllori del traffico aereo sono stati coinvolti in scenari simulati di diversione del volo, caratterizzati da cambi di rotta improvvisi e da elevata complessità decisionale. I partecipanti interagivano con l'interfaccia Autonomous Constrained Flight Planner (ACFP), progettata secondo un approccio “playbook”, che metteva a

¹⁴. “Supervised autonomy” (autonomia supervisionata) adottata da Wisk Aero, è un paradigma operativo che rappresenta una fase cruciale nella transizione verso il volo interamente autonomo. Si osserverà come questa modalità di controllo preveda che la conduzione tecnica del velivolo sia affidata a un computer di bordo, mentre un operatore remoto mantiene l'autorità di supervisione (oversight), intervenendo solo in scenari specifici o di emergenza.

¹⁵. Vernol Battiste (2018), Human-Automation Teaming: Lessons Learned and Future Directions.

disposizione un insieme strutturato di strategie operative predefinite per la gestione di eventi imprevisti nello spazio aereo. Un elemento distintivo dell'interfaccia ACFP era la visualizzazione esplicita della logica decisionale del sistema, realizzata tramite strumenti a cursore (slider bars) che rappresentavano i pesi decisionali (ACFP weights) assegnati a variabili operative critiche quali carburante residuo, tempo stimato di arrivo (ETA) e rischio operativo associato alle diverse opzioni di rotta. Questa modalità di rappresentazione rendeva il processo decisionale dell'autonomo osservabile, interpretabile e negoziabile, consentendo all'operatore di modificare attivamente le priorità del sistema.

I partecipanti hanno riportato che la possibilità di negoziare con il sistema alterando questi pesi decisionali (comunicazione bidirezionale), semplificava il processo di scelta nelle situazioni di diversione e contribuiva a una riduzione del carico di lavoro mentale.

La trasparenza, intesa come capacità del sistema di mostrare i propri "ragionamenti", ha permesso ai piloti di evitare un affidamento acritico, ma di verificare la precisione dei suggerimenti confrontandoli con i dati meteorologici reali. Un ulteriore risultato, ha dimostrato che i piloti preferiscono interfacce dove l'esecuzione finale rimane

sotto la responsabilità umana (Operator Directed Execution). Tale scelta progettuale contribuisce a trasformare l'automazione da entità decisionale opaca a collaboratore cognitivo, configurandosi come un vero e proprio membro del team piuttosto che come una "scatola nera" che impone soluzioni non trasparenti. In continuità con questi risultati, il programma di ricerca della NASA ha formalizzato il paradigma del HAT come risposta strutturale alla perdita di agency nelle Single Pilot Operations (SPO) e nei velivoli altamente autonomi (Battiste et al., 2018). Le simulazioni di diversion scenarios (DS) mostrano che, in assenza di spiegazioni sulle motivazioni alla base delle raccomandazioni automatiche, i piloti manifestano livelli più elevati di sfiducia e incertezza all'affidamento (mancanza di trasparenza).

Al contrario, l'implementazione di un modello HAT basato su tre pilastri (trasparenza, comunicazione bidirezionale ed esecuzione diretta dall'uomo) ha mostrato che i piloti preferiscono sistemi dove possono "chiamare degli schemi" (playbook approach). In questo assetto, l'agency dell'operatore non viene eliminata dall'automazione, ma riconfigurata come funzione di supervisione attiva e strategica: l'essere umano mantiene una supervisione attiva dove l'esecuzione finale dell'azione rimane sotto il suo comando diretto (Battiste et al., 2018).

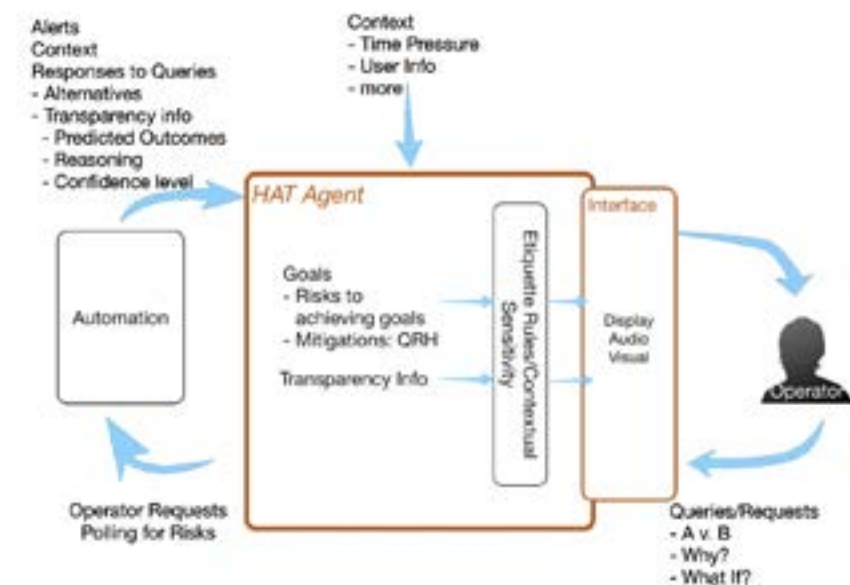


Figura 3.2: Modello iniziale dell'interazione HAT (Battiste et al., 2018).

“Da questi studi è emerso che non vi sono state differenze significative nelle valutazioni del modello HAT riguardanti la consapevolezza situazionale (situation awareness), il carico di lavoro (workload) o la fiducia (trust). Tuttavia, i partecipanti hanno espresso una preferenza netta per le condizioni basate sul sistema HAT rispetto a quelle che ne erano prive (No HAT). Inoltre, come accade per ogni tecnologia emergente, i partecipanti hanno fornito suggerimenti per migliorare l'agente HAT. Tali proposte includono: un'interfaccia vocale potenziata che utilizzi il linguaggio naturale, una migliore etichettatura dei punti di ancoraggio (anchor points) sugli strumenti a cursore (slider tools) e l'integrazione di ulteriori informazioni di supporto per i piloti.”¹⁶

La trasparenza non deve essere intesa come una semplice visualizzazione di dati, ma come una qualità descrittiva dell'interfaccia che supporta la comprensione (Chen et al., 2014). Esperimenti in domini affini, come quello della medicina predittiva, hanno dimostrato che fornire informazioni sui driver di alto livello dell'algoritmo raddoppia la capacità dell'uomo di

¹⁶ Vernol Battiste (2018), Human-Automation Teaming: Lessons Learned and Future Directions.

identificare correttamente situazioni critiche, come l'ipossiemia durante un intervento chirurgico. Questo suggerisce che, anche nell'UAM, la XAI è necessaria per calibrare le aspettative del pilota o dell'operatore a terra, riducendo la "compiacenza indotta dall'automazione" (Chen et al., 2014). Studi condotti dalla NASA nell'ambito delle operazioni con singolo pilota e velivoli autonomi mostrano che la trasparenza adattiva, combinata con comunicazione bidirezionale, migliora

significativamente la fiducia e riduce il carico cognitivo (Battiste et al., 2018). In questi scenari, il sistema non si limita a informare, ma dialoga con l'operatore, supportando la costruzione di modelli mentali aggiornati.

Figura 3.3: Integrazione dell'HAT nella cabina di pilotaggio. Fonte: Battiste et al., 2018.



3.2.2 EVIDENZE SPERIMENTALI: EXPLAINABLE AI E CONTROLLO COGNITIVO

Nell'ambito dei sistemi autonomi, l'obiettivo primario dell'Explainable AI (XIA) non consiste nel ripristinare un controllo operativo diretto da parte dell'essere umano, ma nel rafforzare il controllo cognitivo, ovvero la capacità dell'utente di comprendere, interpretare e anticipare il comportamento del sistema automatizzato. Questo aspetto risulta particolarmente rilevante per la UAM, dove il passeggero non dispone di leve di intervento diretto, ma valuta la sicurezza sulla base di segnali interpretativi, portando a un controllo consapevole (Chen et al., 2014). In scenari di autonomia condivisa, la trasparenza fornita dalla XAI mira a rendere il comportamento dell'automazione non solo osservabile, ma anche prevedibile, supportando la costruzione di modelli mentali adeguati (Chen et al., 2014). Un riferimento istituzionale centrale in questo ambito è rappresentato dal rapporto della Defense Advanced Research Projects Agency (DARPA) sulla Explainable AI (2016), secondo il quale un'interfaccia esplicitiva efficace deve essere in grado di rispondere, almeno, a cinque domande fondamentali poste dall'utente:

- 1_ Perché il sistema ha compiuto quella specifica azione e non un'altra?
- 2_ Quando il sistema ha successo nelle sue operazioni?
- 3_ Quando il sistema fallisce?
- 4_ Quando l'utente può effettivamente fidarsi del sistema?
- 5_ Come può l'utente intervenire per correggere un eventuale errore?

Tali requisiti delineano un quadro funzionale della XAI come strumento di supporto decisionale e cognitivo, piuttosto che come meccanismo di trasparenza tecnica.

Uno studio condotto da Koo et al. (2015), nel contesto della guida autonoma ha dimostrato che la fornitura di spiegazioni sul comportamento del sistema aumenta significativamente il controllo percepito, anche quando l'utente non ha alcuna

possibilità di intervento. I partecipanti esposti a spiegazioni contestuali (ad esempio, motivazioni per una frenata o una deviazione) riportavano livelli inferiori di ansia e una maggiore accettazione dell'automazione rispetto al gruppo di controllo. Risultati analoghi emergono da studi in ambito aeronautico e UAV. In un esperimento condotto da Mercado et al. (2016) su sistemi di supervisione di droni autonomi, gli operatori che ricevevano spiegazioni sulle priorità decisionali dell'AI mostravano una migliore comprensione del sistema, tempi di reazione più rapidi e una fiducia più calibrata. In questi casi, la spiegazione non migliora soltanto la fiducia dichiarata, ma il comportamento operativo, riducendo errori di affidamento. Un tema centrale nella ricerca XAI riguarda il livello di dettaglio delle spiegazioni più efficaci. Studi recenti suggeriscono che una maggiore quantità di informazioni non corrisponde necessariamente a una migliore comprensione. Lim et al. (2009) hanno dimostrato che spiegazioni eccessivamente tecniche possono aumentare il carico cognitivo e generare confusione, soprattutto in utenti non esperti. Al contrario, nel già citato studio di Omeiza et al., evidenzia che spiegazioni specifiche e concrete aumentano la credibilità percepita del sistema, ma possono anche incrementare il desiderio di intervento (takeover),

poiché l'utente si sente legittimato a valutare criticamente la decisione dell'AI. Questo risultato evidenzia una tensione progettuale cruciale per la UAM: le spiegazioni devono essere sufficienti ricche da rassicurare e informare, senza tuttavia incentivare una dinamica di controllo competitivo tra uomo e macchina.

Contrariamente alla percezione diffusa, l'obiettivo della XAI non è rendere l'utente consapevole della struttura matematica dei modelli di machine learning, spesso poco utile e impraticabile per l'utente finale (Benda et al., 2021). Piuttosto, una trasparenza efficace si fonda sulla capacità del sistema di comunicare la propria razionalità operativa, ossia i criteri che guidano le decisioni dell'automazione (Hoffman, 2015; Lee e See, 2004).

Questo approccio si colloca nel secondo livello del modello Situation Awareness-based Transparency (SAT), che mira a supportare la comprensione del ragionamento e dei limiti dell'agente intelligente (Chen et al., 2014). Come discusso nel Capitolo 2, la fiducia nei sistemi automatizzati non dipende dalla conoscenza del codice, ma dalla percezione delle prestazioni, del processo decisionale e dello scopo del sistema (Lee & See, 2004).

In questo quadro, fornire informazioni sui "driver di alto livello" (ad esempio i pesi attribuiti a variabili quali

condizioni meteorologiche, stato della batteria o densità del traffico) consente all'operatore di sviluppare un modello mentale più accurato del comportamento del velivolo (Benda et al., 2021). La trasparenza garantita dalla XAI rappresenta pertanto uno strumento fondamentale per evitare i due estremi della fiducia mal calibrata: misuse (overtrust) e disuse (undertrust). Studi sperimentali mostrano inoltre che Le interfacce XAI che utilizzano la verbalizzazione, traducendo i dati dei sensori in linguaggio naturale, migliorano l'usabilità e favoriscono la costruzione di modelli mentali validi, riducendo l'ansia e aumentando la percezione di controllo (Chen et al., 2014; Lee & See, 2014; Alonso et al., 2018; Arora & Tillieux, 2025). In contesti di alta criticità, l'efficacia della XAI nel migliorare le prestazioni emerge con particolare evidenza in studi dove l'interfaccia sovrappone le previsioni algoritmiche ai driver di rischio specifici (Benda et al., 2018). In tali sistemi, la quantificazione dell'importanza dei predittori (ad esempio attraverso tecniche di post-elaborazione come i valori SHAP) consente agli operatori di identificare correttamente le anomalie con una precisione significativamente superiore rispetto ai sistemi opachi di tipo "scatola nera" (Benda et al., 2018).

3.2.3 EVIDENZE SPERIMENTALI: CONTROLLO PERCEPITO E CALIBRAZIONE

Tuttavia, la trasparenza e gli strumenti di controllo risultano efficaci solo se riescono a integrare il controllo percepito, definito come perceived behavioral control over automation, formalizzato da Walker et al. (2023) come estensione della Theory of Planned Behavior di Ajzen (1991) nel contesto dei sistemi automatizzati. Tale costrutto non si limita alla capacità tecnica di governare direttamente un sistema, ma riguarda l'aspettativa soggettiva dell'utente che l'automazione operi in modo corretto e, soprattutto, che sia possibile riprendere il comando qualora lo si ritenga necessario o desiderato. Lo studio di Walker et al. (2023) aggiunge un tassello fondamentale al modello di Lee e See (2004): non basta "fidarsi" del sistema (atteggiamento), l'utente ha bisogno anche "sentirsi in grado di intervenire" (controllo percepito).

