Journal of Design Intelligence
https://www.ffspub.com/index.php/jdi
ISSN 3104-9001 (print)
ISSN 3136-3032 (online)
E-mail: jdi.office@ffspub.com



Article

Intelligent Cockpit Design through Coordination of Driving Tasks and Attentional Resources

Guanghong Xie †
College of Design and Innovation, Tongji University, Shanghai 200092, China
†*E-mail: 2511753@tongji.edu.cn*

Received: September 27, 2026 / Revised: October 20, 2026 / Accepted: November 11, 2026 / Published online: November 23, 2026

Abstract: The value of an intelligent cockpit depends on the fit between information, interaction modalities, and driving tasks. Drawing on an integrative analysis of design research, scenario-based design, attention, and human–automation interaction, this paper proposes a scenario–task–attention–initiative framework. Scenarios are translated into tasks with explicit responsibility, timing, and consequence constraints. Modality selection is treated as a resource-competition problem, while proactive system behavior is bounded by uncertainty, authorization, and recovery conditions. Event-level evidence from use feeds back into problem framing and AI-assisted design. The contribution is a conceptual model operationalized as an analytical design method: an event record links task constraints, admissible service actions, and revisions to design assumptions. Comparisons with four theoretical foundations identify inherited mechanisms and proposed relationships. Worked scenario and task cards, a qualitative decision procedure, and four operationalized propositions make the synthesis inspectable. No driving experiment was conducted; its incremental benefits, classification reliability, and safety implications remain to be empirically tested.

Keywords: Intelligent Design; Intelligent Cockpit; Scenario-Based Design; Attentional Resources; Human–Automation Interaction; System Initiative
https://doi.org/10.64509/jdi.14.165

1 Introduction

Intelligent cockpits bring navigation, environmental controls, infotainment, assistance-state information, and conversational services into one space. These functions do not occur independently: navigation may coincide with hazard detection, and passenger requests may conflict with drivers' ongoing actions. Design must therefore address not only desired functions but when information is valuable, how much attention it requires, and who decides whether to act. Li *et al*. organize automotive intelligent design around users, technology, and scenarios, connecting requirements, computer-assisted and algorithmic generation, and scenario-oriented design [1]. This broad classification requires further translation into cockpit interaction decisions.

Design research is systematically examining AI's influence on practice [2]. Problem–solution co-evolution shows that generating more alternatives and defining problems more accurately are distinct processes [3]; framing organizes subsequent action [4]. In cockpits, an unresolved link concerns how user profiles, sensor data, and scenario labels determine what should appear, be delayed, or be cancelled now. Moving directly from detected fatigue or predicted preferences to proactive services can omit urgency, uncertainty, and attentional costs.

This gap has a human factors basis. Multiple resource theory relates interference between concurrent tasks to resource structure [5], while situation awareness distinguishes perception, comprehension, and projection [6]. Automation may reduce routine actions while leaving people with less-practised, more difficult anomalies [7]. Fewer interaction steps, higher satisfaction, or greater autonomy therefore cannot alone define design success.

This paper asks how user, technology, and scenario drivers become analysable interaction events; how task requirements, attention, and initiative can be coordinated; and how use evidence informs AI-supported design. Its contribution is a conceptual model and an analytical design method, rather than a new general theory of attention or automation. The proposed addition is a shared event representation that makes decisions conditional on task validity, resource competition,

† Corresponding author: Guanghong Xie
*Academic Editor: Yan Hong
© 2026 The author. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

authorization, and recovery, and links failures to revisions of those conditions. Sections 4 and 5 specify the method; Section 6 states tests of its incremental value.

2 Research Approach and Conceptual Boundaries

2.1 Materials And Analytical Procedure

Starting from the three drivers of automotive intelligent design, the paper uses integrative literature analysis to construct a cross-disciplinary framework [8, 9]. The synthesis is problem-oriented and does not claim an exhaustive systematic review, independent double screening, or meta-analytic effect estimation. Its analytical corpus is the 82 references listed in this paper, published from 1977 to 2025; this publication span is not a recorded database search window. The available records do not establish the original search dates, database coverage, full query histories, or exclusion decisions; consequently, the original retrieval process cannot be reproduced from the bibliography alone.

For the revision, a focused traceability check was conducted on 15 October 2026 using web searches directed to publisher and author-institution sources: ScienceDirect, SAGE Journals, and Microsoft Research. Representative queries were "site.sciencedirect.com Carroll 2000 Five reasons scenario based design", "site.sagepub.com Wickens 2008 Multiple resources mental workload 50 3", and "site.microsoft.com research Horvitz Principles mixed initiative user interfaces 1999". Additional title-and-author queries addressed co-evolution, speech interaction, automation uncertainty, takeover synthesis, and human–AI guidelines. The check used publication records and abstracts and, where accessible, author-hosted full text; it verifies selected interpretive links, not full-text coverage of every cited work. Appendix A records the actual queries and references checked.

Sources enter the synthesis when they provide a mechanism, empirical comparison, review synthesis, or design procedure relevant to at least one of five questions: task relevance, resource competition, responsibility, uncertainty, and recovery. Empirical sources are interpreted within their participants, tasks, comparators, and settings. Reviews identify recurring patterns and heterogeneity; conceptual and design-method sources support transfers of reasoning. Commercial claims, undifferentiated preferences, and productivity results without driving outcomes do not establish cockpit safety. These are the organizing criteria for the present synthesis, not a retrospective claim about an independently logged screening process.

Each analytical entry records the reference, evidence type, supported claim, transfer limit, proposed extension, and linked proposition. The sequence is extraction of a source claim, comparison of its scope with a cockpit event, formulation of an explicit decision condition, and identification of an observable counterexample. Table A1 provides the evidence-to-proposition map. Its extensions are authored hypotheses, not findings attributed to the cited studies. The worked cases in Section 5 are constructed consistency checks and supply no participant or deployment evidence.

PRISMA 2020 emphasizes transparent reporting for systematic reviews [10]. Its reporting logic motivates disclosure of the limits here, without relabelling this integrative synthesis as systematic. The focused verification log and evidence map improve inspectability but do not recover missing original search records or demonstrate completeness.

Naturalistic driving, simulation, and scenario questionnaires provide different evidence. Naturalistic studies capture actual task combinations and consequences, but behaviour and road conditions are interrelated. Simulators enable control of cues and tasks but are affected by perceived risk, motion cues, and familiarity. Questionnaires assess understanding and acceptance without directly measuring vehicle control. Evidence is selected according to the question rather than assigned equal strength. Actual-risk claims require relevant driving evidence; interface meaning also requires participant interpretation.

2.2 Two Levels of Intelligent Design

Intelligent design has two connected levels. During design, AI supports organization, scenario expansion, generation, and prototype analysis. During use, cockpit systems provide information, suggestions, or services based on inputs and states. The participants, goals, and error consequences differ, requiring separate criteria. Design-stage outcomes include framing, alternative differences, and verification costs; use-stage outcomes include task performance, attention, understandability, and recovery.

Human–AI co-creation research addresses complementary capabilities in design [11], while collective-intelligence research indicates that knowledge should not be restricted to one designer or model [12]. Participation alone does not ensure coordination. Designers, drivers and passengers, engineers, and algorithmic systems have different judgement responsibilities, and each judgement should be traceable to specific materials and conditions.

Initiative here means the allocation and exercise of authority to initiate, suggest, execute, interrupt, or reverse an information or service interaction. It is distinct from vehicle motion-control authority, driving-automation level, and the presentation format of a cue. The framework neither allocates steering or braking responsibility nor decides when a takeover request is required. A vehicle-specific safety and control architecture supplies those requirements. The service layer receives their priority and transition state as constraints and must not delay, suppress, reinterpret, or compete with their required cues.

Cockpit task boundaries must also be separated from driving responsibility. Speech recognition, suggestion generation, and proactive service capability do not establish responsibility for motion control. Kyriakidis *et al*. examine human factors challenges in automated driving and the need to coordinate technology with human capability [13]. Rather than label scenarios by intelligence level alone, this paper specifies who continuously monitors, responds to changes, and executes or reverses actions.

During a safety-critical transition, the service policy freezes optional prompts and pending confirmation requests,

suspends interruptible service output, and preserves its dialogue state. Necessary safety cues retain the vehicle-specific priority, timing, and modality rules; service suppression must not suppress them. Resumption requires a transition-complete signal from the responsible system and a fresh task and information-validity check, rather than a fixed timeout or a low-workload inference alone. If transition state is uncertain or unavailable, optional services remain deferred. These are proposed architectural constraints requiring integration and human-factors validation, not a certified safety mechanism.

2.3 Translating Three Drivers into an Analytical Unit

The source review divides user requirements into psychological, physiological, and habitual dimensions, extends computational approaches to automotive styling, and considers two directions of user–machine scenario relations [1]. Body-form optimization and cockpit orchestration nevertheless require different objectives and validation methods. The present paper retains the three-driver structure without transferring all styling techniques directly into interaction rules. Table 1 shows how the three drivers translate into cockpit interaction decisions and identifies the design record associated with each one.

Scenario-based design connects concrete activities and communication among participants [14]. A cockpit interaction event is defined here as a minimal unit: within specified participant relationships and system states, a task generates an information or action requirement, humans and systems select an intervention, and recordable consequences follow. Records include scenario, task, attentional demands, system grounds, initiative, action, and consequences. Broad preferences such as liking voice interaction then become questions about who accomplishes what, in which task, through how much dialogue.

Orchestration means selecting and coordinating the content, modality, timing, and permitted initiator of information and services across concurrent tasks. It includes deferral, cancellation, and state-preserving resumption. “Scenario” denotes a bounded description of actors, conditions, and an unfolding event; contextual variables are fields within that description. “Initiative” denotes observable interaction choices and authorization, whereas the broader concept of agency is not used as a framework component. Table C3 gives codable definitions.

Table 1: Translating three design drivers into cockpit interaction decisions

Design Resource	Original Focus	Additional Judgement	Inspectable Output
Users	Psychology, physiology, and habits	Whether preferences suit the current task	Requirement sources and task cards
Technology	Assistance, generation, and recognition	What is possible and when action is appropriate	Capability boundaries and error conditions
Scenarios	User–machine relationships	Responsibility, timing, and concurrent tasks	Scenario cards and transition records

Note: The first two columns summarize the design drivers in [1]; the final two present this paper’s conceptual extensions.

