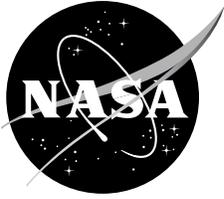# Human Factors Research Needs for NextGen-Airportal Safety

Thomas B. Sheridan
U.S. Department of Transportation, John A. Volpe National Transportation Systems Center
Cambridge, Massachusetts

The NASA STI Program Office . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the Lead Center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA's counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

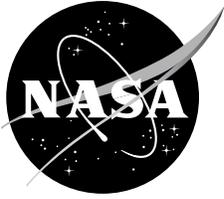- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results . . . even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at *http://www.sti.nasa.gov*

- E-mail your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA Access Help Desk at (301) 621-0134

- Telephone the NASA Access Help Desk at (301) 621-0390

- Write to:
  NASA Access Help Desk
  NASA Center for AeroSpace Information
  7115 Standard Drive
  Hanover, MD 21076-1320

Human Factors Research Needs for NextGen-Airportal Safety

Thomas B. Sheridan
U.S. Department of Transportation, John A. Volpe National Transportation Systems Center
Cambridge, Massachusetts

Available from:

TABLE OF CONTENTS

TABLE OF CONTENTS (cont.)

LIST OF FIGURES

LIST OF TABLES

HUMAN FACTORS RESEARCH NEEDS FOR NEXTGEN-AIRPORTAL SAFETY

Thomas B. Sheridan[1]

Ames Research Center

1. INTRODUCTION

The purpose of this document is to recommend human factors research needs for NextGen airportal safety.

The document first summarizes, from the perspective of human factors engineering, how NextGen is expected to differ from the current National Airspace System (NAS). It then lists some caveats regarding known human operator performance limitations that will affect airportal safety and accordingly imply necessary research application. Finally, it discusses human factors research needs specific to major stages of airport operations (surface and terminal airspace).

2. ASSUMPTIONS OF WHAT KEY FEATURES ARE LIKELY TO BE IN NEXTGEN THAT WILL AFFECT HUMAN FACTORS RESEARCH NEEDS

Total air traffic will steadily increase, possibly doubling by 2025, posing significant workload increase on the air navigation service providers (ANSPs). Major U.S. airports are already beyond capacity at peak hours and especially in poor weather.

A much greater mix of aircraft, including aircraft equipage (automatic dependent surveillance-broadcast (ADS-B), datalink, etc.) and aircraft type (including light jets, helicopters, unmanned aerial vehicles (UAVs), etc.) must share the airspace.

ADS-B technology will permit latitude/longitude surveillance that is far more accurate than radar, while altitude surveillance will improve relatively little. Closer separations will be required.

Four-dimensional (4D) trajectories will be negotiated well before flight time between airlines operations personnel, pilots, controllers and airport managers. These will be modified just before flight time and during the flight due to weather and other demands. Airlines will see tighter constraints on scheduling.

---

[1] U.S. Department of Transportation, John A. Volpe National Transportation Systems Center, 55 Broadway, Cambridge, MA 02142.

Digital data-link will become the major means of air-ground communication, including shared system-wide information (weather, flow constraints, airport surface information, clearances, etc.). Voice communications will be available for particular communication protocols as necessary.

The pilot will assume primary responsibility for self-separation, using improved displays of surrounding traffic, while at the same time adhering to assigned 4D trajectories.

Controllers will assume more flow management responsibilities and less vectoring of aircraft, intervening when needed. There is a need to establish the relative roles and responsibilities of pilots, controllers, airline operations centers (AOCs), and airport authorities in NextGen and how they will differ from current practices.

Many new decision aids will be available to both controllers and aircrew, including: conflict probes and alarms, and an enhanced traffic alert and avoidance system (TCAS); time-based metering for spacing and flow; descent and landing advisories; moving maps and advisories for taxiing and gate assignment; pushback scheduling, etc. There is a need to anticipate how these decision aids will function, simulate their function, and evaluate their usability.

Where commercial aviation safety has been very good compared to other transportation modes, experience has shown that system changes typically precipitate unpredictable human errors and system failures. Therefore policymakers and the public will be especially apprehensive and critical during transition to NextGen. Acceptance of changes will require extensive demonstration by human-in-the-loop (HITL) simulation and demonstration in actual operations.