Questa dinamica è stata indagata attraverso uno studio sull'utilizzo del Cruise Control Adattivo (ACC). I partecipanti sono stati monitorati nel corso di dieci sessioni di utilizzo, analizzando l'evoluzione della fiducia dinamica e del controllo percepito (Walker et. al., 2023). E' stato dimostrato che la fiducia dinamica cresce rapidamente nelle prime fasi di interazione; tuttavia, il controllo percepito aumenta in modo significativo solo quando il sistema rende esplicite le proprie condizioni limite. Ad esempio, la comunicazione chiara di scenari in cui l'ACC non è in grado di rilevare correttamente il veicolo che precede (come curve strette o situazioni di occlusione sensoriale) consente all'utente di sviluppare un modello mentale accurato, anticipando quando l'automazione potrebbe fallire e quando sarà necessario attivare un processo di error recovery (Walker et. al., 2023).

In questo quadro, Walker et al. evidenziano inoltre come i livelli di automazione di Parasuraman influenzano direttamente il controllo percepito. In particolare, i sistemi collocati al livello 7, che operano in autonomia ma informano costantemente l'utente, sono percepiti come più affidabili e generano una fiducia meglio calibrata rispetto ai sistemi di livello 10, che agiscono senza fornire feedback

esplicativi. La differenza non risiede nella performance tecnica, ma nella sensazione di consapevolezza e supervisione garantita all'utente, che riduce il rischio di esclusione cognitiva dal processo decisionale (Walker et al., 2023).

L'esperimento sull'Illusory Control, condotto dalla Manger et al. (2025), che abbiamo trattato nel capitolo 2.5 come bias cognitivo, va a indagare la tendenza degli utenti a sopravvalutare la propria capacità di influenzare situazioni che sono, in realtà, gestite interamente dall'automazione. Il "Controllo Illusorio" è un fenomeno psicologico per cui un individuo crede soggettivamente di avere il controllo personale su un risultato determinato dal caso o da fattori esterni. Nel contesto della guida autonoma, si manifesta quando il conducente crede di avere il controllo della situazione anche durante una corsa completamente automatizzata.

Lo studio ha coinvolto 34 persone (14 femmine e 20 maschi), l'età andava dai 24 ai 66 anni di età. il 69% erano dipendenti, 16% studenti, 9% in pensione e 6% non lavoratori. the group tended to be experienced drivers as 84% of the participants drive daily or multiple times a week. Overall, the participants rated their prior knowledge of automated. Ai partecipanti sono stati mostrati 8 video di situazioni di guida reale

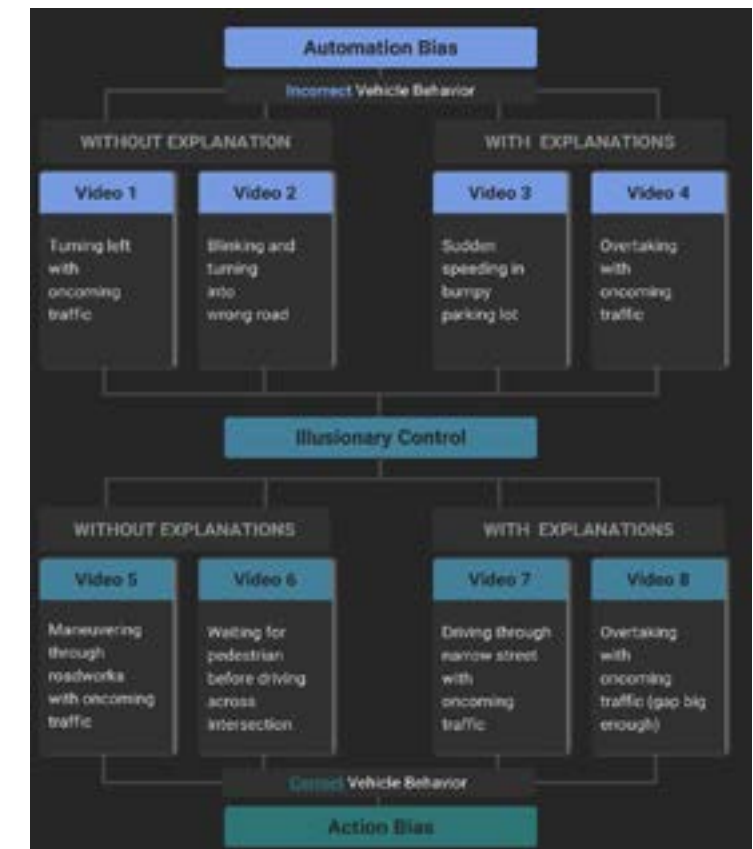


Figura 3.4: Panoramica della ricerca di Manger sui Bias Cognitivi nell'interazione dell'utente con veicoli automatizzati degli otto scenari video presentati ai partecipanti. Metà dei video mostrava un comportamento errato del veicolo per testare il Bias di Automazione, mentre l'altra metà mostrava un comportamento corretto del veicolo per testare il Bias di Azione. Il Controllo Illusorio è stato testato con tutti i video chiedendo informazioni sul controllo percepito (Fonti: Manger et al., 2025).

effettuate con il sistema Full Self.Driving (FSD) di Tesla. In tutti i video, la guida era completamente automatizzata e i partecipanti non avevano alcuna possibilità reale di intervenire o influenzare il veicolo. Dopo ogni video, i partecipanti dovevano valutare l'affermazione "Sento di avere il controllo della situazione" su una scala da -2 (fortemente in disaccordo) a +2 (fortemente d'accordo). Un punteggio alto indicava la presenza del bias di controllo illusorio. I ricercatori hanno voluto verificare se due fattori influenzassero questo bias, La Spiegabilità (explainability) metà del video includeva un interfaccia (HMI) che spiegava a voce e con testo le manovre dei veicoli, mentre l'altra metà mostrava solo una mappa base. Altre al Illusion control sono stati testati anche automation bias e action bias, che non hanno portato risultati rilevanti durante lo studio, come l'illusory. (probabilmente a cause del carico cognitivo, dell'assenza di stress elevato e della natura simulata degli scenari. I risultati relativi a questo effetto si riferiscono alla percezione generale dei partecipanti sul tema della guida automatizzata. piuttosto che a specifiche interazioni o

comportamenti del veicolo.

Nel complesso, i partecipanti hanno percepito un senso di Controllo piuttosto neutro, senza alcuna indicazione di Controllo Illusorio. Tuttavia, i partecipanti si sono sentiti leggermente più in controllo quando è stato presentato il concetto spiegabile (M = 1,06, DS = 0,90) rispetto al concetto di base (M = 0,70, DS = 0,89). Anche in questo caso, il senso di controllo percepito differiva significativamente tra le situazioni specifiche (F(7) = 11,15, p < 0,001). Il video 7, in cui il veicolo stava guidando in una strada stretta con traffico in arrivo (concetto spiegabile), ha prodotto il senso di controllo più basso (M = -1,50, DS = 1,48). Il senso di controllo più alto è stato percepito nel video 1, in cui il veicolo stava svoltando a sinistra a un incrocio con traffico in arrivo senza spiegazioni

(M = 1,46 DS = 0,94). La precisione del modello mentale del partecipante (quanto ne sapeva di tecnologia) non ha influenzato il senso di controllo illusorio sui risultati ottenuti. In pratica, ricevere informazioni sul perché l'auto si muove in un certo modo fa sentire l'utente più "al comando", anche se non lo è affatto, andando a ridurre lo stress e il senso di vulnerabilità. Questa trasparenza è un'arma a doppio taglio, perché informazioni troppo tecniche o statiche possono "abbagliare" l'utente, e il sistema fornisce una spiegazione (anche se il comportamento è errato), l'utente tende a fidarsi di più perché l'informazione è persuasiva, evidenziando un rischio di rafforzamento involontario della disinformazione. Nel complesso, i risultati suggeriscono che trasparenza ed explainability, se non progettate con attenzione, possono avere effetti

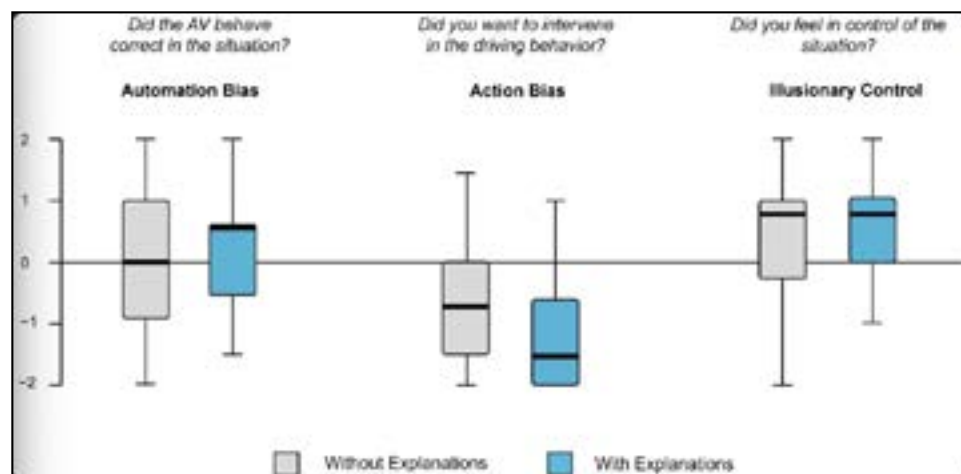


Figura 3.5: Panoramica della percezione soggettiva del comportamento del veicolo autonomo (valutazione del video) in base alle valutazioni dei partecipanti (Fonti: Manger et al., 2025).

persuasivi indesiderati.

Ciò implica la necessità di un approccio progettuale che bilanci informazione, incertezza e limiti del sistema, e di modelli teorici che integrino il ruolo dei fattori situazionali e cognitivi nella fiducia verso i sistemi di autonomi.



Figura 3.6: Gli otto scenari di guida automatizzata presentati nell'esperimento. Metà degli scenari è stata integrata con spiegazioni sul comportamento dell'automazione (colonna di destra). L'assegnazione dei video ai Bias Cognitivi è illustrata schematicamente nella Figura 3.4 (Fonti: Manger et al., 2025).

3.3 UX COME MEDIATORE TRA SICUREZZA TECNICA IN SICUREZZA PERCEPITA

Nel paradigma della Urban Air Mobility (UAM), la sicurezza tecnica costituisce una condizione necessaria ma non sufficiente all'accettazione del servizio. Poiché l'utente finale non dispone né degli strumenti cognitivi né delle competenze tecniche per valutare i complessi protocolli di affidabilità, ridondanza e certificazione che caratterizzano i velivoli eVTOL, tali proprietà non risultano direttamente accessibili all'esperienza. La sicurezza tecnica diviene pertanto comprensibile, valutabile e accettabile solo attraverso la mediazione della User Experience (UX). In questo contesto, la sicurezza percepita non emerge dalla sicurezza tecnica in sé, bensì dalla sua rappresentazione esperienziale. Come discusso nei paragrafi precedenti, la perdita di agency associata all'assenza del pilota umano (3.1) e la necessità di strumenti di trasparenza e controllo cognitivo (3.2) collocano l'interazione uomo-sistema al centro del processo di costruzione della fiducia. La UX si configura, dunque, come il principale canale di traduzione tra le proprietà oggettive del sistema e la loro interpretazione soggettiva da parte dell'utente. È pertanto evidente che l'adozione di architetture ridondanti e algoritmi avanzati non si traduce

automaticamente in accettazione, se non è accompagnata da una percezione di sicurezza individuale (Yang & Kim, 2024). In questa prospettiva, l'Intelligenza Artificiale non deve essere considerata esclusivamente come componente tecnica, ma come variabile esperienziale che può sostenere o compromettere l'agency percepita. Attraverso un'adeguata progettazione dell'interazione, la condizione passiva del trasporto autonomo può essere riconfigurata in una supervisione consapevole, anche in assenza di controllo operativo diretto (Battiste et al., 2018; Lee & See, 2004). La transizione verso i velivoli eVTOL richiede inoltre il superamento di un paradosso aeronautico consolidato: sebbene il volo rappresenti statisticamente una delle modalità di trasporto più sicure, esso continua a generare risposte emotive intense, amplificate da bias cognitivi e dall'elevata salienza mediatica dei rari eventi avversi (Graziato, 2026). In tale scenario, la UX interviene come meccanismo di riduzione della complessità (Luhmann, 1978), rendendo accessibile e "vivibile" la densità algoritmica del sistema. In questa prospettiva, l'obiettivo del Trust Design non è quello di eliminare la percezione di rischio, ma di evitare una sospensione dell'agency, consentendo all'utente di interpretare l'automazione come un'estensione delle proprie capacità cognitive,

CHI È RESPONSABILE DI QUESTA MEDIAZIONE E COME AVVIENE, DAL PUNTO DI VISTA PROGETTUALE?

piuttosto che come un'entità opaca e incontrollabile (Lee & See, 2004; Hoffman, 2015). La fiducia diventa così l'elemento che rende possibile la delega del controllo operativo alla macchina, preservando una forma di autorità decisionale di alto livello (Battiste et al., 2018). Affinché la UAM possa affermarsi come sistema di mobilità accettabile, l'AI non deve essere percepita come un agente estraneo, ma come una componente governabile dell'esperienza. Nel dominio aeronautico persiste una distinzione rilevante tra affidabilità effettiva (actual trustworthiness) e fiducia percepita (perceived trustworthiness). Mentre la prima è garantita da sistemi rigorosi (come l'integrazione tra sensori inerziali e GNSS per il calcolo dei Protection Levels (HPL/VPL)) la seconda è fortemente influenzata dal cosiddetto "risk as feeling", un processo valutativo intuitivo che privilegia la risposta emotiva rispetto al calcolo statistico. In questo quadro, la UX non rappresenta uno strato estetico o accessorio, ma un'architettura cognitiva che consente al passeggero di calibrare la propria vulnerabilità all'interno di una situazione caratterizzata da incertezza. Attraverso la mediazione dell'agency percepita, la UX rende possibile una forma di fiducia funzionale, condizione imprescindibile per l'accettazione dei sistemi di Urban Air Mobility (Hoffman et al., 2015).