3 Complementary Contributions And Unresolved Links

3.1 From Preferences To Situated Task Evidence

Subjective experience is indispensable for understanding requirements. Research on interaction naturalness distinguishes ease and fluency from responsiveness to situations and people [15]. Perceived naturalness or convenience, however, does not establish low driving workload. Questionnaires and interviews explain values and experiences, whereas observation and process measures examine performance and interference. These methods answer different questions.

Physiological sensing demonstrates opportunities to analyse stress during real driving [16], and deep convolutional networks have been used to classify driving signals [17]. Richer process data do not mean that each physiological change identifies one emotion or each behaviour category specifies a service need. Sensor errors, baselines, road conditions, and task content affect interpretation. Observations, inferred meanings, and selected actions should therefore be recorded separately, with alternative explanations retained.

Situated observation and prototype participation can recover experiences lost in labels. Bodystorming emphasizes entering activity settings [18], sensory ethnography examines situated experience [19], and probes, toolkits, and prototypes offer different co-design approaches [20]. These may reveal passenger coordination, noise, reach difficulties, and social meanings of cues. They are proposed evidence-gathering methods here, not fieldwork already conducted.

Attention research distinguishes distraction from broader inattention. Regan *et al.* show that insufficient attention to driving is not always caused by an external device [21]. Cockpits can create competition or support already difficult tasks. Attributing all errors to screens overlooks fatigue, comprehension, and environmental complexity; attributing them to inattentive drivers overlooks interface organization. Research should specify the object, timing, and duration of diverted attention.

Naturalistic evidence moves analysis beyond preferences. Klauer *et al.* compare distraction and risk among novice and experienced drivers [22], while Dingus *et al.* examine crash-related factors and prevalence [23]. These studies support analysing specific behaviours rather than generic device use. Entering, searching, reading, and confirming create different visual and motor requirements. Observed associations

remain influenced by exposure, context, and selection, so risk estimates cannot be directly assigned to untested interfaces.

Requirements analysis should therefore record the goal, acceptable interruption window, abandonment condition, and cost of nonresponse alongside stated preferences. Workload and safety outcomes remain separate from preference ratings. The operational treatment of these fields is specified in Sections 4.2–4.4 and the worked cards in Section 5.1.

3.2 From Additional Modalities To Resource Competition

Automotive multimodal research considers voice, touch, and other channels [24], but multimodality does not require simultaneous activation. Oviatt's analysis of common misconceptions emphasizes actual use rather than intuitive assumptions about combined benefits [25]. Design must determine which information needs complementary presentation, which needs redundant cues, and which should not yet appear.

Voice can reduce visual and manual interaction, but comprehension and response still require cognition. Dual-task simulation and meta-analysis of phone conversations show that hands-free interaction does not eliminate distraction [26, 27]. These findings do not quantify every voice service's risk, but reject the general assumption that speech is attention-free. Modality choices should consider length, response requirements, semantic complexity, and concurrent tasks rather than touchscreen counts alone.

Three relationships are distinguished: complementary channels divide necessary information, redundancy improves detectability of critical cues, and substitution accommodates unavailable or excessively interfering channels. Each depends on task conditions. Deferring nonurgent information may be preferable to delivering it through another modality, whereas excessive simplification of critical information may impair understanding. Multiple resource theory guides analysis, but interface form alone does not establish workload [5].

Early phone-use meta-analysis [28], texting meta-analysis [29], and an updated systematic review and meta-analysis [30] show that costs depend on task form. Reading, typing, dialling, and sustained conversation are not equivalent secondary tasks. Brief confirmations and open-ended dialogues likewise differ. These findings cannot supply a universal effect size for modern assistants, but support decomposing interactions by semantic, operational, and temporal structure.

Gaze position alone does not establish attention. Strayer *et al*. show that looking at the road does not guarantee processing relevant information during phone conversations [31]. Lee *et al*.'s speech-email study also indicates cognitive demands and effects on driving response [32]. Eye movements should be interpreted with event detection, comprehension, and action. Reduced off-road glances accompanied by more missed events cannot establish lower overall interference.

Multimodal value also depends on correspondence among signals. Spence and Ho review multisensory warnings for perception and driving [33], while Ho *et al*. examine in-car combinations [34] and spatial correspondence [35]. These studies shift attention from channel counts to temporal and spatial organization. Cue position may aid directional hazard understanding; elaborate spatial coding for general states may add learning costs. Combinations require task-specific testing.

Detection, identification, and action should be distinguished. Louder sound may improve detection without conveying meaning; haptics may reduce screen reading while requiring learned rhythms; redundant channels may also amplify false-alarm interference. Comparisons should record whether cues are noticed, understood, and followed appropriately. This identifies the level improved and whether added modalities yield net benefit.

3.3 From Autonomy To Appropriate Reliance

Automation research distinguishes acquisition, analysis, decision selection, and implementation [36]. Cockpits may offer strong assistance in some functions while retaining human decisions in others, without one uniform autonomy level. Research on use, misuse, disuse, and abuse further locates problems within design–use relationships rather than willingness to accept technology alone [37].

Trust should support reliance matched to capability rather than ever-higher ratings [38]. Human–robot trust meta-analysis identifies the importance of factors including performance [39], providing a comparative perspective without directly transferable effect sizes. Correct acceptance, correct rejection, erroneous acceptance, and erroneous rejection should be recorded separately rather than concealed within general favourability.

Autonomy also changes state understanding. Out-of-the-loop research identifies complex relationships between automation and recovery [40], while mode-awareness research reveals new demands associated with functions and modes [41]. Changing information initiative without explanation may leave users uncertain why something appeared, whether it has executed, and how to cancel it. These are design tasks, not problems solved by adding explanatory text at the end.

Trust calibration has multiple sources. Hoff and Bashir distinguish disposition, situation, and learning [42], and Schaefer *et al*. compare influences on automation trust [43]. Appearance alone does not explain trust. Introductions, experience, repeated success, and sudden failures may jointly affect reliance. Measurements after brief demonstrations cannot adequately represent long-term judgement.

Körber *et al*. show the relevance of initial system communication to trust and reliance [44]. Verberne *et al*. examine shared driving goals, information, trustworthiness, and acceptance using descriptions of adaptive cruise control [45]. Evaluation of descriptions must be separated from driving performance: greater trust does not establish better recognition of limits. Advertising, onboarding, and in-vehicle information should communicate consistent capabilities.

Parasuraman and Manzey emphasize attention in automation bias and complacency [46]. The question becomes whether systems support effective monitoring, rather than merely instructing drivers to remain vigilant. Displays communicating only normal operation without task limitations make targeted checking difficult; excessive confirmations can also devalue cues. Calibration requires visible boundaries, detectable errors, and recovery paths rather than persistent reminders.

3.4 From Generation Speed To Judgement Quality

Experiments support productivity gains in writing [47], but meta-analysis shows that human–AI outcomes depend on task and comparator [48]. Design-team research also identifies unintended consequences [49]. Text or image generation speed therefore cannot establish requirements quality, and candidate counts cannot stand for innovation.

Creative research shows that improved individual outputs can coexist with reduced collective diversity [50]. Discussions of illusions of understanding caution against allowing tool capability to obscure epistemic limits [51]. Models can expand scenarios and alternatives, but scenario importance, strategy acceptability, and causal interpretation require different evidence. Design must reserve resources for challenge, comparison, and abandonment.

4 A Scenario–Task–Attention–Initiative Framework

4.1 Structure And Theoretical Position

Four connected judgements form the framework. Scenarios specify road, vehicle, participant, and system conditions. Tasks specify goals, responsibility, timing, and error consequences. Attention analysis examines resource use and competition. Initiative determines whether to present, defer, ask, execute, or withdraw. Figure 1 connects use-stage judgement with design-stage revision. These are proposed mechanisms, not verified causal pathways.

The framework inherits situated activity descriptions from scenario-based design [14], interference mechanisms from multiple resource theory [5], uncertainty-sensitive timing and human control from mixed-initiative interaction [60], and revision of problem–solution relationships from co-evolution [3]. None of these elements is claimed as new. Wicked problems [52], design under constraints [53], and designerly knowledge [54] further explain why a single unconstrained optimum is unsuitable. Table B2 compares the foundations by unit of analysis, variables, outputs, scope, and validation.

The proposed relationship is a common event record that carries a task's responsibility, validity window, and consequences into resource assessment and then restricts the admissible initiative set. The selected action and its recovery outcome can, in turn, invalidate the original task assumption. Applying the theories in parallel could reach the same decision; the claimed increment is the explicit dependency and auditable propagation of a changed condition. This yields conditional predictions about unnecessary interruption, recoverable reliance, and cross-scenario failure detection, rather than a claim that earlier theories ignore tasks, attention, or deferral. Section 5.1 isolates this increment against a specified scenario-label implementation.