3. CROSS-CUTTING HUMAN FACTORS KNOWLEDGE THAT NEEDS TO BE APPLIED TO NEXTGEN-AIRPORTAL CHALLENGES

Listed below are multiple caveats about human operator limitations that are well known to human factors engineers. These caveats all apply in one way or another to safety aspects of NextGen, and thus call for application and refinement of available research knowledge. In both this section and the next section, where particular research needs are detailed with respect to phases of flight (except en route), four broad topical categories are used: 1) topics related to inputs to humans (human acquisition of information); 2) topics related to human outputs (decision, response and mental workload); 3) topics related to human-automation interaction (interaction with and trust in decision aids and other automation); and 4) topics related to evaluating human errors and system simulation (performance evaluation).

3.1 Human Acquisition of Information

3.1.1 Shared Situation Awareness
System-wide information management (SWIM) is a major feature of NextGen, made available through modern high bandwidth digital communications between air and ground computers and personnel. One basis for analysis is to consider the network of such "intelligent agents" and the information flows between them (see appendix A1).

A major premise of NextGen is shared situation awareness (SA). SA is a relatively recent topic of human factors study, bringing together a variety of sub-disciplines, including: assumptions and expectations; attention allocation and distraction; signal detection; perception and evaluation of stimulus relevance; prospective memory; prediction capability; and boredom and fatigue. We need to understand what sharing of information implies in the context of various airportal operational tasks. The ideal of shared SA is especially challenging because different operators in the system have differing responsibilities, differing temporal demands, different forms of display, etc.

Humans are very limited in the number of elements of a situation they can focus on at any one time, though given enough time they are good at perceiving patterns in data that a machine, unless specifically programmed to do so, cannot detect.

Humans can easily be overloaded with information, so analysis of what information should be "pushed" and what should be "pulled" is a critical need.

The best measures of SA are based on debriefing either after a set of tasks is completed or by real-time interruption and query. What measures are most appropriate to simulations and real operations?

3.1.2 Monitoring
Airport Surface Detection Equipment, Model X (ASDE-X) surface radar as well as global positioning system (GPS) enables new displays of the airport surface, for example as so-called staffed virtual towers (Hannon et al., 2008). This capability, whether used for tower cab auxiliary display (e.g., in inclement weather) or in tower-less airports, needs to be carefully evaluated.

From a variety of evidence we know that humans are poor monitors, and after 30 minutes of nothing unusual occurring on what they are monitoring (and their own resultant inactivity) they tend to get inattentive, bored, and drowsy.

People tend to see what they expect to see. There is a well known "looked-but-did-not-see phenomenon" (e.g., in highway driving research). We need to know when this might occur, for example when the controller views displays of aircraft.

3.1.3 Distraction
What we call "distraction" can take several forms (Sheridan, 2006):
- Visual distraction, which can be
  (a) unexpected events external to the nominal task that visually and involuntarily "grab" attention, or
  (b) voluntary and intended (looking away from the nominal task to socialize, do side tasks, etc.).
- Cognitive distraction, which can be
  (a) involuntary emotional stress, voluntarily problem solving unrelated to the nominal task, or
  (b) voluntary intentional thinking about things unrelated to the nominal task.
We need to anticipate what forms of distraction will be manifest in airportal operations.

3.2   Decision, Response, and Mental Workload

3.2.1   Decision Making
Humans implicitly use objective functions (value or utility trade-offs) in making decisions and wherever possible tend to make decisions "automatically" based on learned behavior patterns that have worked in the past. They avoid weighing all the possibilities explicitly, as compared, e.g., to computerized Bayesian decision-making.

Such patterns of decision-making work well for routine decisions, but pose problems for abnormal situations.

Humans are poor at estimating absolute probabilities that are very small or very close to one. They are much better at estimating probability ratios. Probability estimation will arise, for example, when controllers judge time-slot availability for runway crossing during taxiing, or for injecting takeoffs into an arrival stream.

3.2.2   Mental Models
A "mental model" is a mental cause-consequence calculation that a human can "run," often in real time, for example by a pedestrian in judging oncoming vehicles in crossing the street (Moray, 1997). It is critical that mental models of different ANSP and aircrew engaged in joint operations correspond. This can be tested in simulations.

3.2.3   Response Time
When human controllers are initially "out of the loop" and are suddenly called upon to step in, understand the situation, and take over control, it can take a long time—a function of situation complexity.