3.3.3 UX COME TRADUTTORE

La funzione di traduzione della UX consiste nel rendere osservabile e prevedibile il comportamento del sistema, superando l'opacità intrinseca dei modelli di deep learning. Essa converte gli output opachi dell'IA in "segnali" (cues) leggibili, permettendo all'utente di formare un modello mentale accurato del comportamento del velivolo. In questo processo, l'interfaccia non è un semplice supporto visivo, ma il medium attraverso cui la fiducia si evolve tramite interpretazioni di tipo analitico, analogico e affettivo delle capacità dell'automazione (Lee & See, 2004). Poiché l'osservazione diretta dei processi algoritmici è spesso impossibile, la percezione dell'affidabilità è mediata dal display; ciò implica che l'appropriatezza della fiducia (ovvero la

corrispondenza tra l'affidamento dell'utente e le reali capacità del mezzo) dipenda strettamente dal contenuto e dal formato dell'interfaccia. La UX deve quindi tradurre i "high-level-drivers" (i pesi decisionali) lungo le dimensioni del dettaglio (performance e scopo) e dell'astrazione (specificità del controllo), trasformando variabili tecniche in informazioni significative (Zhang et al., 2021; Lee & See, 2004).

Questa dinamica di traduzione è stata indagata con successo nelle simulazioni NASA sul sistema Autonomous Constrained Flight Planner (ACFP). La resa visiva della logica decisionale ha permesso agli operatori di non fidarsi ciecamente del sistema, ma di convalidare le raccomandazioni confrontandole con dati reali, raddoppiando l'accuratezza diagnostica in situazioni critiche (Battiste et al., 2018). Attraverso l'integrazione della Explainable AI (XAI), il sistema comunica non solo "cosa" sta facendo, ma il "perché" e il "cosa farà in seguito", sfruttando quello che la letteratura definisce come "Truthiness Effect" (effetto veridicità). Gli esperimenti di Lee e See (2004) confermano che segnali superficiali (surface features), come un layout bilanciato, l'uso di colori freddi o la presenza di fotografie reali, possono indurre fiducia indipendentemente dalle reali capacità del sistema, poiché gli utenti tendono

a rispondere emotivamente a ciò che "sembra vero" a livello intuitivo. Negli studi sull'automazione a supporto del rilevamento di bersagli (target detection). L'aumento del realismo delle immagini ha aumentato la fiducia e ha portato a una maggiore dipendenza dalle informazioni di guida (cueing). Allo stesso modo, la tendenza dei piloti a seguire ciecamente i consigli del sistema aumenta quando l'ausilio include immagini dettagliate come guida (Ockerman, 1999). Così come immagini altamente realistiche possono aumentare la fiducia, immagini degradate possono diminuirla.

Tuttavia, la traduzione deve essere onesta: una trasparenza mal calibrata o l'uso di caratteristiche puramente estetiche potrebbero essere percepiti come manipolatori, degradando l'usabilità o inducendo all'accesso di fiducia (compiacenza). Al contrario, è possibile utilizzare strategicamente la degradazione della qualità dell'immagine o del parlato sintetico per calibrare la fiducia verso il basso quando l'automazione opera in condizioni di incertezza (Lee & See, 2004). La UX come traduttore deve considerare la dimensione sociale. Creare un'interfaccia che agisca come un partner conversazionale può indurre una fiducia simile a quella interpersonale (faccia a faccia), a patto che i parametri vocali (intonazione,

frequenza, coerenza del parlato) siano allineati al contenuto e alle aspettative dell'utente. Sebbene l'uso di segnali sociali e immagini salienti evochi risposte affettive più potenti dei dati astratti (Finucane et al., 2000), una traduzione scorretta può generare false aspettative o aumentare paradossalmente il desiderio di intervento manuale (takeover), qualora le decisioni dell'IA non siano perfettamente allineate allo stile cognitivo del passeggero (Walker, 2023; Omeiza & Manger, 2025).

3.3.2 UX COME REGOLATORE

La UX gestisce l'equilibrio dinamico tra misuse (Over-Trust) e disuse (Under-Trust), garantendo che l'utente rimanga nel ciclo di controllo (in-the-loop). Il concetto cardine è il "controllo comportamentale percepito sull'automazione" (perceived behavioral control over automation), ovvero l'aspettativa dell'utente di poter riprendere il comando qualora necessario o desiderato (Walker, 2023; Lee e See, 2004).

La capacità percepita di gestire errori o malfunzionamenti costruisce una componente integrante della calibrazione della fiducia. La UX deve facilitare la transizione di controllo, garantendo che il conducente o operatore remoto disponga del tempo e delle informazioni necessarie per la ripresa del DDT (Dynamic Driving Task) (Walker, 2023).

I risultati dello studio evidenziano che la fiducia e l'intenzione d'uso non sono influenzate esclusivamente dalle prestazioni oggettive del sistema, ma risultano fortemente condizionate dalla "propensione verso la tecnologia". Tale costrutto è emerso come il fattore con il coefficiente d'influenza più elevato (0.838) e statisticamente significativo (p-value < 0.001) nel determinare l'accettazione del servizio. La UX deve quindi operare come un regolatore adattivo, fornendo interfacce intuitive che valorizzano la competenza percepita del soggetto e riducono l'ansia tecnologica (technophobia).

Un'evidenza sperimentale coerente con questa prospettiva proviene dall'esperimento di Rice et al., che verificare come la pressione temporale rappresenti un fattore determinante nella

rottura di tale equilibrio: quando il tempo disponibile per la decisione si riduce, gli utenti tendono a manifestare una dipendenza forzata dall'automazione (automation dependency), delegando il compito al sistema indipendentemente dalla fiducia effettivamente riposta in esso (Rice et al., 2009). In scenari caratterizzati da elevata pressione temporale, l'utente smette di agire come supervisore critico e assume un ruolo passivo, rendendo l'interfaccia l'unico presidio contro l'accettazione acritica di eventuali errori del sistema. I dati ViRTRAS mostrano infatti che le "preoccupazioni per il volo" (Flying Concerns) e le "preoccupazioni per la sicurezza UAM" (UAM Safety Concerns) agiscono come potenti inibitori, con impatti negativi significativi sull'intenzione d'uso (coefficienti rispettivamente di -0.380 e -0.366) (De Fabiis, 2025; Yang e Kim, 2025).

La sperimentazione dimostra inoltre che la capacità percepita di gestire errori o malfunzionamenti è una componente integrante della calibrazione: quando l'utente percepisce di avere gli strumenti per monitorare il sistema, la sua disponibilità a delegare il controllo aumenta. In particolare, lo studio evidenzia come la percezione della sicurezza sia legata non solo alla tecnologia del velivolo, ma alla fiducia nel fornitore del servizio e nella

gestione dell'infrastruttura. In questo senso, il compito della progettazione è facilitare la transizione di controllo, garantendo che il passeggero disponga delle informazioni necessarie per mitigare la percezione del rischio. Attraverso l'uso di ambienti di realtà virtuale e simulatori avanzati, il progetto ViRTRAS sottolinea l'importanza di profilare gli utenti in base alla loro eterogeneità psicografica, assicurando che la UX possa modulare la comunicazione del rischio e dei benefici al fine di trasformare lo scetticismo in un'accettazione consapevole del servizio.

3.3.3 UX COME NARRATORE

In questo contesto, la UX inquadra il velivolo, come un attore all'interno di un contesto sociale e normativo (Cheung & Biermann, 2020); la narrazione del trust si fonda sulla percezione dello scopo e del beneficio pubblico (Lee & See, 2004). La narrazione inizia con l'onboarding, che deve anticipare le ansie dell'utente fornendo trasparenza sulle rotte, sul meteo e sulla sicurezza dei vertiporti (Obuobie, 2025). Un ruolo centrale in questa strategia narrativa è affidato all'antropomorfizzazione dell'interfaccia, i cui effetti sulla fiducia sono stati documentati sperimentalmente attraverso il sistema "Iris". Nello studio di Waytz et al. (2014), i partecipanti

IN CHE MODO LA USER EXPERIENCE PUÒ MEDIARE LA PERDITA DI AGENCY NEI PAV AUTONOMI PER TRASFORMARE LA SICUREZZA TECNICA IN SICUREZZA PERCEPITA?

sono stati sottoposti a una simulazione di guida autonoma confrontando un veicolo puramente "agente" (capace di controllo autonomo ma privo di identità) con uno dotato di caratteristiche umane: un nome (Iris), un genere (femminile) e una voce umana per comunicare le intenzioni del sistema. I risultati della sperimentazione hanno evidenziato che l'aggiunta di questi segnali antropomorfi non altera solo la percezione superficiale, ma trasforma il modello mentale dell'utente, che tende ad attribuire al velivolo reali capacità di pensiero e pianificazione (capacità cognitive). Questo "ponte sociale" si traduce in un incremento misurabile della fiducia su tre livelli: il primo è quello comportamentale e psicologico, gli utenti riportano livelli di fiducia e sicurezza significativamente più alti quando interagiscono con un'interfaccia umanizzata. Il secondo è quello Fisiologico, dove durante eventi critici e stressanti, come un incidente inevitabile causato da terzi, i soggetti che interagiscono con "Iris" mostrano una risposta fisiologica attenuata, con un minore aumento della frequenza cardiaca e una reazione di trasalimento (startle response) ridotta rispetto a chi utilizza un veicolo privo di voce o identità.

Il terzo livello si presenta in caso di collisione, l'antropomorfizzazione agisce come un mitigatore del bias di colpa; i passeggeri tendono a incolpare meno il sistema e i suoi progettisti per esiti negativi chiaramente causati da agenti esterni, riconoscendo al velivolo una forma di competenza consapevole (Waytz et al., 2014).

Tuttavia, come evidenziato dai dati EASA, l'accettazione sociale non dipende solo dall'interfaccia, ma è strettamente legata all'interesse generale: i casi d'uso per il beneficio della comunità (trasporto medico d'urgenza, soccorso) godono di un supporto superiore rispetto ai voli privati. La UX deve quindi "narrare" questa utilità sociale integrando segnali antropomorfi rassicuranti con protocolli di sicurezza visibili (EASA, 2024; Obuobie, 2025), garantendo che l'attribuzione di una "mente" al sistema non porti a false aspettative ma a una delega consapevole e sicura.

3.4 SINTESI CRITICA: CONDIZIONI PER LA FIDUCIA NELL'UAM

L'analisi condotta in questo capitolo mostra che, nel contesto della Urban Air Mobility, la fiducia del passeggero non può essere interpretata come semplice "accettazione" dell'affidabilità tecnica del velivolo. Al contrario, essa emerge come esito di un processo di costruzione esperienziale in cui la sicurezza tecnica, pur essendo condizione necessaria, resta per l'utente largamente inaccessibile sul piano cognitivo e valutativo.

La transizione verso PAV/eVTOL autonomi introduce infatti una frattura specifica: la delega del controllo operativo a un agente algoritmico riduce l'agency percepita, producendo vulnerabilità e un incremento del rischio come esperienza soggettiva (risk-as-feeling), soprattutto in un ambiente "chiuso" e non interrompibile come il volo. Le evidenze sperimentali e di percezione sociale analizzate (sez. 3.1) convergono nel mostrare che l'assenza del pilota umano non è letta solo come un cambiamento tecnico, ma come perdita di un garante sociale e situazionale, capace di intervenire e gestire l'imprevisto. Tale garante svolge una funzione psicologica di stabilizzazione: riduce l'incertezza, aumenta la percezione di protezione

e sostiene la disponibilità all'adozione. Quando questa figura viene rimossa, i passeggeri ricercano sostituti funzionali della stessa funzione: segnali di supervisione, feedback, possibilità di intervento (anche solo percepito), chiarezza dei limiti e delle condizioni di operatività. In questo scenario, la trasparenza e la XAI (sez. 3.2) assumono un ruolo cruciale non perché restituiscano controllo operativo, spesso impossibile per il passeggero, ma perché consentono la costruzione di controllo cognitivo: comprendere cosa sta facendo il sistema, perché lo sta facendo e cosa è ragionevole aspettarsi. Il punto centrale non è "spiegare la tecnologia", ma rendere l'automazione leggibile e prevedibile per un utente non esperto.

Tuttavia, la letteratura evidenzia anche una tensione progettuale: spiegazioni troppo persuasive o eccessivamente dettagliate possono incrementare credibilità e rassicurazione, ma anche generare effetti collaterali (overtrust, desiderio di takeover, illusione di controllo). Ne deriva che trasparenza ed explainability non sono "aggiunte" informative, bensì leve progettuali da calibrare in funzione del rischio situazionale e del profilo di vulnerabilità dell'utente. Alla luce di queste dinamiche, la UX (sez. 3.3) si configura come meccanismo di mediazione tra sicurezza tecnica

e sicurezza percepita: traduce l'opacità algoritmica in cues comprensibili, regola la calibrazione della fiducia attraverso segnali di supervisione e limiti comunicati, e costruisce una narrazione di scopo e legittimità sociale che sostiene l'accettazione. In altri termini, la UX non "aumenta" semplicemente la fiducia: la rende funzionale, evitando tanto la sospensione totale dell'agency (disuse) quanto l'affidamento acritico (misuse).

Questo passaggio è particolarmente rilevante per i sistemi UAM, dove l'utente finale non può verificare direttamente le proprietà di affidabilità e certificazione, ma valuta la sicurezza attraverso la qualità dell'interazione, la prevedibilità e la possibilità percepita di recupero in caso di errore. Questa sintesi conduce direttamente al Capitolo 4: se la fiducia è un esito esperienziale mediato da agency e controllo cognitivo, allora l'utente diventa il luogo in cui la fiducia si manifesta. Paure, resistenze, bisogni e aspettative non sono semplici "opinioni" o tratti personali, ma segnali osservabili degli effetti dell'autonomia sull'esperienza. Ne consegue che progettare la fiducia nei PATS non significa segmentare utenti in modo descrittivo, ma comprendere quali configurazioni di trasparenza, agency e comunicazione producono specifiche dinamiche di calibrazione. Il capitolo successivo approfondirà quindi la prospettiva dell'utente come risultato concettuale delle interazioni discusse, traducendo tali risultati in implicazioni per il design della fiducia nei sistemi UAM.

SAFETY TRANSPA- RENCY TRUST FRAMEWORK NEI PAV

L'utente come luogo in cui la fiducia si manifesta

Le conclusioni tratte nei capitoli precedenti e i profili di utenti qui delineati devono essere intesi come esiti di un'analisi esplorativa, non generalizzabile in senso statistico, ma orientata alla comprensione di dinamiche ricorrenti nella calibrazione della fiducia. E' stato, inoltre, dimostrato che la fiducia nei sistemi autonomi non coincide con la sicurezza tecnica certificata, ma emerge come esito di un processo cognitivo ed emotivo situato. La fiducia non è una proprietà intrinseca del sistema, bensì un costrutto dinamico che prende forma nell'interazione tra utente, tecnologia e contesto operativo.