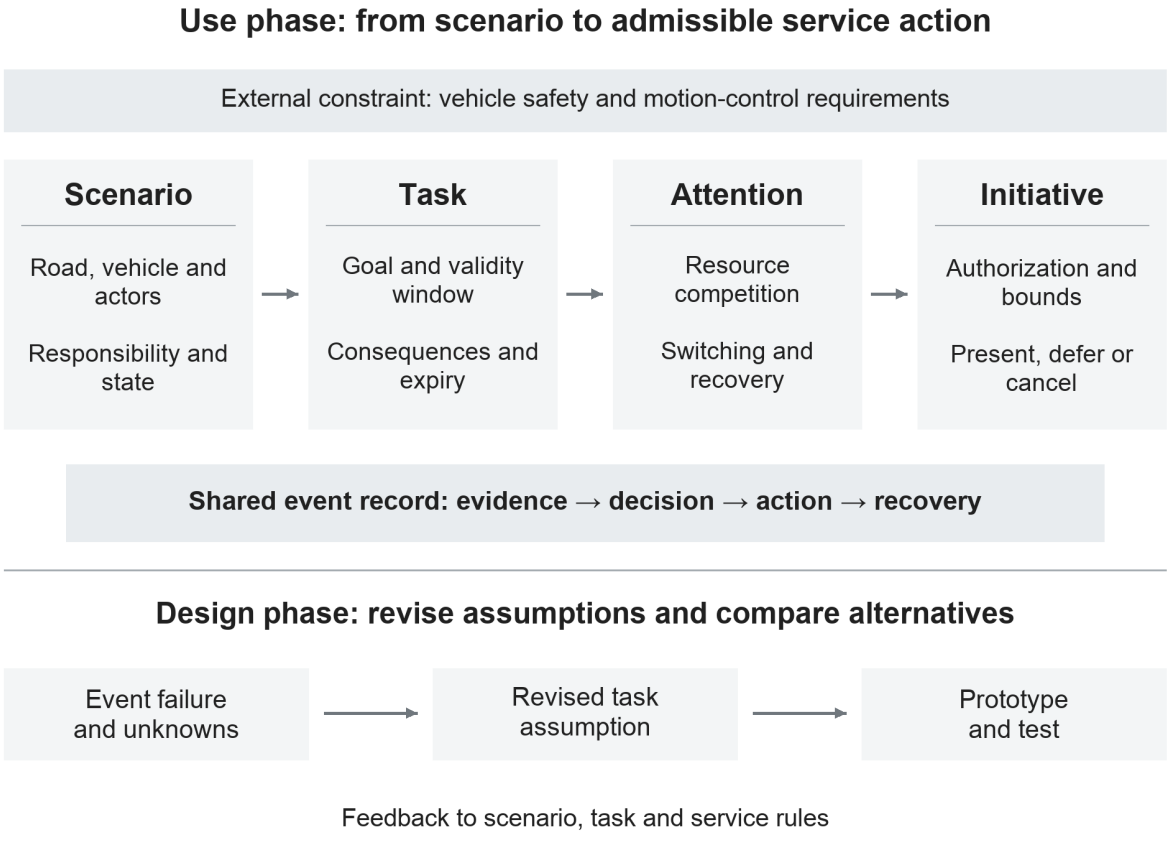


Figure 1: The scenario–task–attention–initiative design framework. (Arrows indicate proposed dependencies and feedback, not verified causal effects. Safety and motion-control requirements constrain service decisions from outside the framework).

4.2 Translating Scenarios Into Tasks

Labels such as commuting, long-distance travel, and rest are insufficient. One commute includes parking, ordinary travel, complex turns, and state changes with different demands. Scenario cards should record motion state, road complexity, responsibility, participants, current tasks, available time, and unknowns. Task cards further specify goals, success criteria, permissible delay, and failure consequences.

Vehicle-control tasks, driving-related information tasks, and non-driving services should be distinguished without imposing a permanently fixed priority order. Route planning may be deferrable until an imminent turn changes its temporal value. Entertainment normally should not compete with critical driving cues. Priorities need explicit grounds rather than defaulting to notification order or commercial goals.

Drivers and passengers must also be distinguished. Passenger spare capacity does not imply driver tolerance for the same display or sound. Shared displays and sound fields should identify both operators and affected people. Even independent passenger displays require assessment of reflections, audio leakage, and passive driver involvement; account or seat separation alone does not establish absence of interference.

Translation must preserve continuity around events. Navigation before a junction and confirmation after turning belong to different action stages despite sharing one function. Records should specify prior activity, interrupted judgements, and tasks needing resumption. Systems should preserve necessary state when dialogue is interrupted rather than require restarting. Whether and where to resume are interaction costs.

For each task, permissible delay ends when the information ceases to support the intended decision, an external deadline is reached, or its state assumptions change. The card specifies the earliest admissible delivery opportunity, expiry condition, and consequence of omission. An expired instruction is cancelled and, if a need remains, regenerated from the current state. Priority conflicts are resolved first by protected safety requirements, then by the consequence and deadline of currently valid tasks, and finally by interference and recovery costs. Arrival order is only a tie-breaker among otherwise equivalent optional services.

4.3 Joint Assessment Of Attention And Information Value

Attention is not a fixed balance directly read by a sensor. It is a design estimate combining task analysis, behaviour, and state inference. Costs include off-road vision, auditory competition, manual occupation, semantic processing, switching, and recovery. Embodied and extended cognition locate task support across bodies, environments, and tools [55]; control placement, road visibility, and feedback therefore contribute to attention organization.

Candidate information should first be assessed for task relevance and consequences of delay, then for current resources. Urgent relevant information enters validated critical-cue channels; relevant but deferrable information enters a queue. High predictive probability alone does not warrant irrelevant notifications. Preferences can select among acceptable options but should not override responsibility or validated constraints.

The output is a testable content–modality–timing combination, not an abstract low-workload interface. Shortening a long explanation requiring immediate confirmation may be insufficient if the information should be deferred altogether. Conversely, removing critical state information reduces display content but may obscure limitations. Evaluation should examine attentional costs and comprehension rather than minimize information alone.

Workload does not directly mirror interface complexity. Paxion *et al*. examine task demands, effort, and performance [56]; Warm *et al*. show that sustained vigilance itself can demand effort and create stress [57]. Fewer actions do not ensure lower cognitive costs, and automation may leave monitoring burdens after reducing manipulation. Resource conditions vary with tasks, experience, and time and cannot be reduced to one score governing all strategies.

Physiology supplements subjective reports but cannot independently determine interruption timing. Charles and Nixon's systematic review identifies conditions and limitations of different measures [58], while another systematic review shows inconsistent validity across tasks [59]. Baselines, environmental change, and movement artefacts require consideration, with measures aligned to explicit task changes. Workload-classification accuracy must also reflect asymmetric error costs: mistaking high workload for low has different consequences from the reverse.

Trade-offs can be tested through contrasting events without first building a complex index. Presenting identical information before, during, and after tasks permits comparison of comprehension, interference, and recovery. Earlier delivery may aid preparation but increase memory demands; later delivery may reduce interruption but cause omissions. Such experiments operationalize attention through timing comparisons rather than use satisfaction to validate an entire orchestration policy.

4.4 Bounded Initiative Under Uncertainty

Mixed-initiative research addresses coordination between automated service and direct manipulation [60]. Here, actions include observation, suggestion, execution after confirmation, execution within authorization, and stopping or rollback. These are alternatives selected according to consequences, evidence, and authorization, not successive maturity levels. Capability communication, correction, and control guidelines can assess their design [61].

A service is eligible for low-consequence, reversible execution only if all of the following are satisfied in the current event: it leaves motion control and safety-relevant visibility or alert audibility unaffected; creates no immediate manoeuvre demand or material external commitment; remains inside an explicit bounded authorization; preserves the prior state; and can be stopped or undone without a new time-critical driver action or lasting harm. Severity and reversibility are assessed separately. If any condition is unknown, the action does not enter this category. An available undo button alone is insufficient.

For example, a previously authorized small cabin-temperature adjustment may qualify when it does not affect

demisting, visibility, alert audibility, or immediate comfort-related distraction. Changing defogging or creating a distracting fan-noise increase does not qualify by virtue of belonging to HVAC. A proactive route change is usually consequential: although the route can be restored on screen, a missed exit, lane change, toll, or time loss may not be reversible. It therefore remains a proposal until suitable confirmation, and is deferred near a manoeuvre or critical transition. Classification attaches to the action-in-context, not the feature name.

Valid authorization identifies the authorizing person, affected users, action type, allowed extent, applicable situation, duration or expiry, and revocation route. It must be explicit and current; inferred preference, silence, or past acceptance of a different action is not consent. Context changes trigger rechecking. The requester cannot authorize an effect outside their scope, such as a passenger request that commandeers the driver's audio channel. Revocation prevents new service execution and stops or restores an ongoing action only through its approved recovery path.

Prediction models and decision policies are distinct technical objects [62, 63]. The qualitative procedure in Table 2 first checks the external safety boundary, then information validity, task urgency, authorization, consequence and recovery, and resource competition. Candidates that fail a necessary condition are excluded before preference or convenience is considered. This is a proposed order of reasoning, not a deployed controller: it supplies neither universal numerical thresholds nor empirically calibrated weights.

When no optional proactive action is admissible, non-intervention is an explicit decision with a recorded reason and review condition. Deferral requires a queue entry and a resumption rule; cancellation requires an expiry or invalidity reason. Neither may remove necessary warnings or required state cues. A critical-cue omission remains a failure even if the service policy otherwise reduces interruption.

Beller *et al*. examine uncertainty information in automated driving [64], offering a reference for communicating limitations. This need not mean continuously displaying probabilities. Drivers need to know whether grounds are adequate for action, what limits apply, and whether environmental checks are needed. Numerical confidence may suit retrospective analysis but not every immediate event. Uncertainty information should be organized around action requirements and evaluated for comprehension and calibrated reliance.

Körber *et al*.'s study of post-takeover explanations makes timing an explicit research question [65]. Pre-takeover cues should support necessary action; later explanations can improve system understanding. Detailed explanation should not be concentrated at urgent moments. Bounded initiative includes deferring explanations, but not permanently hiding reasons: accessible records and opportunities to challenge intervention should follow when tasks permit.

Uncertainty may concern perception, user intent, authorization, or transition status. Records distinguish these sources instead of collapsing them into one confidence score. Uncertain optional services are deferred; clarification is requested only if its expected decision value justifies the interruption and an admissible opportunity exists. Uncertainty about a required cue is handled by the responsible safety system, not by deleting it from a service queue. For conflicting passenger requests, the policy identifies affected users and preserves the driver's protected channels before seeking a cancellable arrangement.

Table 2: Qualitative decision procedure for information and service initiative

Ordered Check	Condition	Decision	Record or Fallback
1 Protect	Required cue or active safety transition	Retain external cue policy; suspend optional output and confirmation	Preserve service state; uncertain transition keeps services deferred
2 Validate	Content expired, superseded, or no longer relevant	Cancel; regenerate only if a current need remains	Record expiry reason and replaced event; never replay a stale cue
3 Prioritize	Multiple valid tasks compete	Protected requirements first; then consequence, deadline, and interference	Retain rejected alternatives and conflict rationale
4 Authorize	Execution requested; scope absent, expired, or ambiguous	No execution; seek confirmation only at an admissible opportunity	Record actor, scope, duration, context and revocation path
5 Bound	All low-consequence and reversibility criteria in Section 4.4 satisfied	Limited authorized execution may be a candidate	Otherwise propose or defer; preserve prior state and approved recovery
6 Match	Grounds adequate but current resource competition too high	Defer optional content; do not merely add another channel	Store reason, task window, expiry and release condition
7 Resume	Critical transition ended and task permits interaction	Recheck relevance, authorization and state before offering pending content	Cancel invalid content; record action and observed recovery separately

Note: Apply the rows in order, re-entering the procedure after a material state change. Protected cues follow the vehicle's own validated safety architecture. Consequence categories and release conditions require vehicle-specific validation.