Human response time for responding to abnormalities and taking corrective action tends to follow a log normal function, meaning there is a long tail on the probability density distribution and to achieve 95% confidence may mean a long period. Expectations of system designers regarding the properties of human response distributions need to be checked.

3.2.4   Mental Workload
The 2007 NGATS ATM Airportal Project Reference Document (Hinton et al., 2007) cited human workload as the "most critical factor needed for NGATS research."

Workload is problematic when the human operator either has too much to do in the available time (the usual concern), or too little, such that he or she gets bored and inattentive, but most critical is the transient from the latter to the former. This can be a serious problem when ANSP are suddenly and unexpectedly called to intervene and "come up to speed" in sorting out a complex or abnormal air traffic management (ATM) situation.

Human (mental) workload became an important factor in aviation starting with the DC9-80 conversion from a three- to a two-person flight crew in the mid 1970s. Early studies (Sheridan and Simpson, 1979) suggested a three-attribute subjective scaling: task-load (busy-ness), stress-load (emotion), and problem complexity load. This was followed by the development of the NASA widely used TLX rating scale (Hart and Staveland, 1988). These methods have been widely used by, e.g., Airbus, and should be applied in NextGen operations.

Physiological measures of mental workload, such as heart rate variability, eye pupil diameter, respiratory rate, etc., have proven unsuccessful due to large human variability and are not recommended.

The imposition of secondary tasks can be applied in human-in-the-loop (HITL) experimental simulations but are inappropriate distractors in real operations. How well the subject does on the secondary task measures "spare capacity" on the primary task, but such measures cannot be used in real aviation operations because they may compromise safety.

3.3 Interaction With and Trust in Decision Aids and Other Automation

3.3.1 Interaction With Decision Aids and Other Automation

Human supervision of automation means that humans must:

- understand what the machine can and cannot do in terms of capability,

- know how to give directives to the machine,

- know how to monitor what the machine is doing, i.e., whether it is following directives, what kind of feedback it is getting from outside, whether it is running into trouble,

- know what to do if intervention is necessary and be able to recover to a stable state, and

- be able to learn from what the machine has done, and work with it better next time.

Appendix A2 makes a distinction between human direct control of a process, human control of a decision aid, and human supervisory control of automation. It also tables the common "levels of automation" which are different options for the degree to which systems are autonomous. Finally, it offers taxonomies with respect to ways automation can fail and temporal aspects of failure. Reviews of the topic are in Sheridan and Parasuraman (2006) and Sheridan (2002).

It is generally believed that a human should always be able to take over control from automation. However, there are examples where that would be dangerous. Certain Airbus aircraft (e.g., Airbus A320) will not allow the aircraft to respond to control actuations that put it into stall. In another area, control rods on nuclear power plants will drop into the core under certain circumstances and the operator is unable to override that action. So a major question for NextGen is when should the human be able to override the automation, and when should the automation be able to override the human.

A decision aid is more accepted by the human if it displays some "etiquette," i.e., is clear in communicating information, does not demand too much from the human, is respectful of the human's response delays and his errors in asking for or giving information, etc. (Miller, 2004). While it is too early for designing etiquette into airportal displays, specification of display requirements will point to needs for display interaction etiquette.

3.3.2 Trust in Automation

Decision aids are judged to be very useful and therefore trustworthy if they reliably work as advertised and reduce operator workload. However a highly reliable decision aid can produce complacency such that its recommendation is followed without question and where incipient failure is not detected. Experience with existing decision aids should be studied to determine the correct balance of too-much-trust versus too-little-trust to build into training programs.

3.3.3 Automation Failure Recovery

Automation failure includes not only the sudden catastrophic failure of automation hardware or software but also failures attributable to poor design that mislead the human operator. The latter are much harder to detect and therefore more insidious.

In judging likelihood of automation failure, humans are prone to probability judgment biases referred to earlier. In judging causality of automation failures, humans have a tendency to fixate on one cause and look for confirming evidence rather than hold multiple hypotheses in mind while evaluating evidence. Computers, on the other hand, are systematic and accurate, but only for hypotheses that have been programmed into them. Means should be sought by which human operators and computers can collaborate in detecting and evaluating incipient failure.

Automation failure recovery includes detection by pilot or controller, acquisition of enough information to decide on response, appropriateness of recovery response, and execution of recovery response. If response is too impetuous and without sufficient understanding of the situation, errors are likely. If too slow, serious consequences can occur before the system has recovered. Failure modes should be studied to reveal how quickly recovery must be made, or alternatively, how best to "buy time."