Nel caso dei Personal Aerial Vehicles (PAV) autonomi, tale dinamica risulta amplificata da tre fattori strutturali: l'assenza del pilota umano come riferimento simbolico di sicurezza, l'elevata vulnerabilità percepita associata al volo e l'opacità decisionale dell'intelligenza artificiale. Come discusso nel modello di Lee e See (2004), la fiducia è strettamente connessa a condizioni di incertezza e vulnerabilità. Nel trasporto aereo autonomo, tali condizioni non sono contingenti, ma strutturali: il passeggero si trova in un ambiente chiuso e tridimensionale, con possibilità limitate di intervento diretto. La perdita di agency percepita rappresenta quindi una variabile centrale nella costruzione della fiducia. In questo scenario, la distanza tra safety tecnica e safety percepita diventa il nodo progettuale primario. Anche in presenza di elevati standard ingegneristici, l'utente può sperimentare ansia, sfiducia o, al contrario, eccessivo affidamento. Comprendere come differenti utenti calibrano la propria fiducia diventa pertanto essenziale per la progettazione dell'esperienza. In questa prospettiva,

l'eventuale criticità dell'UAM non risiede esclusivamente nei limiti tecnologici, ma nei possibili disallineamenti tra sistema e esperienza. L'utente non rappresenta il limite del sistema, bensì il luogo in cui la fiducia si manifesta e si stabilizza o si rompe in funzione della qualità della mediazione progettuale.

4.2 STIPOLOGIE DI UTENTI COME CONFIGURAZIONE PSICOLOGICHE DI CALIBRAZIONE DELLA FIDUCIA

Le tipologie di utenti qui delineate non devono essere interpretate come segmentazioni demografiche o psicografiche in senso tradizionale, bensì come configurazioni psicologiche ricorrenti attraverso cui si manifesta la calibrazione della fiducia nei sistemi autonomi ad alta criticità, quali i Personal Aerial Vehicles (PAV) e gli eVTOL. Esse rappresentano configurazioni dinamiche di affidamento, non tratti stabili della personalità, e descrivono modalità attraverso cui l'utente gestisce l'interazione con l'automazione in condizioni di incertezza e vulnerabilità. In coerenza con la definizione di fiducia proposta da Mayer, Davis e Schoorman (1995) come disponibilità ad accettare una condizione di vulnerabilità, nel contesto UAM tale esposizione risulta

amplificata dall'assenza del pilota umano e dalla delega decisionale all'Intelligenza Artificiale. Le quattro configurazioni non rappresentano "tipi di persona", ma strutture cognitivo-affettive attraverso cui l'utente interpreta il rischio, valuta l'Ability, l'Integrity e la Benevolence del sistema e decide se affidarsi o meno ad esso. Esse riflettono differenti modalità di gestione della Propensity to Trust, del controllo percepito e dei fenomeni di overtrust e undertrust descritti nella letteratura sulla Trust in Automation. In questa prospettiva, le tipologie devono essere lette come strumenti analitici per il design. Non classificano gli utenti, ma mappano le condizioni in cui la fiducia può risultare calibrata, fragile, eccessiva o difensiva, offrendo al progettista un modello interpretativo orientato alla mediazione tra safety tecnica e safety percepita.

4.2.1 UTENTE RAZIONALE ANALITICO

Questa configurazione si manifesta quando la fiducia è regolata prevalentemente da processi cognitivi analitici (dual-process theory). La delega non viene concessa a priori, ma costruita attraverso verifica logica, coerenza procedurale e comprensione del funzionamento del sistema. Nel contesto dei PAV/eVTOL, l'assenza del pilota umano non rappresenta

di per sé una minaccia; ciò che genera diffidenza è l'opacità del processo decisionale algoritmico. La fiducia aumenta quando l'automazione rende visibile la propria Ability tecnica e garantisce coerenza tra comportamento e logica dichiarata (Integrity).

MECCANISMI DELLA FIDUCIA

Ricerca di comprendere il "come" (process) e il "perché" (purpose) tecnico dietro ogni manovra del velivolo e tende a spostarsi rapidamente dalla domanda "perché dovrei fidarmi?" a "in quali casi posso fidarmi?". La sua fiducia cresce quando comprende gli algoritmi che governano il volo e osserva una coerenza tra la manovra eseguita e la logica dichiarata dal sistema. Non sono sufficienti dichiarazioni generiche di sicurezza, in questo contesto, la fiducia emerge quando l'automazione appare prevedibile e sistematica, non semplicemente avanzata.

COSA LO RASSICURA

Lo rassicura la visualizzazione dei driver decisionali di alto livello (ad esempio pesi relativi a meteo, batteria, traffico), che consente la costruzione di una matrice situazionale di sicurezza. È rassicurato da una trasparenza strutturale che espliciti stato attuale, intenzione futura del sistema e motivazione sintetica dell'azione. Comprendere perché il sistema ha preso una decisione lo rassicura più di una prestazione perfetta ma opaca: per questa tipologia, la coerenza è più importante del comfort.

COSA LO DESTABILIZZA

Lo destabilizza l'opacità decisionale (scatola nera), la pressione temporale e le discrepanze tra il comportamento del velivolo e il proprio modello mentale. Spiegazioni vaghe o meramente rassicuranti generano sospetto e possono produrre una percezione di perdita di agency, conducendo a un rifiuto critico oppure a una dipendenza forzata e ansiosa [3.2.2]. Questo utente è particolarmente sensibile alla manipolazione percettiva.

RUOLO DELLA UX

La UX deve agire come traduttore logico-tecnico. È efficace quando offre moduli di trasparenza immediata ("a colpo d'occhio"), comunica non solo ciò che il sistema fa correttamente

ma anche le condizioni in cui le prestazioni possono variare, supporta la costruzione del modello mentale attraverso micro-onboarding, spiegazioni contestuali e feedback coerenti nel tempo, e consente una verifica rapida dell'azione. Fallisce quando fornisce dettagli eccessivi senza gerarchia, utilizza un'estetica non funzionale che aumenta il carico cognitivo o nasconde i limiti del sistema, producendo non solo sfiducia ma una vera e propria rottura epistemica che può innescare il desiderio di intervento manuale ("takeover feeling").

RISCHIO

I rischi principali sono l'action bias, l'under-trust strutturale e una complacency cognitiva forzata. Se percepisce anche una minima deviazione dalla "traiettoria logica" o un aumento del rischio percepito, questo utente può sviluppare un bisogno compulsivo di riprendere il controllo manuale, anche quando il sistema opera in sicurezza, per evitare di rimanere "out-of-the-loop". In termini di calibrazione, questa configurazione è esposta al rischio di undertrust qualora la trasparenza sia percepita come insufficiente o incoerente con la performance osservata. L'obiettivo progettuale diventa quindi l'allineamento tra modello mentale dell'utente e logica operativa del sistema.

4.2.2 UTENTE ANSIOSO - IPERVIGILANTE

Questa configurazione emerge quando la percezione di vulnerabilità domina l'elaborazione cognitiva. La fiducia è mediata principalmente da processi affettivi ("risk-as-feeling") e dalla necessità di sicurezza percepita prima ancora che di comprensione tecnica. La delega decisionale non viene rifiutata per scetticismo razionale, ma per attivazione emotiva. La valutazione del sistema si fonda meno sull'Ability tecnica astratta e più sulla percezione di Benevolence e stabilità situazionale.

MECCANISMI DELLA FIDUCIA

La fiducia è di tipo affettivo e orientata allo scopo (purpose). Prima di capire, deve percepire stabilità. Essa si costruisce sulla fluidità dell'esperienza, sull'assenza di segnali di incertezza e sulla percezione di continuità. La dimensione centrale non è l'abilità tecnica astratta, ma la riduzione dell'ansia. Si fida quando l'interazione è piacevole e quando percepisce l'integrità benevola dei progettisti. Per questo motivo è particolarmente sensibile all'assenza del pilota umano. Ogni micro-interruzione, variazione brusca o cambiamento improvviso attiva un picco di vigilanza. La coerenza del sistema viene valutata in termini fisiologici: "mi sento stabile?".

COSA LO RASSICURA

Un'estetica curata che trasmetta professionalità, un linguaggio gentile e un certo grado di antropomorfismo dell'interfaccia. Una voce empatica o una comunicazione anticipatoria che riduca l'ambiguità contribuiscono a diminuire la percezione di vulnerabilità. È importante percepire che "qualcuno" sta supervisionando la situazione (ridondanza e controllo) [3.3].

COSA LO DESTABILIZZA

Immagini vivide di fallimenti, allarmi urgenti e stridenti, errori minimi ma visivamente evidenti. Il bias di disponibilità può amplificare l'impatto emotivo di un singolo evento negativo. Il senso di isolamento in una cabina chiusa aumenta la vulnerabilità percepita [3.1.1]. Livelli di automazione elevati senza feedback continuo o spiegazioni tecniche complesse possono incrementare l'incertezza. Il rischio maggiore è l'undertrust cronico, che può innescare un circolo vizioso in cui la percezione del rischio aumenta la vigilanza, portando a interpretazioni negative anche di segnali neutri e a un ulteriore calo di fiducia.

RUOLO DELLA UX

La ux deve agire come un narratore e regolatore emotivo, riducendo l'ambiguità senza aumentare la complessità. È efficace quando crea un'atmosfera di sicurezza percepita attraverso segnali sensoriali coerenti, onboarding proattivo e comunicazione anticipatoria dello stato del sistema. Fallisce quando adotta un approccio eccessivamente tecnico e freddo, mostrando errori senza contestualizzazione, alimentando la percezione di affidabilità globale del veicolo.

RISCHIO

È vulnerabile all'undertrust e all'abbandono del servizio. Può partire da una positività bias iniziale (schema di automazione perfetta), ma la fiducia può collassare improvvisamente dopo un singolo errore, anche non critico. Il rischio è percepito in base alla gravità delle conseguenze immaginate. Questa configurazione è particolarmente esposta a fenomeni di undertrust cronico: un singolo evento negativo può produrre un crollo improvviso

della fiducia, indipendentemente dall'affidabilità statistica del sistema.

4.2.3 IL SUPERVISORE IN CERCA DI CONTROLLO

Questa configurazione si attiva in presenza di un elevato bisogno di agency e controllo situazionale. La fiducia non viene negata, ma subordinata alla possibilità di monitoraggio e intervento. Non rifiuta l'automazione in sé, ma la perdita di ruolo simbolico nel processo decisionale. La delega è accettabile solo se accompagnata da trasparenza di processo (Process transparency) e da meccanismi di supervisione percepita.

MECCANISMI DI FIDUCIA

La fiducia si fonda sulla performance osservata e sulla trasparenza del processo. Non è mai concessa a priori, ma viene calibrata progressivamente attraverso la comprensione dei limiti della macchina (understandability). La fiducia non è guidata principalmente dall'ansia, ma dalla necessità che il sistema riconosca il suo ruolo di supervisore. Non desidera necessariamente intervenire, ma vuole sapere di poterlo fare. Ha bisogno di informazioni sullo stato attuale del sistema, sulle sue intenzioni future e sulla possibilità di

interrogare o verificare le decisioni dell'automazione. Il rischio principale è una oscillazione instabile della fiducia: l'utente può alternare momenti di affidamento a momenti di conflitto uomo-macchina qualora l'automazione limiti eccessivamente l'intervento percepito.

COSA LO RASSICURA

Il controllo comportamentale percepito (perceived behavioral control over automation). È rassicurato dalla presenza di meccanismi di intervento diretto e da interfacce di Explainable AI che mostrino i driver decisionali di alto livello, come meteo, batteria o traffico, consentendogli di monitorare la logica dell'AI.

Necessita di un equilibrio tra logica e rassicurazione: la comprensione tecnica deve coesistere con la percezione di stabilità operativa.

COSA LO DESTABILIZZA

Può sperimentare frustrazione e conflitto con il sistema quando l'automazione appare opaca. La mancanza di feedback in tempo reale e l'impossibilità di intervento manuale (lack of directability) costituiscono fattori critici. Anche la pressione temporale e qualsiasi manovra dell'AI non perfettamente allineata alle sue aspettative situazionali possono generare una percezione di perdita di controllo.

RUOLO DELLA UX

La UX deve agire come mediatore di agency. È efficace quando fornisce una forma calibrata di controllo simbolico (illusory control) che soddisfi il bisogno di supervisione senza sovraccaricare l'utente. Deve tradurre l'autonomia tecnica in un equilibrio percepito tra controllo umano e decisione automatica. Fallisce quando offre un controllo eccessivo che espone discrepanze insignificanti tra la logica della macchina e la preferenza umana, oppure quando non chiarisce adeguatamente i limiti di intervento.

RISCHIO

Questo profilo presenta un alto rischio di disuse, ovvero rifiuto del sistema per eccessiva sfiducia, e di sovraccarico cognitivo dovuto alla necessità compulsiva di verificare costantemente ogni parametro, specialmente in situazioni percepite come rischiose. Può intervenire manualmente durante manovre automatizzate sicure ma percepite come divergenti dal proprio stile, generando potenziali instabilità operative o conflitti di autorità uomo-macchina.

4.2.4 UTENTE FIDUCIOSO PASSIVO

Questa configurazione è caratterizzata da un'elevata propensione alla delega e da una rapida attribuzione di competenza al sistema, spesso basata su analogie con tecnologie quotidiane percepite come sicure (ascensore, autopilota automobilistico). Tende a operare attraverso il sistema esperienziale e a basare la propria fiducia sulla comodità e sulla somiglianza con sistemi già noti, come ascensori, aerei di linea o veicoli Tesla. Mostra una predisposizione positiva verso le nuove tecnologie: se l'esperienza "sembra corretta", delega completamente il controllo. L'interazione con l'eVTOL è caratterizzata da accettazione spontanea, in cui l'AI è percepita come facilitatore trasparente e benevolo. La vulnerabilità non è vissuta come minaccia primaria, ma come condizione delegabile.