4.5 Feeding Use Evidence Back Into Design

Event records should return to design: what triggered action, what grounds were used, whether users understood, what was interrupted, and how recovery occurred. These records revise task cards, generation constraints, and evaluation questions. Understandability, feedback, and mapping principles help explain failures [66]. Heuristic evaluation can identify evident problems but cannot replace behavioural validation in driving [67].

This loop centres on traceable judgement rather than automatically returning all use data to training. Model updates, interface changes, and rule revisions should have separate versions and impact checks. Goals should include well-being and control rather than only interaction counts or service duration [68]. When feedback invalidates the problem definition, removing a function may better reflect co-evolution than perpetually optimizing it [3].

The event-record template contains: event and rule-version IDs; scenario state and its provenance; task ID, responsible actor, goal, deadline and expiry; competing resource demands; candidate actions and rejected alternatives; authorization and consequence checks; selected action and reason; queue state and release condition; observed detection, understanding and action; error and recovery outcome; and the task or design assumption to revise. Unknown values and unobserved outcomes are explicitly labelled. Section 5.1 completes this template for a constructed event; it does not fill observational fields with hypothetical successes.

5 Scenario Analysis And Implementation

5.1 Worked Scenario And Task Cards

The constructed case concerns a driver and passenger travelling through a city with an active route, a passenger music request, and an optional route-service suggestion. The same request is analysed while parked (S0), while approaching a complex junction (S1), and during a system-defined assistance transition (S2). These are alternative states of an illustrative event, not observed trials or an assumption that every vehicle supports automated driving. The scenario cards in Table 3 and task cards in Table 4 specify the conditions under which the decision changes.

The specified baseline is a coarse scenario-label policy: “urban trip” permits one brief spoken route suggestion when a predicted preference matches and no safety transition is active. The task-conditioned policy uses the same candidate wording, modality, preference, and transition safeguard,

Table 3: Completed scenario cards

State and participants	Responsibility and concurrent demand	Decision and unresolved condition
S0 Parked before departure; driver and passenger; route comparison requested	Vehicle stationary in the constructed case; no active control transition; route comprehension is the current task	Offer comparison and obtain explicit route-change confirmation. Parking does not itself authorize execution
S1 Approaching a complex urban junction; same people and candidate proposal	Driver monitors traffic and executes manoeuvre; visual, manual and semantic demand; current turn cue remains necessary	Retain necessary cue; defer optional route comparison and competing music output. Release requires post-junction task recheck
S2 System-defined assistance transition; same optional requests pending	Responsibility and required action come from the vehicle safety architecture; attention must reorient to the required task	Freeze optional prompts; preserve dialogue state. Resume only after an authoritative transition-complete signal and validity checks; unknown state means continued deferral

Table 4: Completed task cards for event E1

Task and success criterion	Timing and expiry	Resource and consequence	Authorization and recovery
T1 Current manoeuvre information; driver receives and understands the required current-route cue	Within the vehicle-specific decision window; obsolete after the relevant turn or route-state change	Must support traffic monitoring; omission and stale instructions are separately coded failures	Required cue follows external architecture; expired instruction is not replayed; current guidance is regenerated if necessary
T2 Optional route comparison; user understands alternatives before any route change	Can wait until after the junction; expires if alternatives, location or intent change	Semantic comparison competes with manoeuvre; changed route may entail irreversible travel costs	No execution authorization in E1; store original route and pending proposal; recheck and ask later or cancel
T3 Passenger music request; service delivered only within affected users’ permissions	Deferrable; expires on withdrawal, changed request, or lost relevance	Shared audio can mask cues or involve the driver; passenger input alone does not remove competition	Passenger request is not authorization to disrupt protected audio; queue state is retained and revalidated before resumption

Note: Cards S0–S2 and T1–T3 are conceptual design records. Success criteria specify what a study would test; they are not achieved outcomes. Event E1 is the completed decision record described in Section 5.1.

adding the current manoeuvre, relevance window, and recovery state. At S1, the baseline presents the suggestion, while the proposed policy defers it because an optional route comparison would compete with junction monitoring. The necessary current-route cue is retained under the same vehicle-specific rule in both conditions. At S0, both may present a proposal for confirmation; at S2, both suppress optional services. This comparator tests the added task condition, not scenario-based design as a whole.

For S1, event E1 records: rule version V1; driver responsible for the current manoeuvre; turn-relevant navigation valid only before the decision point; a route-comparison proposal that remains optional; high visual and semantic competition; no current authorization to change the route; and a reversible proposal but potentially irreversible route consequences. The decision is to retain the required navigation cue, defer the proposal, and avoid a new confirmation dialogue. A passenger music request is queued if shared audio would compete. Neither the route change nor shared-audio execution is inferred from preference.

E1's recovery path preserves the proposal's source, original route, pending status, and expiry condition. After the junction and any transition have ended, the policy rechecks the current route and requester intent. A still-valid proposal may be offered at an admissible opportunity; a passed-turn instruction or obsolete proposal is cancelled and, if needed, regenerated. The user must not be told that a pending action has executed. Detection, comprehension, actual interference, and successful recovery are marked "not observed" because E1 is constructed. A later test challenges V1 if required cues are missed, a stale message is resumed, or users mistake deferral for execution.

5.2 Takeover As a Process

Nonurgent and urgent transitions require different temporal assumptions [70], and study conditions systematically influence takeover time [71]. Evaluation should distinguish noticing a request, understanding the task, beginning control, and stabilizing performance. First steering-wheel contact or button presses alone omit subsequent control quality.

Experiments on resuming control support attention to transition processes [72, 73]. Reviews also show that workload and situation awareness under automation depend on factors including non-driving tasks [74]. These sources motivate attention recovery but do not validate this framework. Figure 2 therefore presents stages and comparisons without hypothetical effect values.

Lu *et al*. conceptualize transitions between human and automated states, including monitoring and control relationships [75]. McDonald *et al*. synthesize empirical takeover research and driver models, noting that determinants of response time and subsequent control are not identical [76]. Evaluation should distinguish request detection, attention reorientation, traffic understanding, action initiation, and stable

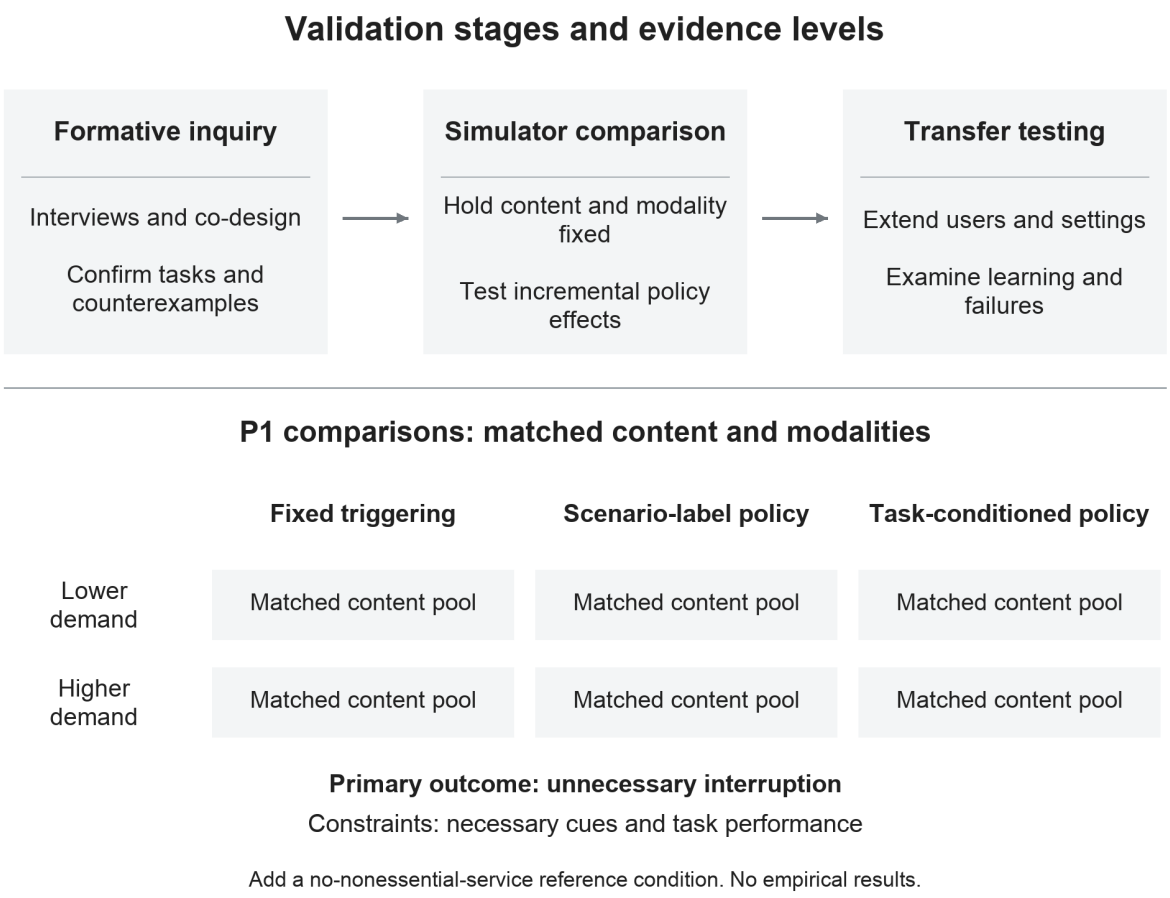


Figure 2: Staged validation and comparison of triggering strategies. (Note: This proposed protocol contains no measured results. Modality and initiative effects require separate subsequent comparisons.)

control. Earlier action may indicate responsiveness or premature intervention.