3.4 Performance Evaluation

3.4.1 Human Error

Human error has been a controversial topic since the 1970s regarding its definition, measurement, classification, and prediction. Best known treatises are Rasmussen (1982), Reason (1990), and Senders and Moray (1991). Recently a report by Reason et al., (2006) revisited the famous Swiss Cheese model (that accidents occur when multiple human, procedural, managerial, hardware and software defenses have "holes" that "line up") and questions whether this model has been overemphasized, with the quip "Is Swiss Cheese past its sell-by date?" (Reason, 2006).

The criterion of what constitutes an "error" is arbitrarily a function of what is regarded as an acceptable response. Many safety researchers prefer to talk about human action variability rather than human error. For some operations the margin of what is acceptable should be set generously to allow for human variability that will still not cause disruptions.

Errors can be classified by various taxonomies, which are useful to infer causation: slip vs. mistake; body function (sensory vs. memory vs. decision vs. motor); omission vs. commission; etc. Such error taxonomies can be gleaned from HITL experiments at early design stages and will suggest preventative measures.

Current thinking among safety researchers asserts that there should be less emphasis on counting "errors" and more on ensuring avenues of recovery once "errors" are made (since errors *will* be made and no two errors occur in exactly the same way). Thus, recovery techniques should be designed for all errors that can be anticipated.

There is a trend by safety engineers toward less emphasis on the "sharp end" of human-machine operations (real-time operations) and more on the "blunt end" (management practices: planning, allocation of resources, anticipation of latent weaknesses, policies that encourage individuals to be critical of current practices, sharing of assumptions, etc.). Overly relaxed supervision and not anticipating failures is regarded by some as the most critical problem (Hollnagel et al., 2006).

Various studies on safety in human-machine interaction suggest that safety is less a matter of the limits of human performance (human capabilities generalized from experimental studies) and more on individual human "behavior" (considerations individuals use in making decisions: what is important to them, what they remember, what they learned and how they were trained, what assumptions they make of their tasks, what they attend to, how they respond, how consistent they are, etc.). The latter factors are more difficult to research than the former. Unfortunately people do not normally operate at peak performance levels.

3.4.2   Simulation of Human-machine Interaction

HITL simulation has been called for repeatedly above. It is the means by which many NextGen ideas must be evaluated, given that for humans (mostly) one cannot do analysis by solving equations, as in applied physics. HITL must not wait until the validation/verification stage, but be employed at all stages of design and development, initially using preliminary and relative crude system simulations.

Fast-time simulation, where both automation and human are modeled in the computer, can and should be used when human models are available and the task is relatively simple. For example, good human models are already available for thresholds of vision and hearing, for biomechanics and anthropology of workspaces, for manual control of aircraft, and for pilot workload (Gore and Corker, 2000).

4.   HUMAN FACTORS RESEARCH NEEDS AT MAJOR STAGES OF TERMINAL AIR AND SURFACE OPERATIONS

Research needs are lumped under: 1) preflight planning and negotiation; 2) gate departure and taxi; 3) take-off and climb-out; 4) descent and landing; and 5) taxi to gate. Research needs are categorized under the same headings used above.

4.1   Preflight Planning and Negotiation

This includes all the interactions between AOCs, Federal Aviation Administration (FAA) flight managers, ANSP, and pilots and the various computer tools they use preflight, from 6 months to minutes before push back.

4.1.1   Human Acquisition of Information

Special displays/interfaces for negotiation must be developed so that all parties can quickly and easily determine the implications of the constraints and what options remain for the AOC (up to some point prior to departure), and to the flight crew as departure time nears. Final negotiations may occur due to weather, etc., after the flight is in progress.

Satisficing, as articulated by Simon (1969) is appropriate for a formal information sharing/display/ negotiating process (see appendix A3). Various proposals have been made by the Joint Planning and Development Office (JPDO) for 4D negotiation systems. Though trajectory negotiation has always been done by informal communication between pilots and controllers, formal 4D negotiation in flight is something new.

4.1.2   Decision, Response, and Mental Workload

There is not likely to be excessive workload until minutes before preflight when final 4D routes are loaded into the flight management system (FMS). Presumably an FAA computer will make available to the AOC (or flight crew) an "optimized" 4D trajectory with some range of options. As the departure deadline nears there will necessarily be less flexibility and less opportunity for negotiation, and hence greater workload. If the aircraft is not ready, or airportal traffic is otherwise constrained so that departure must be delayed significantly, there must be increasingly strict lines of decision flexibility and authority. The timelines for imposing such restrictions must be researched.