MECCANISMI DI FIDUCIA

La fiducia si fonda su una fede iniziale e sulla reputazione del

sistema. Spesso utilizza analogie cognitive (“se funziona sugli aerei di linea, funzionerà anche qui”). Manifesta un forte automation bias: tende a preferire i suggerimenti del veicolo e a ridurre il monitoraggio ambientale. La modernità, l'estetica e la reputazione del brand fungono da proxy cognitivi della competenza. La delega è percepita come benefico, e l'automazione decisionale come sollievo cognitivo.

COSA LO RASSICURA

Un'esperienza di volo fluida, priva di interruzioni comunicative. È rassicurato da interfacce minimali che utilizzano persuasione estetica e toni assertivi, confermando il suo modello mentale di infallibilità tecnologica. Per questa tipologia, il comfort emotivo e la familiarità costituiscono indicatori di sicurezza più forti delle statistiche tecniche. Minore frizione equivale a maggiore fiducia.

COSA LO DESTABILIZZA

Raramente viene destabilizzato da errori minori, che tende a ignorare o razionalizzare (bias di conferma). La destabilizzazione avviene in modo improvviso e intenso solo in presenza di un guasto catastrofico o di un comportamento palesemente illogico, per il quale non possiede un modello mentale interpretativo. Può sperimentare frustrazione quando il sistema restituisce responsabilità o comunica segnali di incertezza

che mettono in discussione la sua aspettativa di infallibilità.

RUOLO DELLA UX

La UX deve agire come regolatore della vigilanza. È efficace quando inserisce una trasparenza leggera ma intenzionale, attraverso micro-indicatori di incertezza (uncertainty cues), segnali di confidenza e reminder sottili in situazioni ambigue, mantenendo l'utente cognitivamente nel loop. Fallisce quando l'esperienza è eccessivamente fluida e invisibile, alimentando una fiducia cieca che maschera i limiti dell'Operational Design Domain (ODD).

RISCHIO

È il profilo più vulnerabile alla complacency indotta dall'automazione e all'overtrust. In caso di evento anomalo, l'assenza di un modello mentale solido può generare shock improvviso e perdita drastica di fiducia. La mancanza di monitoraggio attivo può condurre a una ridotta situazione awareness, rendendo lento e inefficace un eventuale recupero manuale in emergenza. Questa configurazione è la più esposta a fenomeni di overtrust e complacency indotta dall'automazione. L'obiettivo progettuale non è aumentare la fiducia, ma mantenerla entro livelli appropriati, preservando la consapevolezza situazionale.

Tabella 9: Sintesi comparativa delle quattro tipologie di utenti nei PAV/ eVTOL, con indicazione della base psicologica della fiducia, dei bisogni prevalenti, delle leve UX prioritarie, dei rischi di disallineamento e degli obiettivi di calibrazione progettuale.

Profilo Utente	Base della fiducia	Tratto Psicologico Prevalente	Soluzioni UX/HMI Suggeste	Rischio Correlato
Razionale - Analitico	Analitica (Performance, Process)	Bisogno di logica e coerenza procedurale.	XAI Dettagliata Visualizzazione dei pesi decisionali (slider), dati meteo in tempo reale, logica delle rotte.	Sfiducia se i dati sono percepiti come insufficienti (Disuse)
Ansioso - Ipervigilante	Affettiva (Risk-as-Feeling)	Sensibilità al rischio e vulnerabilità emotiva.	Comfort Design: Toni caldi, interfacce vocali (Iris), feedback costanti di "stato ok", narrazione dei benefici sociali.	Takeover immotivato: Tentativo di intervento per ansia situazionale.
Profilo Utente Control-Seeking Supervisor	Operative (Agency-based)	Bisogno di controllo, supervisione attiva	Pulsante di intervento chiaro; indicatori di controllo condiviso; modalità "supervised autonomy"; log decisionale consultabile	Conflitto uomo-macchina; Illusory Control; oscillazione fiducia
Fiducioso - Passivo	Analogica e disposizionale	Tendenza alla delega basata su esperienze pregresse.	Engagement Tools: Segnali di attenzione (nudge), checklist di supervisione, interfacce che richiamano il volo di linea.	Misuse/Complacency: Eccessivo affidamento e perdita di consapevolezza situazionale.

4.3 MAPPATURA DEI PATTERN RICORRENTI

Le tematiche comuni che emergono da quelle in alto sono la gestione della Trasparenza (Cosa mostrare e come). Tutti gli utenti reagiscono alla trasparenza, ma con bisogni opposti. Il problema comune è il grado di apertura della “scatola nera” dell’algoritmo. Il conflitto nasce quando se mostri troppo, sovraccarichi l’utente razionale e spaventi l’ansioso. Se mostri troppo poco, perdi il supervisore e induci l’ottimista in una “falsa calma” pericolosa. Una manovra brusca è un errore logico per il razionale, una minaccia vitale per l’ansioso e un brusco risveglio per il passivo.

La seconda dimensione riguarda l’agency e il senso di controllo, sintetizzabile nella domanda implicita: “chi comanda?”. Il fenomeno del takeover feeling accomuna l’utente razionale, l’ansioso e il supervisore, sebbene per motivazioni differenti. Il primo desidera intervenire per validare la coerenza del sistema, il secondo per ridurre l’ansia, il terzo per mantenere il proprio ruolo simbolico di supervisione. L’utente fiducioso, al contrario, tende ad abbandonare l’agency, correndo il rischio di finire fuori dal loop cognitivo. L’agency non rappresenta quindi un bisogno uniforme, ma una

variabile distribuita lungo uno spettro che va dall’eccesso di controllo alla completa delega.

La terza dimensione concerne la prevedibilità e la coerenza tra aspettativa e comportamento del sistema. Ogni profilo entra nel velivolo con un modello mentale preesistente, derivato da esperienze analogiche: l’aereo di linea, l’ascensore, l’automobile tradizionale o un veicolo semi-autonomo. La discrepanza tra ciò che il velivolo fa e ciò che l’utente si aspetta costituisce il principale trigger di sfiducia. Se il comportamento fisico del mezzo non è sincronizzato con il comportamento informativo dell’interfaccia, si genera un disallineamento cognitivo. Per il razionale si tratta di un errore logico; per l’ansioso di un segnale di pericolo; per il fiducioso di una rottura improvvisa della propria fiducia analogica.

La quarta dimensione riguarda la calibrazione della fiducia. Nessuna tipologia possiede una fiducia perfettamente proporzionata alle capacità del sistema. L’utente razionale e quello ansioso tendono verso forme di undertrust, seppur per motivazioni differenti; l’utente fiducioso manifesta invece un rischio di overtrust e complacency; il supervisore oscilla tra controllo e delega. Il design non deve quindi aumentare indiscriminatamente la

DALL’ANALISI DELLE QUATTRO TIPOLOGIE EMERGONO ALCUNE TEMATICHE TRASVERSALI CHE COSTITUISCONO I NODI STRUTTURALI DEL RAPPORTO TRA UTENTE E SISTEMA AUTONOMO.

fiducia, ma calibrare, spingendo chi diffida eccessivamente verso una maggiore serenità e richiamando chi delega troppo a una vigilanza attiva.

Oltre a queste quattro macro-dimensioni. Trasparenza, Agency, Prevedibilità e Calibrazione emergono ulteriori tematiche trasversali che incidono sulla stabilità del sistema uomo-macchina. E’ necessario tenere in considerazione il modello mentale (l’ancoraggio al passato). Tutti e quattro gli utenti entrano nel velivolo cercando di confrontarlo con qualcosa che già conoscono.

[Il Razionale, cerca la logica di un software professionale.](#)
[L’Ansioso, cerca la protezione di un pilota umano \(che qui manca\).](#)
[Il Supervisore cerca la dinamica di guida di un’auto o un aereo manuale.](#)
[Il Fiducioso, lo paragona a un ascensore o a una modello di macchina con guida automatica integrata.](#)

In questi casi, la UX fallisce per tutti se non aiuta l’utente a costruire un nuovo modello mentale specifico per l’eVTOL. Senza un onboarding che spieghi le nuove regole del funzionamento, ciascun utente giudicherà il volo sulla base di parametri inadeguati (ad esempio, l’Ansioso potrebbe interpretare una correzione del vento come un guasto). Un’ulteriore dimensione trasversale riguarda la gestione della vulnerabilità. Anche se le tipologie distinguono tra chi tende ad accettarla (Fiducioso) e chi la teme (Ansioso), la vulnerabilità rappresenta un elemento strutturale per tutti.

[Il Razionale e il Supervisore cercano di “risolvere” la vulnerabilità con i dati e il controllo.](#)
[L’Ansioso e il Fiducioso la gestiscono a livello emotivo \(uno scappando, l’altro ignorandola\).](#)

Tutti, tuttavia, necessitano di una rete di sicurezza percepita. Che si tratti di un grafico dei pesi decisionali o di una voce calma che comunica “Stato OK”, la UX deve costantemente rispondere alla domanda implicita: “Cosa succede se qualcosa va storto?”.

Un altro elemento critico è la sensibilità alla discrepanza, ovvero il trigger della sfiducia. Per tutti gli utenti, la fiducia non è statica, ma estremamente fragile di fronte all'incoerenza. Se il velivolo vira bruscamente (comportamento fisico) e l'interfaccia non lo segnala o non lo giustifica (comportamento informativo), si genera un disallineamento cognitivo.

La sincronia diventa quindi un requisito fondamentale. Per il Razionale, la mancata coerenza rappresenta un errore logico; per l'Ansioso, un segnale di pericolo; per il Fiducioso, una rottura improvvisa della propria fiducia nel sistema. La UX deve pertanto essere perfettamente sincronizzata con il movimento fisico del mezzo.

Ciò riporta l'attenzione a tre esigenze progettuali fondamentali. La necessità di un onboarding cognitivo: tutti devono apprendere “come pensare” questo nuovo mezzo. La mitigazione del senso di isolamento: tutti, sebbene per ragioni differenti (ansia o bisogno di controllo), temono di essere fuori dal loop. Infine, un linguaggio della conferma: tutti necessitano di un feedback che validi

la coerenza tra la realtà fisica del volo e la rappresentazione digitale dell'interfaccia.

Le differenze tra le configurazioni non indicano categorie separate di utenti, ma polarità lungo uno spettro continuo di calibrazione della fiducia. Tutte condividono la medesima condizione strutturale: l'accettazione di vulnerabilità in assenza di controllo diretto. Ciò che varia è la strategia cognitiva ed emotiva attraverso cui tale vulnerabilità viene gestita.

4.4 IL SAFETY-TRANSPARENCY TRUST FRAMEWORK COME STRUMENTO PROGETTUALE

L'obiettivo del Safety-Transparency Trust Framework non è massimizzare la fiducia dell'utente, bensì di calibrarla dinamicamente rispetto alle reali capacità operative del sistema e ai modelli mentali dell'utilizzatore. In coerenza con la definizione di fiducia proposta da Mayer, Davis e Schoorman (1995), La fiducia, è intesa come disponibilità ad accettare una condizione di vulnerabilità in un contesto di incertezza. Nel caso dei PAV/eVTOL, tale vulnerabilità è strutturale: l'utente delega la propria sicurezza a un sistema autonomo operante in uno spazio tridimensionale, con limitate possibilità di intervento diretto. Il framework assume quindi che l'obiettivo progettuale non sia aumentare la disponibilità alla

vulnerabilità, ma allinearla alla reale affidabilità del sistema. Questo implica non portare l'utente a fidarsi sempre e comunque, bensì portare la fiducia ad allinearsi alle effettive capacità operative del velivolo, evitando sia la diffidenza ingiustificata (undertrust), sia l'affidamento acritico (overtrust), entrambi descritti nella letteratura sulla Trust in Automation. Qui il ruolo del design, non è persuasivo né meramente rassicurante.

Il framework risponde al quesito metodologico centrale della ricerca: “Quali sono le conoscenze necessarie al designer per ottimizzare l'esperienza utente all'interno dei PAV?”. Il modello non si configura come soluzione normativa prescrittiva, ma come strumento integrativo capace di connettere dimensioni tecnica, trasformando l'interazione uomo-macchina in una cooperazione strutturata, il framework allinea tre pilastri fondamentali:

[Capacità reale del sistema \(prestazioni oggettive\);](#)
[Comprensione dell'utente \(modello mentale\);](#)
[Comportamento di reliance effettivo \(azione dell'utente\).](#)

Attraverso tale allineamento, il modello mira a mitigare fenomeni disfunzionali quali l'automation bias, la riduzione del monitoraggio, il decremento della vigilanza umana e l'erosione delle competenze manuali nel lungo periodo. Il principio cardine della Safety-Transparency che funge qui da mediatore è la trasparenza delle interfacce agisce come un “traduttore” capace di rendere leggibile l'autonomia, trasformando la performance tecnica in un'esperienza di sicurezza percepita. Attraverso la modulazione adattativa dell'agency e la prevenzione dei bias cognitivi, il framework permette all'utente non di affidarsi ciecamente alla tecnologia, ma di calibrare il proprio affidamento sulla base dei reali limiti operativi del sistema.

Tramite il Questo framework va a trattare un obiettivo non ancora trattato. Nei diversi framework visitati, che vanno a toccare tematiche calde della UAM, esiste una lacuna logica, un obiettivo non ancora esplorato nelle fonti è la creazione di un framework unificato che colleghi la maturità analitica (TCMM) alle specifiche di implementazioni della trasparenza (SAT) in tempo reale. Inoltre, si nota che la maggior parte degli obiettivi degli

esperimenti analizzati è tarata su scenari di breve durata (simulazioni). Un altro obiettivo è quello di dovrebbe estendersi alla gestione della fiducia a lungo termine per evitare l'erosione delle abilità manuali identificata come rischio [Capitolo 3]. La UX deve agire come il mediatore supremo che assorbe la computazione algoritmica e la restituisce all'utente sotto forma di comprensione, controllo e valore sociale.

Tabella 10: Dalla teoria alla decisione progettuale. Questa mappa dimostra che il framework non è arbitrario, ma deriva direttamente dalla sintesi critica della letteratura.