Gold *et al*. demonstrate the relevance of traffic density [77]. Alongside takeover-time meta-analysis [71] and workload and situation-awareness reviews [74], this suggests that available time, feasible road actions, and prior tasks jointly shape transitions. Mean response times from individual experiments should not become universal thresholds; scenarios, control tasks, cues, and permissible risk conditions must be explicit.

Repetition may also change results. First exposure requires rule comprehension, whereas later trials may encourage anticipation. Aggregate means can conceal whether an interface improves initial understanding or skilled operation. Future work should distinguish first exposure, learning, and stable performance, counterbalancing scenario order where necessary. Unexpected-failure tests should avoid predictable repetition that reveals event timing.

5.3 Implementation For Design Teams

Implementation begins with task and responsibility assumptions, followed by minimal prototypes comparing fixed presentation, scenario-label policies, and coordinated task–attention policies. AI can construct counterexamples involving missing sensors, simultaneous speech, incorrect preferences, or sudden road change, but researchers must check relevance and retain provenance and edits.

Comparisons should not give an intelligent condition richer content, better wording, and more modalities while simplifying the baseline. A first round holds content and modality constant and varies triggering; a second holds triggering constant and compares modality relationships; only later are components combined. This separates information selection and resource coordination from wording improvements.

Three inspectable outputs are recommended: evidence-linked task cards, strategy comparisons retaining rejected options, and changes connecting versions to test results. These prevent generation speed from obscuring judgement and support cross-role discussion. Graphic polish, conversational fluency, and model size may influence presentation but do not replace these outputs.

6 Propositions And Validation

6.1 Four Falsifiable Propositions

P1, task mediation: With the same candidate content and modalities, a policy conditioned on task relevance, validity, and responsibility is expected to reduce unnecessary interruption relative to the scenario-label policy specified in Section 5.1, particularly under high concurrent task demand. Support also requires that necessary-cue omission and primary-task performance remain within prespecified acceptable bounds. A reduction in prompts alone is not support. The proposition concerns the task layer's incremental value within this comparator.

Manipulate triggering policy and scenario task demand while holding the candidate pool, wording, modality, alert safeguards, and exposure opportunities constant. The unit is an eligible interaction event nested within driver and scenario. The primary endpoint is the proportion of eligible events with an unnecessary interruption. Code an interruption when an optional system output demands a response or coincides with a task switch; code it as unnecessary when independent, pretrial task annotations show that it has no current task need or could wait without losing required value. Separately report system-initiated prompts and observed behavioural interruption. Code a necessary-cue omission when a prespecified required cue is absent, late beyond its task window, or not detected or understood; report these causes separately. Secondary endpoints include hazard response, resumption errors, and workload. Evidence of smaller interruption with bounded omission and performance costs supports P1; a meaningful increase in either critical cost contradicts its benefit claim.

P2, resource matching: The effect of a modality arrangement is expected to interact with concurrent resource demand. When an optional message competes with the ongoing task, delayed delivery is expected to reduce primary-task interference relative to immediate delivery without making the message unusable. Complementary, redundant, and substitutive arrangements are distinct comparisons, not interchangeable examples of "more modalities". Benefits attributable only to message length or practice do not support a resource-matching explanation.

Use event-level repeated measures within drivers. First compare visual and auditory substitution with matched meaning and response requirements across visual-manual and auditory-verbal concurrent demands. Test redundancy and complementarity in separate planned contrasts with the same essential semantic content. Independently cross immediate versus deferred delivery where the task permits delay. Match or explicitly factorially manipulate word count and semantic complexity; train cue meanings to a stated criterion and measure familiarity, rather than confounding them with channel number. The primary endpoint is a prespecified scenario-relevant primary-task interference measure; comprehension and resumption cost are secondary, and expired delivery is a validity failure. The predicted resource-demand interaction must persist after these controls. A precise absence of the planned interaction, worse interference, or benefits wholly explained by length or familiarity disconfirms the proposed mechanism in that scope.

P3, bounded initiative: Under perceptible service errors, explicit and current authorization, visible capability limits, and accessible recovery are expected to improve appropriate reliance and recovery relative to a service interface without this added support. The basic vehicle safety architecture is identical in both conditions. Support does not require maximal trust or satisfaction, and additional confirmation workload can restrict the benefit.

Manipulate the support package and valid versus invalid service recommendations with matched error exposure; later component studies can isolate authorization, explanation, and recovery. The unit is a recommendation or recovery episode nested within driver. Define recommendation validity beforehand from the scenario requirements and authorized scope. Report correct acceptance and incorrect rejection among valid recommendations, and correct rejection and incorrect acceptance among invalid recommendations, each with its own

denominator. Nonresponse is reported separately and is not automatically coded as rejection. Co-primary endpoints are incorrect acceptance of invalid advice and successful restoration of an acceptable state after error; secondary outcomes include the other decision categories, recovery time and actions, confirmation workload, trust, and satisfaction. Improved discrimination and recovery without unacceptable primary-task costs support P3. Higher trust without improved discrimination, persistent unsafe acceptance, or recovery costs that outweigh the benefit restrict or contradict the claim.

P4, feedback translation: With equal time and source materials, design teams using event records and explicit task-assumption revisions are expected to identify more valid failures in held-out scenarios than teams receiving aggregate feedback from the same underlying events. Traceable judgement is a separate process outcome. The proposition predicts useful changes to design reasoning and failure detection, not more documentation alone or demonstrated road-safety gains.

Randomize whole teams to feedback formats, balancing prior design experience where feasible; keep the brief, event source, tool access, time, and feedback exposure comparable. Evaluate teams on unseen scenarios and a prespecified failure rubric. Independent raters, blinded to allocation where feasible and not involved in producing the designs, assess each claim for a cited event, an explicit task assumption, a justified decision change, and a testable consequence. Record correctness and false positives, rather than rewarding record length. The team is the unit of allocation and inference; ratings and scenarios are nested observations. The primary endpoint is valid held-out failure detection, with traceability, false-positive rate, and prototype performance as secondary endpoints. Analyse team-level summaries or hierarchical models accounting for team, scenario, and rater dependencies. More traceable records without improved valid failure detection do not support the principal benefit; imprecise results remain inconclusive.

Table 5 summarizes the four propositions and their disconfirming conditions. Comparisons should specify competing explanations. Fewer cues may reduce interference or omit necessary information; dynamic strategies may increase trust through capability fit or merely fluent presentation; feedback groups may benefit from longer discussion. Content, exposure, and basic functions should be comparable while rules or explanations are varied. Otherwise results establish package-level differences rather than specific mechanisms.

6.2 Staged Validation And Measures

Validation has three stages. Interviews, co-design, and low-fidelity prototypes first examine whether task categories capture actual problems. Thematic analysis can retain interpretive links from materials to themes [78], without equating frequency with importance. Simulator studies then manipulate policy and task load to assess behaviour and interaction costs. Controlled studies subsequently examine transfer to more complex settings and longer use, clearly distinguishing experimental findings from actual road risk.

P1 could cross policy with road-task load, using the same content pool and a reference condition without nonessential services. Outcomes include unnecessary interruption, missed critical cues, hazard response, and task resumption, supplemented by off-road vision and workload. NASA-TLX can measure subjective workload [79], but lower ratings alone do not establish improved performance. Situation-awareness measures should specify timing and target knowledge rather than substitute trust for understanding [6].

For P1–P3, preregister the event-coding manual, scenario-specific required cues, acceptable recovery state, primary contrasts, and meaningful-benefit and harm bounds. Calibrate coders on separate practice material, then independently code test events and report agreement and adjudication [82]. Use participant and scenario effects for repeated measures, and the team-based analysis specified for P4. An uncertain confidence interval is not evidence of equivalence; “no harm” requires a justified bound and adequate precision. The present paper supplies operational definitions, not arbitrary universal numerical cut-offs.

Table 5: Principal comparisons and disconfirming conditions

Proposition	Comparison and unit	Primary endpoint	Disconfirming evidence
P1 Task mediation	Task-conditioned vs specified scenario-label policy; driver-event	Unnecessary interruption; omission and task performance as constraints	No meaningful reduction, or unacceptable critical-cue or task cost
P2 Resource matching	Modality arrangement × concurrent demand; driver-event	Prespecified primary-task interference; planned interaction	Precise absence of interaction; length or familiarity explains benefit
P3 Bounded initiative	Support package with matched valid/invalid advice; decision or recovery episode	Incorrect acceptance; successful recovery	Trust alone improves, inappropriate acceptance persists, or costs dominate
P4 Feedback translation	Event records vs aggregate feedback; randomized team	Valid failure detection in held-out scenarios	More documentation without improved failure detection

All propositions remain to be tested. P1–P3 evaluate use strategies; P4 evaluates design processes. Process benefits are not equated with road-safety benefits.

Sample size should be justified beforehand through the primary comparison, smallest meaningful effect, and precision goals rather than customary counts [80]. Primary endpoints, exclusions, missing-data handling, and multiplicity should be prespecified. Effect sizes and uncertainty should accompany equal consideration of null and opposite results. No participant counts, significance tests, accuracies, or improvement percentages are reported because no sample was collected.

Measures should follow questions rather than available equipment. Allocation questions may combine eye tracking and event detection; understanding requires state judgements or action choices; recovery requires post-error processes. Workload, physiology, and control data can complement one another, with primary outcomes specified to avoid selective significance. Disagreements should be interpreted by measurement level rather than resolved by standardizing and averaging everything.

Lakens links effect-size reporting with cumulative evidence [81]. Repeated-measures studies should identify effect-size definitions and calculation conditions and report uncertainty. Significance does not establish meaningful benefit, and nonsignificance does not establish equivalence. Sample justification [80] should reflect meaningful differences, precision, or resources rather than copy prior studies. Equivalence claims require appropriate tests and prespecified bounds.

Observer-agreement methods can assess whether scenario and error coding is reproducible [82], but agreement does not ensure adequate coverage. Initial materials can refine definitions, independent materials assess application, and records explain resolution of disagreement. Passenger conflicts, speech errors, and service resumption may span multiple levels; forcing one label may discard process information.

Transfer from simulation should examine relevant conditions rather than realism alone. Cue recognition may initially require comprehensible content and signals; action timing requires accurate task rhythms and control feedback; long-term reliance requires repeated use and actual limitations. Different questions require different fidelity. Each stage should explain which extrapolation limit it addresses while retaining comparable primary outcomes.