4.1.3   Interaction With and Trust in Decision Aids and Other Automation

Primary negotiation systems must be backed up by alternate means to effect the 4D routing and negotiation, especially critical as departure time nears. As a last resort it may be the computer or the ANSP specifying the beginning segment of a 4D flight plan to allow the aircraft to get off the ground, a refinement of the flight plan to be firmed up after the aircraft is en route with potential alternative routes having been evaluated.

4.1.4 Performance Evaluation

Presumably negotiation will be mediated by computer systems as well as voice communication. Insufficient or undisciplined verbal communication or poor display/interface design can lead to misassumptions and confusion. AOC, ANSP, and flight crew may find themselves beginning to operate with differing assumptions about what was agreed to unless the presentation of the 4D trajectory is clear. Negotiation procedures and displays need to be designed and tested for speed, clarity, and effectiveness under various constraining circumstances.

Hypothetical negotiation rules, initial display formats, and hypothetical negotiation scenarios need to be set up and tried out though HITL among salient participants to begin to define what this process might involve.

4.2 Gate Departure and Taxi

This includes communications between AOC, ramp controller, ground controller, flight crew, and ground vehicles. Presumably a surface management computer will anticipate and communicate the "optimized" taxi route instructions to parties concerned (at least the flight crew, ramp controller, and ground controller or their ANSP equivalents), and provide final clearance to push back. For equipped aircraft taxi instructions may be sent to the FMS and displayed to the flight crew on an airport map with a moving "bug" to follow. Eventually further automatic control of taxing may be imposed. Unequipped aircraft will be given verbal taxi instructions in the usual manner by ANSP voice communication but based on the "optimized" taxi algorithms. Proposed procedures and displays need to be evaluated by airportal HITL.

4.2.1 Human Acquisition of Information

Ground controllers should have horizontal situation displays showing all aircraft with ID tags, runways and taxi routes, clearances, wake information, gate availability, etc., and most of this information should be available as needed to local controllers and AOC.

4.2.2 Decision, Response, and Mental Workload

It is expected that taxiing will occur at higher speeds and be continuously directed, possibly even with a two-dimensional (2D)-plus-time trajectory. Ground ANSP should observe an ASDE-X–like display of all taxi instructions and how well aircraft are following those instructions, so that all discrepancies can be spotted and verbal communication initiated. From time to time aircraft will wander from the planned taxi trajectory, and automation should alarm these discrepancies. Display requirements need to be specified.

Aircraft granted clearance for takeoff will be shown on the same display so that crossing active runways by other aircraft will accordingly be blocked (e.g., by a color change) and/or alarms sounded if that aircraft approaches too close to the cross threshold. Flight crew of equipped aircraft will have to monitor their (head-down) flight deck surface displays as well as maintain vigilance outside (head-up) chores probably shared between the two crewmembers—to maintain separation from other aircraft and stay on taxiway centerlines—tasks made more difficult by the higher required taxi speeds.

For unequipped aircraft permissions and prohibitions (holds) for active runway crossing can be provided by radio as auditory alerts to the flight crew. Decision aids should also indicate wake vortex separation requirements. Unequipped aircraft, though they will not have the decision aids to assist in taxi routing and speed maintenance, will nevertheless have to keep up with the other (higher speed taxiing).

During taxi the flight crew may access and communicate with ANSP regarding takeoff clearance (equipped aircraft via datalink, unequipped by voice). Keeping track of which aircraft are data-link equipped and which are not could prove to be a major chore for ground control as well as locals control (assuming that distinction remains), and the coordination between the two will become greater because of increased crossing of active runways and tighter timing to execute those crossings safely. HITL evaluation of such systems is critical.

In cases of bad weather, and especially when taxi to a deicing pad is required, the taxi planning will have to be modified, and the workload and getting back "on schedule" will be exacerbated.

4.3   Interaction With and Trust in Decision Aids and Other Automation

Computer optimization and decision aids may occasionally malfunction such that taxi traffic may have to be slowed down, especially as visibility conditions deteriorate. Because of the density of traffic (in a super-density airport at peak times) a sudden failure of data link or of the taxi routing software would probably force a temporary freeze on departure taxi operations, while arrival taxi operations would have to continue as first priority, with departure taxiing handled manually as best possible. This would also cause a perturbation in the 4D en route scheduling.