Fondamento teorico (Cap. 2–3)	Rischio individuato	Implicazione progettuale	Asse del Framework
Trust in Automation (Lee & See)	<u>Overtrust / Undertrust</u>	Calibrazione proporzionata alle capacità reali	Adaptive Trust Calibration
Risk-as-Feeling (Slovic)	Amplificazione emotiva del rischio	Mediazione narrativa e spiegabilità contestuale	Emotional Risk Mediation
Explainable AI (XAI)	Black-box effect	Esplicitazione del "perché" decisionale	Active Transparency
Human–Autonomy Teaming (HAT)	Out-of-the-loop	Supervisione attiva e micro-interazioni	Active Supervision
Irony of Automation	Skill degradation	Automazione adattiva e uncertainty cues	Anti-Automation Bias
Perceived Control (Walker)	Vulnerabilità percepita	<u>ricostruzione</u> dell'agency	Agency Reconstruction
Learned Trust	Fiducia inerziale	Feedback storico e specificità temporale	Longitudinal Trust

Caratteristiche	Massimizzazione del Trust	Calibrazione del Trust (Framework)
Obiettivo Finale	Indurre l'utente a fidarsi sempre e comunque	Allineare la fiducia alle reali capacità del sistema
Rischio Principale	Eccesso di fiducia: affidamento cieco e automation bias	Nessuno (mira all'equilibrio ottimale)
Ruolo del Design	Riduzione della vigilanza umana	Trasparente e informativo
Risultato Operativo	Rischio di sottovalutazione dei limiti tecnici	Appropriate reliance e monitoraggio attivo

4.4.2 LINEE GUIDA STRATEGICHE PER IL DESIGN UAM

Tabella 11: Rappresentazione delle differenze tra massimizzazione del trust e Calibrazione del Trust.

La UX come traduttore del pilota inesistente. Non avendo più i piloti presenti a bordo, la UX deve "umanizzare" l'autonomia per colmare questo vuoto.

Integrazione suggerita: si propone che tale ruolo possa essere traslato sulle surface features dell'interfaccia. Se l'utente ricerca la protezione umana, il design deve rispondere attraverso una narrazione visiva che emuli la comunicazione del pilota (ad esempio toni di voce calmi, feedback visivi fluidi e colori freddi come il blu e il bianco), sfruttando l'Effetto Veridicità (Truthiness) per indurre una percezione di competenza anche in assenza di un operatore in carne e ossa.

Il collegamento tra il rischio come sentimento e la calibrazione della fiducia

Integrazione suggerita: se l'utente valuta il rischio su basi emotive, la UX deve agire come un "mediatore supremo" che trasforma il calcolo probabilistico dell'IA in una spiegazione logica e comprensibile. Ad esempio, per bilanciare l'approccio di "supervised autonomy" di Wisk, il sistema non deve limitarsi a operare silenziosamente, ma deve esplicitare il proprio "perché" (es. "Virata in corso per evitare corrente ascensionale"),

trasformando l'utente da passeggero passivo e ansioso a supervisore informato.

Per superare l'atteggiamento di attesa prudente ("wait-and-see")

Integrazione suggerita: per superare l'atteggiamento di attesa prudente, la tesi propone di differenziare la comunicazione in base ai modelli mentali degli utenti. Si identificano diversi pattern: il profilo "Ansioso", che richiede una maggiore simulazione di presenza umana, e il profilo "Razionale", per il quale la fiducia si costruisce attraverso la trasparenza dei dati tecnici (TCMM).

Per non portare l'utente a un'erosione delle abilità

Integrazione suggerita: il rischio di "erosione delle abilità manuali", citato nel Capitolo 3 della tesi, diventa qui fondamentale. Il framework deve garantire che la "supervised autonomy" non diventi compiacenza (automation bias), mantenendo l'utente in un loop di monitoraggio attivo attraverso micro-interazioni di controllo.

Sfruttare l'effetto Veridicità

Integrazione suggerita: gli utenti tendono a fidarsi di ciò che appare "esteticamente vero" e professionale a livello intuitivo. L'uso di immagini ad

alta risoluzione o di display dettagliati aumenta la credibilità percepita del sistema. È fondamentale fornire feedback costanti di "stato OK" attraverso toni calmi e rassicuranti, al fine di mantenere la percezione di sicurezza.

Integrazione suggerita: gli utenti tendono a fidarsi di ciò che appare "esteticamente vero" e professionale a livello intuitivo. L'uso di immagini ad alta risoluzione o di display dettagliati aumenta la credibilità percepita del sistema. È fondamentale fornire feedback costanti di "stato OK" attraverso toni calmi e rassicuranti, al fine di mantenere la percezione di sicurezza. In sintesi, integrare la "supervisione attiva e strategica" nel framework significa progettare una UX che non si limita a rassicurare, ma che stimola l'utente a rimanere un partner vigile, prevenendo la degradazione delle sue capacità critiche e manuali.

4.4.3 ASSI PROGETTUALI

Il Safety-Transparency Framework si organizza in cinque assi progettuali macro-direttrici, che non vanno a rassicurare passivamente l'utente, ma lo mantiene partner vigile del sistema, prevenendo la degradazione delle capacità critiche e manuali e trasformando l'autonomia in una forma di cooperazione strutturata tra uomo e macchina.

Il primo asse, [Emotional Risk Mediation & Explainability](#), affronta la distanza tra rischio affettivo e comprensione situazionale. Poiché l'utente tende a valutare il rischio su base emotiva piuttosto che probabilistica, un'autonomia opaca può amplificare la vulnerabilità percepita. L'obiettivo sistemico consiste quindi nel trasformare il calcolo tecnico dell'intelligenza artificiale in una comprensione situazionale assimilabile. Ciò avviene attraverso l'applicazione del modello SAT (Situation Awareness-based Agent Transparency), l'esplicitazione del "perché" operativo dell'IA, l'introduzione di micro-narrazioni contestuali, come l'indicazione di una virata per evitare una corrente ascensionale e la traduzione degli high-level drivers in una narrativa situazionale accessibile. In questo modo, il rischio percepito viene progressivamente convertito in comprensione situazionale, consentendo all'utente di passare da passeggero passivo a supervisore informato.

Il secondo asse, [Adaptive Trust Calibration](#), che si muove tra la fiducia modulare e differenziata, si fonda sul superamento del paradigma "wait-and-see" e dell'idea di una fiducia uniforme e indifferenziata. La fiducia non è una variabile statica, ma deve essere calibrata in funzione dei modelli mentali dell'utente. L'obiettivo è costruire una fiducia dinamica e proporzionata, attraverso una segmentazione per profilo cognitivo ed emotivo. Il profilo ansioso richiede una maggiore simulazione di presenza umana e continuità comunicativa; il profilo razionale, invece, costruisce la propria fiducia sulla trasparenza tecnica e sull'accesso ai dati (come nel caso del TCMM). A ciò si affianca una modularità comunicativa capace di garantire l'allineamento tra capacità reale del sistema, modello mentale dell'utente e

comportamento di reliance effettivo. In questa prospettiva, la fiducia non viene imposta, ma calibrata progressivamente, determinando una transizione da comunicazione massiva a comunicazione modulare e adattiva.

Il terzo asse, [Active Transparency & Cognitive Visibility](#), (Autonomia come processo leggibile), affronta il rischio di una “supervised autonomy” opaca e del cosiddetto black-box effect. Quando il processo decisionale dell’automazione rimane invisibile, si genera disallineamento cognitivo e possibile misuse o disuse. L’obiettivo sistemico è dunque evitare l’invisibilità decisionale, rendendo l’autonomia leggibile. Ciò implica l’esplicitazione delle decisioni critiche, la fornitura di feedback in tempo reale, la visualizzazione dei fattori che guidano le scelte operative e la garanzia di continuità informativa lungo l’intero processo. L’autonomia, per essere legittima, deve essere visibile; solo in questo modo può essere percepita come partner cognitivo e non come entità distante e impenetrabile.

Il quarto asse, [Active Supervision & Anti-Automation Bias](#), (Preservazione delle competenze e prevenzione della compiacenza) risponde al problema dell’erosione delle abilità manuali e all’ironia dell’automazione. L’eccessiva delega può generare compiacenza e automation bias, riducendo la capacità dell’utente di intervenire

in caso di necessità. L’obiettivo è mantenere l’utente cognitivamente “in-the-loop”, preservandone il ruolo attivo. A tal fine, il framework integra micro-interazioni di controllo, loop di supervisione attiva, richieste di conferma situazionale e forme di automazione adattiva basate sul mental workload. L’inserimento di uncertainty cues e la trasparenza sulla salute del sistema contribuiscono a ricordare costantemente che l’automazione opera entro limiti definiti. La supervisione sostituisce così la delega totale, preservando le capacità cognitive e la responsabilità operativa.

Il quinto asse, [Agency Reconstruction & Longitudinal Trust](#), (Controllo percepito e gestione temporale della fiducia) si concentra sulla dimensione temporale della fiducia. In ambienti altamente automatizzati, l’utente può sperimentare una percezione di vulnerabilità e perdita di controllo; inoltre, la fiducia tende a consolidarsi nel tempo come fiducia appresa. L’obiettivo sistemico è restituire senso di agency e garantire un aggiornamento dinamico della fiducia in base alle prestazioni reali del sistema. Questo avviene attraverso feedback bidirezionali, indicatori di stato comprensibili, continuità informativa, onboarding cognitivo evolutivo e monitoraggio delle performance storiche. La visualizzazione del purpose e dei

benefici sociali rafforza ulteriormente la stabilità relazionale tra utente e sistema. In questa prospettiva, la sicurezza percepita nasce dalla percezione di controllo, e la fiducia sostenibile si costruisce nel lungo periodo attraverso trasparenza, coerenza e aggiornamento continuo.

4.4.5 ARCHITETTURA LOGICA DEL FRAMEWORK INTEGRATO

I cinque macro-assi del Safety–Transparency Framework operano su tre dimensioni strutturali interconnesse:

La [dimensione cognitiva](#) comprende gli assi di Emotional Risk Mediation e Active Transparency e riguarda la costruzione della comprensione situazionale. Essa interviene sul modo in cui l’utente interpreta il rischio, elabora le informazioni e costruisce il proprio modello mentale rispetto al funzionamento dell’autonomia.

La [dimensione regolativa](#) integra gli assi di Adaptive Trust Calibration e Active Supervision e concerne la modulazione del comportamento di reliance. In questa dimensione, la fiducia non viene considerata come stato emotivo generico, ma come variabile da calibrare in funzione delle capacità reali del sistema e delle condizioni operative.

La [dimensione evolutiva](#), infine, include l’asse di Agency Reconstruction & Longitudinal Trust e introduce la prospettiva temporale. Essa si concentra sulla gestione della fiducia nel lungo periodo, sulla prevenzione dell’erosione delle competenze e sulla ricostruzione del senso di agency dell’utente all’interno di ambienti altamente automatizzati.

Nel loro insieme, queste tre dimensioni producono un framework integrato in cui l’obiettivo non è semplicemente “rendere l’IA rassicurante”, bensì costruire un ecosistema di cooperazione strutturata tra utente e autonomia. In tale configurazione, la fiducia non è statica, ma dinamica e calibrata; l’autonomia non è opaca, ma leggibile e spiegabile; l’utente non è passivo, ma rimane cognitivamente attivo; la sicurezza percepita è ancorata alla sicurezza reale; e la relazione uomo–macchina si configura come processo evolutivo che si sviluppa nel tempo.

Fase Temporale	Stato psicologico prevalente	Rischio sistemico	Intervento UX
Primo Utilizzo	Incertezza, vulnerabilità	Sfiducia e ansia	Onboarding cognitivo
Uso ripetuto	Familiarizzazione	Eccesso di fiducia latente	Micro-indicatori di incertezza
Uso Prolungato	Automazione cognitiva	Skill loss	Supervisione attiva
Evento critico	Discrepanza	Rottura della fiducia	Spiegazione immediata + trasparenza situazionale
Post-evento	Ricalibrazione	Fiducia distorta	Feedback storico e performance review

Tabella 12: Mappa longitudinale della Fiducia sulla gestione del tempo.

4.5 APPLICAZIONE DEL FRAMEWORK NELLE CINQUE

(User-Adaptive Flight Model)
 In base ai concetti progettuali definiti, il modello non prevede la creazione di quattro interfacce distinte — soluzione costosa e complessa da gestire — bensì la progettazione di un'interfaccia “elastica”, capace di adattarsi dinamicamente al comportamento e al profilo dell'utente. Se l'utente non interagisce attivamente con il sistema, riceve un livello base di rassicurazione, adeguato ai profili Ansioso e Fiducioso. Qualora, invece, l'utente desidera approfondire, l'interfaccia si espande progressivamente, offrendo accesso ai dati tecnici necessari ai profili Razionale e Supervisore.

Preflight & Onboarding:

Costruzione del Modello Mentale

L'obiettivo di questa fase è allineare le aspettative dell'utente con la realtà tecnologica del velivolo, favorendo la costruzione di un modello mentale coerente.

Soluzione UX: introduzione di un “check-in digitale” che non sia esclusivamente procedurale, ma informativo ed educativo.

Per il profilo Razionale/Supervisore, è prevista la visualizzazione dello “stato di salute” del sistema (livello della batteria, condizioni meteo, rotta ottimizzata). Per il profilo Ansioso, può essere proposto un breve video con audio spaziale e tono calmo, che simuli suoni e movimenti del decollo

al fine di favorire una graduale desensibilizzazione. Per il profilo Fiducioso, è sufficiente una notifica sintetica di “pronti al via”, caratterizzata da un'estetica familiare, simile alle applicazioni di ride-sharing.

Elemento comune a tutti i profili è l'anteprima della traiettoria: mostrare dove e come si volerà riduce l'ambiguità e rafforza la prevedibilità del sistema.

Take-off & Climb: Gestione della Vulnerabilità e Agency

La fase di decollo e salita rappresenta il momento a più alto carico emotivo e cognitivo, in quanto coincide con il distacco dal suolo.

Soluzione UX: implementazione di un “countdown informativo”.

Per il profilo Razionale/Supervisore, è prevista la visualizzazione dei vettori di spinta o di una telemetria semplificata.

Per il profilo Ansioso, può essere introdotto un messaggio di conferma della ridondanza (es. “Tutti i sei motori operativi al 100%”). Per il profilo Fiducioso, un'animazione fluida può valorizzare il decollo come momento di comfort e continuità.