6.3 Connecting Evidence And Decisions

Validation should explain conflicts rather than only rank composite scores. Fewer touches may accompany longer voice responses; greater satisfaction may accompany poorer state understanding. Task analysis should identify acceptable gains and boundary-crossing costs. Safety-related performance should not be simply offset by entertainment satisfaction in a weighted score.

After testing, event-level assumptions should change, not only appearance. Revisions should identify reductions in information, changed timing, substituted modalities, or restricted initiative and retain unresolved failures. Scope should expand only when benefits persist in new conditions; otherwise conclusions remain local. The framework is thus revisable by evidence rather than requiring every result to fit it.

7 Discussion

7.1 Contributions To Automotive Intelligent Design

The principal contribution is the operational connection between a task's changing validity and the set of admissible service actions. Scenario descriptions [14] already include activities; multiple resource theory [5] already explains interference; mixed initiative [60] already addresses attention, uncertainty, and deferral. The framework adds a cockpit-specific event representation and decision sequence that make their dependencies inspectable, including which constraint changed an otherwise plausible intervention into deferral or cancellation.

A second relationship links the chosen action and recovery outcome back to a revisable task assumption. This operationalizes co-evolution [3] through an inspectable record rather than asserting a new general design theory. The baseline comparison, worked records, and propositions distinguish this proposed increment from inherited concepts. If an existing integrated method produces the same traceable decisions and outcomes with less effort, the proposed method's practical advantage would be weakened.

The predictions are therefore conditional on the event taxonomy and comparator: fewer unnecessary interruptions with necessary cues preserved (P1), a resource-demand interaction after content controls (P2), better discrimination and recovery with bounded costs (P3), and more valid failures found in unseen scenarios (P4). Each can fail independently; support for one cannot establish the whole framework or road-safety effectiveness.

7.2 Design Value And Evaluation Boundaries

Human-centred cockpit design cannot mean satisfying immediate preferences alone. Passenger needs conflict, and convenience may undermine understanding and recovery. Informed choice and control must therefore be considered together. Interfaces should support accurate capability understanding rather than imply unlimited competence through anthropomorphism or continual proactive service.

During design, AI expands candidates, reveals assumptions, and reduces some repetitive work; during use, it supports particular tasks. These values require separate evidence. General generative-work and collaboration studies illuminate conditional benefits but do not establish cockpit effectiveness. Design research and driving human factors remain the main grounds for domain-specific judgements.

Long-term use can alter tasks. Explanations become repetitive with familiarity, while updates may invalidate prior understanding. Evaluation should include learning, stable use, and capability change. Repeated sessions with comparable core scenarios can examine whether detection, reliance, and recovery change together. Personalization needs inspectable and reversible settings so users can distinguish learned behaviour from temporary states.

Accessibility and capability differences are integral to task analysis. Hearing, touch, language, and motor differences affect cue accessibility and learning. Means should not conceal persistent failures among particular participants. Teams

should define target scope, record alternative-channel requirements and failures, then consider configurable solutions. Personalization should address demonstrated differences rather than automatically change critical cues from inferred labels.

Implementation costs also require testing. Excessively detailed scenarios may be unusable for teams, while changing orchestration rules may increase learning burdens despite reducing individual interruptions. Projects can begin with navigation, state cues, and service recovery, comparing problem discovery and revisions before and after framework use. Expansion is justified when added analysis supports verifiable decisions.

7.3 Limitations And Future Research

This study is an integrative conceptual synthesis without exhaustive searching, independent double coding, or empirical validation. Incomplete documentation of the original retrieval process limits reproducibility; the revision's focused source check cannot remove selection bias or certify full coverage. The scenario and event records demonstrate internal logic only. Their classifications, authorization checks, queue rules, and critical-transition interfaces require independent coding and vehicle-specific testing. Classic driving studies and current conversational cockpits differ in devices, tasks, and social practices, so no universal thresholds or safety improvements are inferred.

Attention estimation is unstable, and capability, language, culture, and bodily conditions alter interaction costs. Future research should test category reliability, low-confidence strategies, adaptation, and passenger negotiation, with alternatives for differing capabilities. Practical value requires connected evidence across prototypes, controlled experiments, and studies closer to actual use.

8 Conclusion

The scenario–task–attention–initiative framework is a conceptual model operationalized as an analytical method for cockpit information and service design. Its proposed contribution is the explicit connection from task validity and responsibility, through resource competition and admissible initiative, to event-based revision of design assumptions. The theoretical comparison, qualitative procedure, and worked cards make that connection inspectable. Four propositions specify how its incremental value could be tested against stated comparators. No experiment establishes its performance or safety benefits. Original-search documentation, classification reliability, transition safeguards, recovery feasibility, and transfer beyond constructed scenarios remain unresolved validation boundaries.

Supplementary Information

Appendices A–C provide the focused verification log, evidence and theory comparisons, and operational definitions within this manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors.

Author Contributions

Guanghong Xie is the sole author and was responsible for the study design, literature synthesis, and writing and revision of the manuscript.

Conflict of Interest

The author declares no conflict of interest.

Data Availability

No participant-level dataset was generated. The conceptual scenario cards, decision procedure, evidence map, and source-check record are included in the manuscript and appendices.

References

[1] Li, Y., Qi, L., Hao, R.: Current status and development of user-, technology-, and scenario-driven automotive intelligent design methods (In Chinese). Art & Design Research **2024**(6), 117–123 (2024)

[2] Figoli, F.A., Bruggeman, R., Rampino, L., Ciuccarelli, P.: *AI-against-design* map: A systematic review of 20 years of AI-focused studies in design research. Design Studies **96**, 101279 (2025). https://doi.org/10.1016/j.destud.2024.101279

[3] Dorst, K., Cross, N.: Creativity in the design process: Co-evolution of problem–solution. Design Studies **22**(5), 425–437 (2001). https://doi.org/10.1016/S0142-694X(01)00009-6

[4] Dorst, K.: The core of 'design thinking' and its application. Design Studies **32**(6), 521–532 (2011). https://doi.org/10.1016/j.destud.2011.07.006

[5] Wickens, C.D.: Multiple resources and mental workload. Human Factors: The Journal of the Human Factors and Ergonomics Society **50**(3), 449–455 (2008). https://doi.org/10.1518/001872008X288394

[6] Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Human Factors: The Journal of the Human Factors and Ergonomics Society **37**(1), 32–64 (1995). https://doi.org/10.1518/001872095779049543

[7] Bainbridge, L.: Ironies of automation. Automatica **19**(6), 775–779 (1983). https://doi.org/10.1016/0005-1098(83)90046-8

[8] Torraco, R.J.: Writing integrative literature reviews: Guidelines and examples. Human Resource Development Review **4**(3), 356–367 (2005). https://doi.org/10.

1177/1534484305278283

[9] Snyder, H.: Literature review as a research methodology: An overview and guidelines. Journal of Business Research **104**, 333–339 (2019). https://doi.org/10.1016/j.jbusres.2019.07.039

[10] Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., *et al.*: The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ **372**, n71 (2021). https://doi.org/10.1136/bmj.n71

[11] Wu, Z., Ji, D., Yu, K., Zeng, X., Wu, D., Shidujaman, M.: AI creativity and the human-AI co-creation model. In proceedings of Human-Computer Interaction. Theory, Methods and Tools, pp. 171–190 (2021). https://doi.org/10.1007/978-3-030-78462-1_13

[12] Luo, S., Zhang, D., Liang, C.: Collective-intelligence innovation design in the age of artificial intelligence [In Chinese]. Art & Design Research **2024**(2), 91–97 (2024).

[13] Kyriakidis, M., de Winter, J.C.F., Stanton, N., Bellet, T., van Arem, B., Brookhuis, K., Martens, M.H., Bengler, K., Andersson, J., Merat, N., *et al.*: A human factors perspective on automated driving. Theoretical Issues in Ergonomics Science **20**(3), 223–249 (2019). https://doi.org/10.1080/1463922x.2017.1293187

[14] Carroll, J.M.: Five reasons for scenario-based design. Interacting with Computers **13**(1), 43–60 (2000). https://doi.org/10.1016/S0953-5438(00)00023-0

[15] Cao, J., Zhang, J., Zhang, L., Wang, X.: Psychological structure and effects of interaction naturalness [In Chinese]. Acta Psychologica Sinica **55**(1), 55–65 (2023). https://doi.org/10.3724/SP.J.1041.2023.00055

[16] Healey, J.A., Picard, R.W.: Detecting stress during real-world driving tasks using physiological sensors. IEEE Transactions on Intelligent Transportation Systems **6**(2), 156–166 (2005). https://doi.org/10.1109/TITS.2005.848368

[17] Shahverdy, M., Fathy, M., Berangi, R., Sabokrou, M.: Driver behavior detection and classification using deep convolutional neural networks. Expert Systems with Applications **149**, 113240 (2020). https://doi.org/10.1016/j.eswa.2020.113240

[18] Oulasvirta, A., Kurvinen, E., Kankainen, T.: Understanding contexts by being there: Case studies in bodystorming. Personal and Ubiquitous Computing **7**, 125–134 (2003). https://doi.org/10.1007/s00779-003-0238-7

[19] Pink, S.: Doing Sensory Ethnography, 2nd edn. SAGE, London, UK (2015). https://doi.org/10.4135/9781473917057

[20] Sanders, E.B.-N., Stappers, P.J.: Probes, toolkits and prototypes: Three approaches to making in codesigning. CoDesign **10**(1), 5–14 (2014). https://doi.org/10.1080/15710882.2014.888183

[21] Regan, M.A., Hallett, C., Gordon, C.P.: Driver distraction and driver inattention: Definition, relationship and taxonomy. Accident Analysis & Prevention **43**(5), 1771–1781 (2011). https://doi.org/10.1016/j.aap.2011.04.008

[22] Klauer, S.G., Guo, F., Simons-Morton, B.G., Ouimet, M.C., Lee, S.E., Dingus, T.A.: Distracted driving and risk of road crashes among novice and experienced drivers. New England Journal of Medicine **370**(1), 54–59 (2014). https://doi.org/10.1056/nejmsa1204142