4.3.1   Performance Evaluation

Even though display and warning aids will help provide taxi guidance, maneuvering taxiing aircraft at higher speeds may still lead to missed taxiways or runway crossing violations. Any one aircraft deviation may upset the taxi optimization for some (hopefully brief) time period, and the system will have to re-optimize. Unequipped aircraft whose identity cannot be confirmed (even though ASDE-X surface radar has them located) may have to be inserted into a taxi optimization manually by ANSP.

Receiving clearances while taxiing may cause memory errors for pilots of unequipped aircraft (since no unambiguous clearance signal would be registered on an available display).

Taxi algorithms, taxi displays, and taxi procedures need to be tested in HITL simulations. Dynamic taxi control simulations should test ability to taxi at higher speeds, and the efficacy evaluated for moving map (or moving "bug") displays in the flight deck. There may be limits to how fast certain aircraft can turn and maneuver, especially in crosswinds.

Recovery procedures (workload and SA transitions) need also be simulated with multiple pseudo-pilots and both ground and local controller.

4.4 Take-Off and Climb-Out

4.4.1 Human Acquisition of Information
Wake vortices will be a serious constraint because of efforts to space arrivals and departures as closely as possible. Wake warnings should be available to aircrew and controllers if practical.

4.4.2 Decision, Response, and Mental Workload
For equipped aircraft the take-off clearance, contingent upon revocation, as well as clearance delivery and necessary frequencies, etc., may be preloaded into the FMS during taxi out. Non-equipped aircraft will have to be cleared by voice.

To save time, aircraft will probably move from the taxiway to the runway and immediately commence roll. Taxi-into-position and hold (TIPH) will necessarily be minimized.

Concurrent crossing of aircraft on the same runway or a crossing runway downfield will necessitate a high degree of coordination in timing, and thus takeoff rolls must be initiated promptly on schedule, with an eye out for need to abort takeoffs. If parallel runways are available one may be devoted to departures, a second to arrivals; alternatively arrivals and departures may occur in "bunches" to save time, but this will require a queue of aircraft waiting for departure. HITL will clarify what works best.

During climb-out there will probably be tighter tolerances on spacing relative to leading aircraft, sensitivity to wake vortex effects, and attention to ensuring conformance to agreed-upon and displayed 4D en route trajectory.

When and what acknowledgment procedures for clearances, and what forms of handoffs between ANSP personnel, have yet to be worked out. Such procedures should be more explicit than rather loose readback-hearback techniques now communicated over voice channels.

4.4.3 Interaction With and Trust in Decision Aids and Other Automation
Automation of takeoff and initial climb-out is not recommended because of need to keep pilot in the loop—until handoff to en route ANSP and climb to altitude and synchronization with 4D trajectory. Failure of automation at that point can be handled by manual control and fallback autopilot modes.

4.4.4 Performance Evaluation
Potential errors center around runway incursions due to miscommunication between pilot and controller or between local and ground control, or with clearance delivery, much as at present except that tolerances will be tighter and automation will be advising or guiding certain operations. (The assumption is that no airportal operations will include automatic control of the aircraft except possibly autopilot control to heading, altitude, and speed once takeoff has been achieved). HITL should be used extensively for this evaluation.

4.5 Descent and Landing

4.5.1 Human Acquisition of Information

Flight deck and ANSP should both have horizontal situation displays of terminal airborne traffic with separation boundaries and collision prediction probes, including an indication of runway assignments, clearances, anticipated taxi turn-offs, possibly even gate assignments and go-around instructions and paths as appropriate.

Efficacy of profile synthetic vision displays (virtual visual flight rules (VFR)) should be evaluated.

4.5.2 Decision, Response, and Mental Workload

This has always been the flight phase of highest crew workload and most likely will continue to be. The flight crew must transition from what is essentially automatic flight with relatively low workload monitoring activity to a relatively high workload: getting the weather and active runway information, anticipating the airport geometry and looking up maps on the electronic flight bag (EFB) as necessary, planning the descent trajectory in 4D according to datalinked advisories, watching for other traffic in the pattern, concern for wake vortices of leading aircraft, worrying about icing in particular weather conditions, getting clearance, looking for the runway, etc.