Elemento comune è la sincronia sensoriale: l'interfaccia deve rappresentare graficamente la salita nel momento in cui il corpo la percepisce, riducendo il rischio di disorientamento, nausea o panico da perdita di controllo.

Cruise: Calibrazione della Fiducia e Vigilanza

La fase di crociera è la più lunga e può generare, a seconda del profilo, noia (Fiducioso) o iper-controllo (Razionale).

Soluzione UX: implementazione di “layer di trasparenza a richiesta”.

Per il profilo Razionale/Supervisore, è previsto l'accesso a un menu avanzato contenente dati sul traffico aereo circostante e informazioni spiegabili secondo logiche XAI. Per il profilo Ansioso, può essere introdotta una linea temporale che mostri il

tempo residuo al termine del volo e la localizzazione dell'area di atterraggio alternativa più vicina, presentata come “porto sicuro” anziché come “emergenza”. Per il profilo Fiducioso, si possono inserire nudge di attenzione periodici, come brevi suggerimenti contestuali (“Stiamo sorvolando...”), al fine di mantenerlo nel loop cognitivo senza generare ansia.

Elemento comune è il feed-forward: ogni variazione di quota o direzione dovrebbe essere notificata alcuni secondi prima che avvenga, migliorando la prevedibilità.

Landing & Transition: Prevedibilità e Chiusura del Ciclo

La fase di atterraggio coincide con il ritorno alla stabilità fisica e con la conferma finale dell'affidabilità del sistema.

Soluzione UX: integrazione di elementi di realtà aumentata per l'atterraggio.

Per il profilo Razionale/Supervisore, è prevista la visualizzazione del punto esatto di atterraggio (vertiporto) con indicazione della tolleranza d'errore.

Per il profilo Ansioso, possono essere forniti messaggi testuali calmi e anticipatori (“Atterraggio morbido previsto tra 30 secondi”). Per il profilo Fiducioso, la transizione può essere orientata verso i servizi

successivi (es. indicazioni per il trasporto a terra).

Elemento comune è il feedback di successo: una sintesi finale del volo (es. “Volo completato con parametri di sicurezza rispettati”) rafforza il modello mentale in vista di esperienze future.

Evento inatteso (Turbolenza / deviazione rotta)

L'obiettivo in caso di evento inatteso è prevenire la rottura della fiducia, mantenendo l'utente nel loop informativo senza sovraccaricarlo cognitivamente.

Il profilo Razionale può notare la discrepanza tra la rotta prevista e quella effettiva e, in assenza di spiegazioni, ipotizzare un malfunzionamento software.

Il profilo Ansioso può interpretare la turbolenza come segnale di imminente cedimento strutturale, lasciando prevalere la logica del risk-as-feeling.

Il profilo Supervisore può percepire una perdita di agency e cercare attivamente strumenti per intervenire.

Il profilo Fiducioso può sperimentare uno shock cognitivo, dovuto alla rottura dell'aspettativa di perfezione del sistema.

Elemento comune è il feedback di successo: Per rispondere a tali dinamiche, il modello non prevede interfacce separate, ma un'interfaccia

elastica e adattiva. Se l'utente non interagisce, riceve un messaggio rassicurante di base; se richiede maggiori informazioni, l'interfaccia si espande offrendo dettagli tecnici. Portando come messaggio universale “Correzione di rotta per vento forte. Il sistema sta stabilizzando il volo in sicurezza”.

Questa formulazione consente di mantenere equilibrio tra rassicurazione e trasparenza, preservando la fiducia senza generare sovraccarico informativo.

CONCLUSIONI

Il Safety–Transparency Trust Framework nasce dalla constatazione che, nel contesto dei Personal Aerial Vehicles autonomi, la fiducia non rappresenta una variabile accessoria dell'esperienza utente, bensì il nodo strutturale attorno al quale si organizza l'interazione uomo–macchina. L'analisi delle configurazioni psicologiche di calibrazione della fiducia ha evidenziato come la distanza tra safety tecnica e safety percepita costituisca il vero terreno progettuale della mobilità aerea urbana. In assenza del pilota umano e in presenza di un'autonomia algoritmica opaca, la vulnerabilità dell'utente non è contingente ma strutturale. In tale scenario, il framework proposto non persegue la massimizzazione indiscriminata della fiducia, ma la sua calibrazione dinamica, intesa come allineamento tra capacità reale del sistema, modello mentale dell'utente e comportamento di reliance effettivo. Questa impostazione consente di superare sia il rischio di undertrust, che può tradursi in rifiuto del sistema o takeover immotivati, sia quello di overtrust e complacency, fenomeni altrettanto critici in ambienti altamente automatizzati.

L'articolazione in cinque assi progettuali, permette di integrare dimensione cognitiva, regolativa ed evolutiva in un'unica architettura coerente. La trasparenza non è concepita come mera esposizione di dati, ma come mediazione situazionale capace di trasformare il calcolo algoritmico in comprensione assimilabile; l'agency non è restituita come controllo totale, ma come supervisione significativa; la fiducia non è trattata come stato emotivo stabile, ma come processo adattivo che evolve nel tempo. Uno dei contributi centrali del framework risiede

proprio nell'introduzione esplicita della dimensione temporale. La maggior parte delle sperimentazioni analizzate in letteratura, e anche molte delle simulazioni discusse nei capitoli precedenti, sono state condotte in contesti di breve durata, spesso limitati a sessioni singole o scenari simulati. Tali studi risultano fondamentali per comprendere le dinamiche iniziali di accettazione, ma non sono sufficienti per intercettare fenomeni progressivi quali l'erosione delle abilità manuali, la normalizzazione del rischio, l'assuefazione alla trasparenza o la formazione di una fiducia inerziale (learned trust).

La fiducia nei sistemi autonomi, infatti, non si stabilizza nel momento del primo utilizzo: essa si costruisce, si consolida, può deteriorarsi e talvolta rompersi lungo traiettorie temporali complesse. Per questo motivo, il framework propone di estendere l'analisi oltre la fase di onboarding o di volo singolo, introducendo una prospettiva longitudinale che osservi l'interazione uomo-autonomia nel medio e lungo periodo. Solo attraverso studi empirici prolungati nel tempo sarà possibile verificare se le strategie di trasparenza adattiva, supervisione attiva e ricostruzione dell'agency riescano effettivamente a prevenire la degradazione delle competenze e a mantenere una vigilanza appropriata.

In conclusione, la sfida della UAM non consiste esclusivamente nel perfezionamento tecnico dei velivoli autonomi, ma nella progettazione di un ecosistema relazionale capace di sostenere nel tempo una fiducia appropriata. Se la sicurezza reale costituisce il fondamento ingegneristico del sistema, la trasparenza calibrata ne rappresenta la traduzione esperienziale. Solo integrando queste due dimensioni lungo un orizzonte temporale esteso sarà possibile evitare che l'autonomia produca, paradossalmente, nuove forme di vulnerabilità cognitiva.

La fiducia, in questo quadro, non è un punto di arrivo statico, ma un equilibrio dinamico che deve essere costantemente monitorato, regolato e ricostruito. Ed è proprio in questa gestione evolutiva della relazione uomo-macchina che il framework trova la propria rilevanza teorica e applicativa.

Figura 4.1 Mappa concettuale del Safety-transparency Framework.

EMOTIONAL RISK MEDIATION & EXPLAINABILITY

Axis 1

Transform "fear of the unknown" into "understanding the maneuver."

ADAPTIVE TRUST CALIBRATION

Axis 2

Create an "elastic" interface that adapts based on who is on board.

ACTIVE TRANSPARENCY & COGNITIVE VISIBILITY

Axis 3

Eliminate the "Black Box" effect. AI must be a readable partner.

ACTIVE SUPERVISION & ANTI-AUTOMATION BIAS

Axis 4

Prevent the user from falling asleep or becoming overly trusting (complacency).

AGENCY RECONSTRUCTION & LONGITUDINAL TRUST

Axis 5

Make the user feel in control and build trust over time.

TAKE-OFF & CLIMB

Gestione della Vulnerabilità e Agency

The take-off and climb phase represents the most emotionally and cognitively charged moment, as it coincides with liftoff. Profile

Rational	Supervisor	Anxious	Confident
----------	------------	---------	-----------

Sensory synchrony The interface must graphically represent the climb as the body perceives it, reducing the risk of disorientation, nausea, or panic due to loss of control.

Design Strategy: Transform technical AI calculations into assimilable situational understanding.

Implementation: Apply the SAT model. Explicitly state the "why" behind AI operations through contextual micro-narratives. Translate high-level drivers into an accessible situational narrative.

UX solution: Implementation of an "information countdown."

Design: Textual and graphic micro-narratives.

LANDING & TRANSITION

Predictability and Closing the Loop

The landing phase coincides with the return to physical stability and the final confirmation of system reliability. Profile

Rational	Supervisor	Anxious	Confident
----------	------------	---------	-----------

feedback A final summary of the flight, reinforces the mental model for future experiences.

Design Strategy: Restore the user's sense of agency and provide dynamic trust updates based on real performance.

Implementation: Utilize bidirectional feedback, comprehensible status indicators, evolutionary cognitive onboarding, and historical performance monitoring.

UX solution: Integration of augmented reality elements for landing.

Design: Historical feedback and "Agency" logic. The user must feel that their monitoring has value.

PREFLIGHT & ONBOARDING

Mental Model Construction

The goal of this phase is to align the user's expectations with the technological reality of the aircraft, promoting the construction of a coherent mental model. Profile

Rational	Supervisor	Anxious	Confident
----------	------------	---------	-----------

Trajectory preview Showing where and how the flight will take place reduces ambiguity and strengthens the system's predictability.

Design Strategy: Build proportionate trust through segmentation by cognitive and emotional profile.

Implementation: Requires increased simulation of human presence and continuous communication.

UX solution: Introduction of a "digital check-in" that is not exclusively procedural, but informative and educational.

Design: Modular levels of detail. The interface must have two main "states" based on the user profile.



UNEXPECTED EVENT

Turbulence, Route Deviation

The goal of an unexpected event is to prevent a breach of trust by keeping the user in the information loop without cognitively overloading them. Profile

Rational	Supervisor	Anxious	Confident
----------	------------	---------	-----------

Success feedback To respond to these dynamics, the model does not provide separate interfaces, but rather a flexible and adaptive interface.

UX solution: Implementation of an "information countdown."

Design: Elastic and adaptive interface.

CRUISE

Calibrating Confidence and Vigilance

The cruise phase is the longest and can generate, depending on the profile, boredom or hyper-control. Profile

Rational	Supervisor	Anxious	Confident
----------	------------	---------	-----------

Feed-forward Any change in altitude or direction should be notified a few seconds before it occurs, improving predictability.

Design Strategy: Keep the user cognitively "in-the-loop" by preserving their active role.

Implementation: Integrate active supervision loops, situational confirmation requests, and adaptive automation based on mental workload. Insert uncertainty cues and "system health" transparency to remind the user of the system's operational limits.

UX solution: Implementation of "on-demand transparency layers."

Design: Active interaction loops. Insert micro-tasks that require user confirmation, especially at critical stages.

ARCHITECTURE DIMENSIONS: ■ Cognitive ■ Regulatory Evolutionary

5 AXIES

SAFETY-TRANSPARENCY TRUST FRAMEWORK

Il Safety-Transparency Framework si organizza in cinque assi progettuali macro-direttrici, che non vanno a rassicurare passivamente l'utente, ma lo mantiene partner vigile del sistema,

prevenendo la degradazione delle capacità critiche e manuali e trasformando l'autonomia in una forma di cooperazione strutturata tra uomo e macchina

BIBLIOGRAFIA

- Abdulla, A., et al. (2025). eVTOLs as Aerial Taxis in Urban Environments: Human-Centred Design and Operational Considerations for Next-Generation Air Transport. *OxJournal*.
- Allianz. (2025). Urban Air Mobility: Public perception and the importance of takeover availability.
- Alonso, V., & de la Puente, P. (2018). System Transparency in Shared Autonomy: A Mini Review. *Frontiers in Neurorobotics*, 12, 83. <https://doi.org/10.3389/fnbot.2018.00083>
- Arafat, M. Y., Alam, M. M., & Moh, S. (2023). Vision-based navigation techniques for unmanned aerial vehicles: Review and challenges. *Drones*, 7(2), 89. <https://doi.org/10.3390/drones7020089>.
- Battiste, V., Lachter, J., Brandt, S., Alvarez, A., Strybel, T. Z., & Vu, K.-P. L. (2018). Human-Automation Teaming: Lessons Learned and Future Directions. NASA Ames Research Center; San Jose State University Foundation; California State University Long Beach.
- Benda, N. C., Novak, L. L., Reale, C., & Ancker, J. S. (2020). Trust in AI: Why we should be designing for appropriate reliance. *Journal of the American Medical Informatics Association*, 27(5), 709–716. <https://doi.org/10.1093/jamia/ocaa038>
- Benson, B. (2016, 1 settembre). Cognitive bias cheat sheet. Medium. <https://buster.medium.com/cognitive-bias-cheat-sheet-55a472476b18>
- Booz Allen Hamilton. (2018, October 5). Urban air mobility (UAM) market study. <https://ntrs.nasa.gov/citations/20190001472>
- Chen, J. Y. C., Lakhmani, S. G., Stowers, K., Selkowitz, A. R., Wright, J. L., & Barnes, M. (2018). Situation awareness-based agent transparency and human–autonomy teaming effectiveness. *Theoretical Issues in Ergonomics Science*, 19(3), 259–282. https://www.researchgate.net/publication/323371993_Situation_awareness-based_agent_transparency_and_human-autonomy_teaming_effectiveness/citations
- Committee on Science, Space, and Technology. (2018, July 24). Urban Air Mobility—Are Flying Cars Preparing for Takeoff? (Hearing before the Committee on Science, Space, and Technology, House of Representatives, 115th Congress). U.S. Government Publishing Office.
- Coppola, P., Silvestri, F., & De Fabiis, F. (2024). Heterogeneity in users' intention-to-use Urban Air Mobility services. *Transportation Research Procedia*, 78, 460–466. <https://doi.org/10.1016/j.trpro.2024.02.058>
- Dassault Systèmes. (2025). 5 elementi essenziali per volare alto nel mercato della mobilità aerea avanzata [E-book].
- De Backer, G. (2024). Cognitive biases: Complete list of 151 biases. Gust de Backer. <https://gustdebacker.com/cognitive-biases/>
- De Winter, J. C. F., Happee, R., Martens, M. H., and Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence. *Transp. Res. Part F: Traff. Psychol. Behav.* 27(PB), 196–217. doi: 10.1016/j.trf.2014.06.016
- Dietrich, A. M. (2018). Summary statement: Prepared for the House Committee on Science, Space, and Technology. Terrafugia Inc.
- EASA (European Union Aviation Safety Agency). (2021). Study on societal acceptance of Urban Air Mobility in Europe. Full Report. <https://www.easa.europa.eu/sites/default/files/dfu/uam-full-report.pdf>
- EHang. (2020, 15 gennaio). The future of transportation: White paper on urban air mobility systems.
- Embention. (2025). Embention becomes the first POA-Certified manufacturer specializing in autopilots for drones and eVTOLs, approved by AESA.
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32–64.
- Frąckiewicz, M. (2025). Urban air mobility 2025: Air-taxi pilot programs, launch cities, and regulatory hurdles. Techstock.
- Graziato, I. (2026, 24 Gennaio). Perché il cervello sovrastima il rischio degli incidenti aerei. SkyConfidence.
- Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3), 407–434. <https://doi.org/10.1177/0018720814547570>
- Hoffman, R. R., Mueller, S. T., Klein, G., & Litman, J. (2019). Metrics for explainable AI: Challenges and prospects [Technical Report]. Air Force Research Lab (AFRL).
- Hoesterey, H., & Onnasch, L. (2022). Trust in automation: The role of risk and level of automation. [Studio ViRTS].
- Jump, M., White, M. D., Padfield, G. D., Fua, P., et al. (2011, Settembre). myCopter: Enabling technologies for personal air transport systems - An early progress report. Atti del 37° European Rotorcraft Forum (ERF 2011), Gallarate, Italia.
- Kadel, G. R. (2025). eVTOL Certification: Where we stand in August 2025. LinkedIn.
- Kim, Y.-W., Ji, Y., Yoon, S., Colley, M., & Meinhardt, L.-M. (2023, Settembre). The 3rd Workshop on User Experience in Urban Air Mobility: What Could We Learn From AutomotiveUI? 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Ingolstadt, Germania.
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80.
- Luhmann, N. (1979). Trust and Power (G. Poggi, Ed.; H. Davis, J. Raffan, & K. Rooney, Trans.). John Wiley & Sons.