[23] Dingus, T.A., Guo, F., Lee, S., Antin, J.F., Perez, M., Buchanan-King, M., Hankey, J.: Driver crash risk factors and prevalence evaluation using naturalistic driving data. Proceedings of the National Academy of Sciences **113**(10), 2636–2641 (2016). https://doi.org/10.1073/pnas.1513271113

[24] Schnelle-Walka, D., McGee, D.R., Pfleging, B.: Multimodal interaction in automotive applications. Journal on Multimodal User Interfaces **13**(2), 53–54 (2019). https://doi.org/10.1007/s12193-019-00295-x

[25] Oviatt, S.: Ten myths of multimodal interaction. Communications of the ACM **42**(11), 74–81 (1999). https://doi.org/10.1145/319382.319398

[26] Strayer, D.L., Johnston, W.A.: Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. Psychological Science **12**(6), 462–466 (2001). https://doi.org/10.1111/1467-9280.00386

[27] Horrey, W.J., Wickens, C.D.: Examining the impact of cell phone conversations on driving using meta-analytic techniques. Human Factors: The Journal of the Human Factors and Ergonomics Society **48**(1), 196–205 (2006). https://doi.org/10.1518/001872006776412135

[28] Caird, J.K., Willness, C.R., Steel, P., Scialfa, C.: A meta-analysis of the effects of cell phones on driver performance. Accident Analysis & Prevention **40**(4), 1282–1293 (2008). https://doi.org/10.1016/j.aap.2008.01.009

[29] Caird, J.K., Johnston, K.A., Willness, C.R., Asbridge, M., Steel, P.: A meta-analysis of the effects of texting on driving. Accident Analysis & Prevention **71**, 311–318 (2014). https://doi.org/10.1016/j.aap.2014.06.005

[30] Caird, J.K., Simmons, S.M., Wiley, K., Johnston, K.A., Horrey, W.J.: Does talking on a cell phone, with a passenger, or dialing affect driving performance? An updated systematic review and meta-analysis of experimental studies. Human Factors: The Journal of the

Human Factors and Ergonomics Society **60**(1), 101–133 (2018). https://doi.org/10.1177/0018720817748145

[31] Strayer, D.L., Drews, F.A., Johnston, W.A.: Cell phone-induced failures of visual attention during simulated driving. Journal of Experimental Psychology: Applied **9**(1), 23–32 (2003). https://doi.org/10.1037/1076-898x.9.1.23

[32] Lee, J.D., Caven, B., Haake, S., Brown, T.L.: Speech-based interaction with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. Human Factors: The Journal of the Human Factors and Ergonomics Society **43**(4), 631–640 (2001). https://doi.org/10.1518/001872001775870340

[33] Spence, C., Ho, C.: Multisensory warning signals for event perception and safe driving. Theoretical Issues in Ergonomics Science **9**(6), 523–554 (2008). https://doi.org/10.1080/14639220701816765

[34] Ho, C., Reed, N., Spence, C.: Multisensory in-car warning signals for collision avoidance. Human Factors: The Journal of the Human Factors and Ergonomics Society **49**(6), 1107–1114 (2007). https://doi.org/10.1518/001872007x249965

[35] Ho, C., Santangelo, V., Spence, C.: Multisensory warning signals: When spatial correspondence matters. Experimental Brain Research **195**(2), 261–272 (2009). https://doi.org/10.1007/s00221-009-1778-5

[36] Parasuraman, R., Sheridan, T.B., Wickens, C.D.: 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 (2000). https://doi.org/10.1109/3468.844354

[37] Parasuraman, R., Riley, V.: Humans and automation: Use, misuse, disuse, abuse. Human Factors: The Journal of the Human Factors and Ergonomics Society **39**(2), 230–253 (1997). https://doi.org/10.1518/001872097778543886

[38] Lee, J.D., See, K.A.: Trust in automation: Designing for appropriate reliance. Human Factors: The Journal of the Human Factors and Ergonomics Society **46**(1), 50–80 (2004). https://doi.org/10.1518/hfes.46.1.50_30392

[39] Hancock, P.A., Billings, D.R., Schaefer, K.E., Chen, J.Y.C., de Visser, E.J., Parasuraman, R.: A meta-analysis of factors affecting trust in human-robot interaction. Human Factors: The Journal of the Human Factors and Ergonomics Society **53**(5), 517–527 (2011). https://doi.org/10.1177/0018720811417254

[40] Endsley, M.R., Kiris, E.O.: The out-of-the-loop performance problem and level of control in automation. Human Factors: The Journal of the Human Factors and Ergonomics Society **37**(2), 381–394 (1995). https://doi.org/10.1518/001872095779064555

[41] Sarter, N.B., Woods, D.D.: How in the world did we ever get into that mode? Mode error and awareness in supervisory control. Human Factors: The Journal of the Human Factors and Ergonomics Society **37**(1), 5–19 (1995). https://doi.org/10.1518/001872095779049516

[42] Hoff, K.A., Bashir, M.: Trust in automation: Integrating empirical evidence on factors that influence trust. Human Factors: The Journal of the Human Factors and Ergonomics Society **57**(3), 407–434 (2015). https://doi.org/10.1177/0018720814547570

[43] Schaefer, K.E., Chen, J.Y.C., Szalma, J.L., Hancock, P.A.: A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. Human Factors: The Journal of the Human Factors and Ergonomics Society **58**(3), 377–400 (2016). https://doi.org/10.1177/0018720816634228

[44] Körber, M., Baseler, E., Bengler, K.: Introduction matters: Manipulating trust in automation and reliance in automated driving. Applied Ergonomics **66**, 18–31 (2018). https://doi.org/10.1016/j.apergo.2017.07.006

[45] Verberne, F.M.F., Ham, J., Midden, C.J.H.: Trust in smart systems: Sharing driving goals and giving information to increase trustworthiness and acceptability of smart systems in cars. Human Factors: The Journal of the Human Factors and Ergonomics Society **54**(5), 799–810 (2012). https://doi.org/10.1177/0018720812443825

[46] Parasuraman, R., Manzey, D.H.: Complacency and bias in human use of automation: An attentional integration. Human Factors: The Journal of the Human Factors and Ergonomics Society **52**(3), 381–410 (2010). https://doi.org/10.1177/0018720810376055

[47] Noy, S., Zhang, W.: Experimental evidence on the productivity effects of generative artificial intelligence. Science **381**(6654), 187–192 (2023). https://doi.org/10.1126/science.adh2586

[48] Vaccaro, M., Almaatouq, A., Malone, T.: When combinations of humans and AI are useful: A systematic review and meta-analysis. Nature Human Behaviour **8**, 2293–2303 (2024). https://doi.org/10.1038/s41562-024-02024-1

[49] Zhang, G., Raina, A., Cagan, J., McComb, C.: A cautionary tale about the impact of AI on human design teams. Design Studies **72**, 100990 (2021). https://doi.org/10.1016/j.destud.2021.100990

[50] Doshi, A.R., Hauser, O.P.: Generative AI enhances individual creativity but reduces the collective diversity of novel content. Science Advances **10**(28), eadn5290 (2024). https://doi.org/10.1126/sciadv.adn5290

[51] Messeri, L., Crockett, M.J.: Artificial intelligence and illusions of understanding in scientific research. Nature **627**, 49–58 (2024). https://doi.org/10.1038/s41586-024-07146-0

[52] Buchanan, R.: Wicked problems in design thinking. Design Issues **8**(2), 5–21 (1992). https://doi.org/10.2307/1511637

[53] Simon, H.A.: The Sciences of the Artificial, 3rd edn. MIT Press, Cambridge, MA, USA (1996).

[54] Cross, N.: Designerly Ways of Knowing, 1st edn. Springer London, London, UK (2006). https://doi.org/10.1007/1-84628-301-9

[55] Clark, A.: Supersizing the Mind: Embodiment, Action, and Cognitive Extension. Oxford University Press, New York, USA (2008). https://doi.org/10.1093/acprof:oso/9780195333213.001.0001

[56] Paxion, J., Galy, E., Berthelon, C.: Mental workload and driving. Frontiers in Psychology **5**, 1344 (2014). https://doi.org/10.3389/fpsyg.2014.01344

[57] Warm, J.S., Parasuraman, R., Matthews, G.: Vigilance requires hard mental work and is stressful. Human Factors: The Journal of the Human Factors and Ergonomics Society **50**(3), 433–441 (2008). https://doi.org/10.1518/001872008x312152

[58] Charles, R.L., Nixon, J.: Measuring mental workload using physiological measures: A systematic review. Applied Ergonomics **74**, 221–232 (2019). https://doi.org/10.1016/j.apergo.2018.08.028

[59] Tao, D., Tan, H., Wang, H., Zhang, X., Qu, X., Zhang, T.: A systematic review of physiological measures of mental workload. International Journal of Environmental Research and Public Health **16**(15), 2716 (2019). https://doi.org/10.3390/ijerph16152716

[60] Horvitz, E.: Principles of mixed-initiative user interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 159–166, (1999). https://doi.org/10.1145/302979.303030

[61] Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P.N., Inkpen, K., *et al.*: Guidelines for human-AI interaction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–13, (2019). https://doi.org/10.1145/3290605.3300233

[62] Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press, Cambridge, MA, USA (2016). https://www.deeplearningbook.org/

[63] Russell, S.J., Norvig, P.: Artificial Intelligence: A Modern Approach, 3rd edn. Pearson Education Limited, London, UK (2010).

[64] Beller, J., Heesen, M., Vollrath, M.: Improving the driver–automation interaction: An approach using automation uncertainty. Human Factors: The Journal of the Human Factors and Ergonomics Society **55**(6), 1130–1141 (2013). https://doi.org/10.1177/0018720813482327

[65] Körber, M., Prasch, L., Bengler, K.: Why do I have to drive now? Post hoc explanations of takeover requests. Human Factors: The Journal of the Human Factors and Ergonomics Society **60**(3), 305–323 (2018). https://doi.org/10.1177/0018720817747730

[66] Norman, D.A.: The Design of Everyday Things, revised and expanded edition, Basic Books, New York, USA (2013).