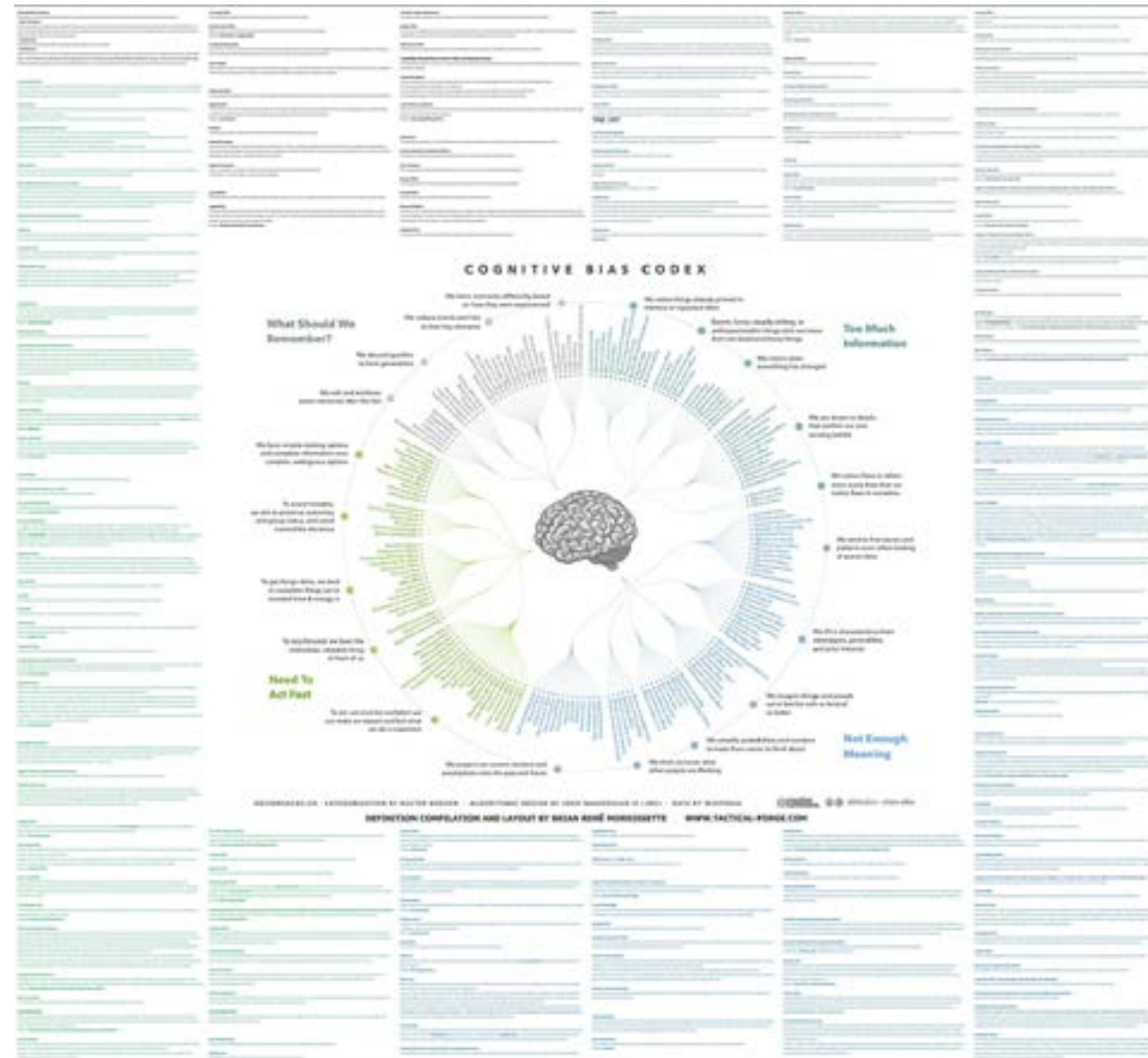
- Manger, C., Vogl, A., & Riener, A. (2025). Cognitive biases in user interaction with automated vehicles: The influence of explainability and mental models. *Applied Sciences*, 15(20), 11030.
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709–734. <https://www.jstor.org/stable/pdf/258792.pdf>
- Mueller, E., Kopardekar, P., & Goodrich, K. (2017, Giugno). Enabling airspace integration for high-density mobility operations. 17th AIAA Aviation Technology, Integration, and Operations Conference.
- myCopter. (s.d.). Enabling technologies for personal air transport systems. <http://www.mycopter.eu/>
- Omeiza, D., Webb, H., Jirotko, M., & Kunze, L. (2021). Effects of explanation specificity on passengers in autonomous driving. *arXiv*. https://www.researchgate.net/publication/372074575_Effects_of_Explanation_Specificity_on_Passengers_in_Autonomous_Driving
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 30(3), 286-297.
- Perfect, P., White, M. D., Padfield, G. D., Burgess, R. K., Jump, M., & Harris, N. (2011, Settembre). Pilot sensitivity to handling qualities-related design parameters for a future personal aerial vehicle concept. 37th European Rotorcraft Forum, DLR.ù
- Rice, S., Keller, D., Hunt, G., & Trafimow, D. (2009). Automation Dependency Under Time Pressure. *International Symposium on Aviation Psychology*, 611-616.
- Rice, S., Winter, S. R., Deaton, J., & Mehta, R. (2019). Predicting consumer willingness to fly in unmanned commercial aircraft. *Technology in Society*, 59, 101150. <https://doi.org/10.1016/j.techsoc.2019.101150>
- Winter, S. R., Rice, S., & Lamb, T. L. (2020). A prediction model of consumer's willingness to fly in autonomous air taxis. *Journal of Air Transport Management*, 89, 101926. <https://doi.org/10.1016/j.jairtraman.2020.101926>
- Rondeau Gomez, A. (2024). Numerical modeling of a coaxial rotor configuration in hover using ROSITA [Tesi magistrale, Politecnico di Milano].
- Saletti, A. (s.d.). Mappa dei bias cognitivi. *Neuromarketing e Psicologia Digitale*.
- Shively, R. J., Lachter, J., Brandt, S. L., Matessa, M., Battiste, V., & Johnson, W. W. (2017). Why Human-Autonomy Teaming? *International Conference on Applied Human Factors and Ergonomics (AHFE)*.
- Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk, and Rationality. *Risk Analysis*, 24(2), 311-322.
- Tech4Future. (2026). Urban Air Mobility, dall'idea dei droni taxi alla sicurezza dei voli aerei. The Edge Company. (2021). Urban Air Mobility: The Edge Company. Tech4Future. The Society of Automotive Historians in Britain (SAHB). (s.d.). Snapshot 295: 1946 Fulton Airphibian.
- Uber Elevate. (2016, 27 Ottobre). Fast-forwarding to a future of on-demand urban air transportation.
- U.S. House of Representatives, Committee on Science, Space, and Technology. (2019). Urban air mobility—Are flying cars ready for take-off? Hearing before the Committee on Science, Space, and Technology, House of Representatives, 115th Congress, 2nd Session (July 24, 2018). U.S. Government Publishing Office.
- Walker, F., Forster, Y., Hergeth, S., Kraus, J., Payre, W., Wintersberger, P., & Martens, M. (2023). Trust in automated vehicles: constructs, psychological processes, and assessment. *Frontiers in Psychology*, 14, 1279271.
- Wang, S., Zhan, X., Zhai, Y., Zheng, L., & Liu, B. (2022). Enhancing navigation integrity for Urban Air Mobility with redundant inertial sensors. *Aerospace Science and Technology*, 126, 107631.
- Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, 52, 113–117.
- Sarah Wells (2021). Aerospace America. The human question. <https://aerospaceamerica.aiaa.org/features/the-human-question/>
- Wikipedia. (s.d.). Personal air vehicle. https://en.wikipedia.org/wiki/Personal_air_vehicle
- Yang, Y., & Kim, M.-Y. (2025). Promoting Sustainable Transportation: How People Trust and Accept Autonomous Vehicles. *Sustainability*, 17(1), 125.
- Yedavalli, P., & Mooberry, J. (2020). An Assessment of Public Perception of Urban Air Mobility (UAM). Airbus UTM.
- Yıldız, E. (2024). Emniyet, Verimlilik ve Performans: Havacılıkta İnsan-Bilgisayar Etkileşiminin Gücü. *Journal of Aerospace Science and Management*, 2(1), 81-95.
- Zhang, Z., Tian, R., & Duffy, V. G. (2021). Trust in automated vehicle: A meta-analysis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 65(1), 1793–1797. <https://doi.org/10.1177/1071181321651398>

ALLEGATI

ALLEGATO A DESCRIPTION OF LEVELS OF DRIVING AUTOMATION

The Level of Driving Automation	Autonomous Driving Level	Label	Vehicle Autonomy's Main Explanation	Dynamic Driving Task (DDT)			DDT Fallback (In the Case of Limited Performance Failure of ADS)	DDT Responsibility	Example of Dynamic Driving Task (DDT)
				Order (Object and Event Detection and Response)	Sustained Lateral and Longitudinal Vehicle Motion Control	Scanning and Reacting to the Driving Environment			
No-Automation (Fully manual) The human driver retains full control and responsibility for all aspects of the driving process	Level 0	No Driving Automation	The vehicle is manually controlled, and human drivers perform all driving tasks.	Human Driver	Human Driver	Eyes on Hands on Feet on	Human Driver	Human Driver	None
	Level 1	Driver Assistance	While human driver control is retained, the vehicle has basic driver assistance with a single automated system to aid the human driver in specific tasks.	Human Driver	Human Driver and System (either lateral or longitudinal)	Eyes on Hands on Feet on	Human Driver	Human Driver	Adaptive Cruise Control; Lane Keeping Assistance; Automatic Emergency Braking; Forward Collision Warning
Low-level Automation (Human driver monitors the road) The autonomous driving system relies heavily on the driver, with the system providing basic assistance features while the driver maintains a predominant role throughout the process.	Level 2	Partial Driving Automation	Combining human driving with automated functions, human drivers monitor the environment and can take control at any time.	Human Driver	System (both lateral and longitudinal)	Eyes on Hands off Feet off	Human Driver	Human Driver	Simultaneous Execution of Tasks; driving maneuvers on highways, including automatic lane changes and navigating through intersections
	Level 3	Conditional Driving Automation	Although vehicles can make informed decisions for themselves through 'environmental detection' capabilities, human driver override is still required.	System	System	Eyes off Hands off Feet off	Fallback ready by use (becomes the human driver during fallback)	Human driver and System	Environmental Detection; Autonomous Task Handling; Human Driver Override; Complex Situation Intervention
High-level Automation (Automated driving system monitors the road) The automation level gradually increases as the system becomes capable of autonomously handling most situations. Human driver intervention becomes more limited, eventually reaching the advanced collaboration stage with complete autonomy.	Level 4	High Driving Automation	Although vehicles can operate in self-driving mode without human interaction, there is still the option for human drivers to override manually.	System	System	Eyes off Hands off Feet off Mind off	System	System	Fully Autonomous Operation; Environment Perception; Autonomous Decision Making; Human Override Option; Safety Monitoring
	Level 5	Full Driving Automation	Although vehicles possess the capability to perform all driving tasks under any conditions, drivers have the option to operate the vehicle alternatively.	System	System	System	System	System	Full Autonomy; Unrestricted Operating Conditions; No Human Intervention Required; Alternative Human Operation Option; Adaptive Driving Abilities; Communication and Interaction; Redundancy and Safety Monitoring

ALLEGATO B.1 COGNITIVE BIAS CODEX WITH DEFINITIONS
 AN EXTENSION OF THE WORK OF JOHN MANOOGIAN BY BRIAN MORRISSETTE



ALLEGATO B.2 RAPPRESENTATE 220 COGNITIVE BIASES
 COMPILED, ORGANIZED AND DESCRIBED IN 5 COLUMNS BY BUSTER BENSON AND JOHN BRADY