[67] Molich, R., Nielsen, J.: Improving a human-computer dialogue. Communications of the ACM **33**(3), 338–348 (1990). https://doi.org/10.1145/77481.77486

[68] Calvo, R.A., Peters, D.: Positive Computing: Technology for Wellbeing and Human Potential. MIT Press, Cambridge, MA, USA (2014). https://doi.org/10.7551/mitpress/9764.001.0001

[69] Naujoks, F., Wiedemann, K., Schömig, N., Hergeth, S., Keinath, A.: Towards guidelines and verification methods for automated vehicle HMIs. Transportation Research Part F: Traffic Psychology and Behaviour **60**, 121–136 (2019). https://doi.org/10.1016/j.trf.2018.10.012

[70] Eriksson, A., Stanton, N.A.: Takeover time in highly automated vehicles: Noncritical transitions to and from manual control. Human Factors: The Journal of the Human Factors and Ergonomics Society **59**(4), 689–705 (2017). https://doi.org/10.1177/0018720816685832

[71] Zhang, B., de Winter, J., Varotto, S., Happee, R., Martens, M.: Determinants of take-over time from automated driving: A meta-analysis of 129 studies. Transportation Research Part F: Traffic Psychology and Behaviour **64**, 285–307 (2019). https://doi.org/10.1016/j.trf.2019.04.020

[72] Merat, N., Jamson, A.H., Lai, F.C.H., Daly, M., Carsten, O.M.J.: Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. Transportation Research Part F: Traffic Psychology and Behaviour **27**(Part B), 274–282 (2014). https://doi.org/10.1016/j.trf.2014.09.005

[73] Gold, C., Damböck, D., Lorenz, L., Bengler, K.: "Take over!" How long does it take to get the driver back into the loop? Proceedings of the Human Factors and Ergonomics Society Annual Meeting **57**(1), 1938–1942 (2013). https://doi.org/10.1177/1541931213571433

[74] de Winter, J.C.F., Happee, R., Martens, M.H., Stanton, N.A.: Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. Transportation Research Part F: Traffic Psychology and Behaviour **27**(Part B), 196–217 (2014). https://doi.org/10.1016/j.trf.2014.06.016

[75] Lu, Z., Happee, R., Cabrall, C.D., Kyriakidis, M., de Winter, J.C.: Human factors of transitions in automated driving: A general framework and literature survey. Transportation Research Part F: Traffic Psychology and Behaviour **43**, 183–198 (2016). https://doi.org/10.1016/j.trf.2016.10.007

[76] McDonald, A.D., Alambeigi, H., Engström, J., Markkula, G., Vogelpohl, T., Dunne, J., Yuma, N.: Toward computational simulations of behavior during automated driving takeovers: A review of the empirical and modeling literatures. Human Factors: The Journal of the Human Factors and Ergonomics Society **61**(4), 642–688 (2019). https://doi.org/10.1177/0018720819829572

[77] Gold, C., Körber, M., Lechner, D., Bengler, K.: Taking over control from highly automated vehicles in complex traffic situations: The role of traffic density. Human Factors: The Journal of the Human Factors and Ergonomics Society **58**(4), 642–652 (2016). https://doi.org/10.1177/0018720816634226

[78] Braun, V., Clarke, V.: Using thematic analysis in psychology. Qualitative Research in Psychology **3**(2), 77–101 (2006). https://doi.org/10.1191/1478088706qp063oa

[79] Hart, S.G., Staveland, L.E.: "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research" in Human Mental Workload, volume 52 of Advances in Psychology,pp. 139–183. North-Holland (1988). https://doi.org/10.1016/S0166-4115(08)62386-9

[80] Lakens, D.: Sample size justification. Collabra: Psychology **8**(1), 33267 (2022). https://doi.org/10.1525/collabra.33267

[81] Lakens, D.: Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. Frontiers in Psychology **4**, 863 (2013). https://doi.org/10.3389/fpsyg.2013.00863

[82] Landis, J.R., Koch, G.G.: The measurement of observer agreement for categorical data. Biometrics **33**(1), 159–174 (1977). https://doi.org/10.2307/2529310

Appendix

A Source Verification And Evidence Mapping

The revision check on 15 October 2026 was a targeted verification of eight existing references, not a new database review. The queries below document web retrieval directed to primary publisher or author-institution records. There was no publication-date filter. The reference corpus itself spans 1977–2025. The dates reported here refer only to this revision check; they do not reconstruct the original retrieval dates or screening history.

[14] Query: site.sciencedirect.com Carroll 2000 Five reasons scenario based design. Source inspected: ScienceDirect publication page and abstract.

[5] Query: site.sagepub.com Wickens 2008 Multiple resources mental workload 50 3. Source inspected: SAGE publication page and abstract.

[60] Query: site.microsoft.com research Horvitz Principles mixed initiative user interfaces 1999. Source inspected: Microsoft Research publication page and paper.

[3] Query: site.sciencedirect.com Dorst Cross 2001 Creativity design process co evolution problem solution. Source inspected: ScienceDirect publication page and abstract.

[32] Query: site.journals.sagepub.com Speech-Based Interaction In-Vehicle Computers Caven Haake Brown 2001. Source inspected: Author-institution paper record returned by web search; University of Iowa.

[64] Query: site.journals.sagepub.com Improving driver automation interaction uncertainty Beller Heesen Vollrath 2013. Source inspected: SAGE publication page and abstract.

[71] Query: site.sciencedirect.com Determinants takeover time automated driving meta-analysis 129 studies 2019. Source inspected: ScienceDirect publication page, abstract and methods excerpts.

[61] Query: site.microsoft.com research guidelines human AI interaction Amershi 2019. Source inspected: Microsoft Research publication materials and paper.

The mapping below separates empirical studies (E), research syntheses or theory reviews (R), established conceptual or design-method sources (T), and extensions proposed in this paper (X). A source can have both empirical and conceptual roles. The table is an analytical audit trail, not a ranking of evidence quality or a claim that all source findings transfer to current cockpits.

Table A1: Evidence roles and proposed extensions

Proposition and sources	Supported source-level claim	Transfer limit	Proposed extension X
P1: [14] T; [5] R/T; [32] E	Situated activities structure design; concurrent tasks can interfere; speech interaction can impose driving demands	No source directly tests this event record or the scenario-label comparator	Task validity and responsibility condition triggering; test interruption with omission constraints
P2: [5] R/T; [26, 32, 34, 35] E; [27, 30, 33] R	Interference depends on resource and task structure; warning arrangements and concurrent speech have conditional effects	Results concern specific tasks and signals; not all assistants, channels or timing policies	Compare resource matching and permitted delay while controlling semantics, length and familiarity
P3: [60] T; [38, 42, 43] R; [44, 64, 65] E; [61] design guidelines	Initiative involves uncertain goals and attention; communication and error experience affect reliance	Trust ratings and local simulator findings do not validate a complete authorization-and-recovery policy	Join event-specific authorization, consequence classification and recovery; separate four decision outcomes
P4: [3] E/T; [4, 20] T; [49] E; [48] R	Problem and solution framing evolve; human–AI outcomes depend on task and collaboration	No evidence here establishes that these cockpit event records improve design teams	Connect event failure to a revised assumption and test valid failure detection in unseen scenarios
Boundary across P1–P3: [69] guidance; [71, 74–76] R; [72, 73, 77] E	Takeover performance depends on transition and task conditions; response onset is not the whole recovery process	Published response times cannot set a universal service-release or safety threshold	Protect externally specified necessary cues; revalidate services after a completed transition

B Comparison With Theoretical Foundations

Table B2: Approaches, units of analysis, core variables, and scope/validation relationships

Approach and unit of analysis	Core variables and decision outputs	Scope and validation relationship
Scenario-based design [14]. Unit: situated activity and use scenario	Actors, goals, activities and consequences; scenarios and design alternatives	Supports multiple perspectives and design reflection. Inherited here: activity context. Added: an explicit event-validity and authorization record. Test against a specified implementation, not the entire tradition
Multiple resource theory [5]. Unit: concurrent task demands	Processing resources and overlap; predictions of interference and possible reallocations	Assesses overload under defined task conditions. Inherited: resource competition. Added: content validity and permissible delay before modality choice. P2 tests the planned resource-demand interaction
Mixed-initiative interaction [60]. Unit: user–system service interaction	Uncertain goals, attention, action value and human control; initiate, defer or seek input	Already includes attention-sensitive deferral. Inherited: uncertainty and initiative principles. Added: cockpit-specific task windows, external cue protection and recoverability checks. P3 tests the support package
Problem–solution co-evolution [3]. Unit: design episode and evolving frames	Problem and solution representations; reframing and revised alternatives	Explains reciprocal development in design. Inherited: revisable problem framing. Added: event-to-assumption links with held-out failure tests. P4 evaluates teams rather than vehicle safety
Proposed framework. Unit: cockpit event, plus team-level design revision	Task validity, responsibility, resource competition, authorization and recovery; admissible action, queue state and revised assumption	Conceptual analytical method within external safety constraints. Increment lies in inspectable dependency between fields. P1–P4 test separate predictions; no superiority has been demonstrated

C Operational Definitions

Table C3: Terminology, definitions, and observable records

Term	Definition	Observable record
Scenario	Bounded description of actors, road and vehicle conditions, responsibility, and event progression	Scenario ID; state and transition fields; known and unknown conditions
Task	Goal-directed activity with a responsible actor, success criterion, validity window and failure consequence	Task card and time-stamped state changes
Attention	Task-related allocation and competition of perceptual, cognitive and action resources; not a directly sensed scalar	Concurrent tasks, gaze, event detection, response and workload, interpreted separately
Initiative	Who may initiate, propose, execute, stop or reverse a service interaction	Initiator, action type, authorization and actual execution state
Interruption	Optional output requiring a response or associated with a switch from the ongoing task	Prompt onset, response demand, observed switch and task return; necessity coded independently
Recovery	Restoration of an acceptable task or service state after interruption or error	Restored state, correctness, time, actions, lost context and residual consequences
Appropriate reliance	Acceptance or rejection matching independently specified recommendation validity and scope	Correct/incorrect acceptance and rejection; validity-specific denominators; nonresponse separate
Non-intervention	Reasoned absence of optional proactive output or execution while necessary cues remain protected	Decision reason; defer versus cancel; review or expiry condition; omitted necessary cues as failure
Orchestration	Joint selection of content, modality, timing and permitted initiator across concurrent tasks	Candidate set, decision rule, selected action, alternatives and recovery state