For ANSP there will be some handoff between the Terminal Control Facility (TRACON) and the local controller. Descent management tools such as Center Tracon Automation System (CTAS) or its successors will no doubt be used for: 1) time of arrival at a final approach fix; 2) descent advisor; and 3) final approach spacing tool. Metering will be time based. CTAS is already known to have increased controller workload and, with super-density workload at this phase, will continue to do so. If go-arounds are required, that will produce major perturbations in the system and increase both aircrew and ANSP workload further.

4.5.3 Interaction With and Trust in Decision Aids and Other Automation

Recovery from automation failure at this stage depends very much on the nature of the failure. If a CTAS-like system is being depended upon by the TRACON, if and when it (or some part of the infrastructure feeding it data) fails, there must be an immediate reversion to spacing aircraft manually. We assume that equipped aircraft will carry both ADS-B and radar transponders, though it is unclear whether full radar coverage will be fully maintained in the future. For transponder failure of individual aircraft, that aircraft can always report raw GPS position data to ANSP and there can be means to continually update position and maintain radio contact.

Primary phased-array surface radar should be evaluated as failure backup.

4.5.4 Performance Evaluation

When there are many things to do and think about, as in this stage of flight, any interruption may throw pilots "off track" and cause them to forget what they were planning to do on the checklist. HITL experiments might be used to evaluate unexpected interposition of unexpected events such as icing or other system failures.

With parallel runways there is the danger of lining up on the wrong runway (or a taxiway), especially in poor visibility, so ANSP must have proper displays by which to monitor glide-slope lateral position.

4.6 Taxi to Gate

4.6.1 Human Acquisition of Information
In small- and medium-size airports taxi routes will be standard for commercial aircraft, and either the ground controller will give taxi-to-gate instructions or give specific instructions that are well known and experienced by the flight crew.

Gate assignments and anticipated gate delays should be data-linked to pilots prior to landing.

4.6.2 Decision, Response, and Mental Workload
At some airports (e.g., Frankfort) most gates are used for any and all aircraft, independent of fleet operator. This idea needs policy evaluation.

Workload is as noted previously in 4.2 for departure taxi and higher taxi speeds.

4.6.3 Interaction With and Trust in Decision Aids and Other Automation
Same as 4.2.

4.6.4 Performance Evaluation
Tendency to turn off on wrong taxiway from runway, or take wrong taxi route, especially in bad weather, must be evaluated; when lost aircraft stops (awaiting further instructions) it may hold up other aircraft.

5. REFERENCES

Clarke, J. P. and Hansman, R. J.: Systems Analysis of Noise Abatement Approach Procedures Enabled by Advanced Flight Guidance Techniques. AIAA Paper no. 1997-0490, 35th AIAA Aerospace Sciences Meeting, Reno, Nev., Jan. 1997.

Charny, L. and Sheridan, T. B.: Adaptive Goal-Setting in Tasks with Multiple Criteria. Proc. 1989 IEEE International Conference on Cybernetics and Society, Cambridge, Mass., 1989.

Gore, B. F. and Corker, K. M.: Human Performance Modeling: Identification of Critical Variables for National Airspace Safety. Proc. 14th Triennial International Ergonomics Association (IEA) and the Human Factors and Ergonomics Society 44th Annual Meeting, Santa Monica, Calif., 2000.

Hannon, D.; Lee, J. T.; Geyer, M.; Mackey, S.; Sheridan, T.; Francis, M.; Woods, S.; and Malonson, M.: Feasibility Evaluation of a Staffed Virtual Tower, J. Air Traffic Control Association, in press, 2008.

Hart, S. G. and Staveland, L. E.: Development of the NASA TLX (Task Load Index): Results of Empirical and Theoretical Research. In Hancock, P. A. and Meshkati, N. (Eds.), Human Mental Workload, North Holland, 1988, pp. 239–250.

Hinton, D. et al.: FAA: NGATS ATM Airportal Project Reference Document, 2007.

Hollnagel, E.; Woods, D. D.; and Leveson, N.: Resilience Engineering: Concepts and Precepts. Ashgate Publishing, Burlington, Vt., 2006.

Miller, C. (Ed.): Human-Computer Etiquette [Special Issue], 2004. Communications of the ACM, vol. 47, no. 4, pp. 30–61.

Moray, N.: Models of Models of Mental Models. In Sheridan, T. and Van Lunteren, A. Eds., Perspectives on the Human Controller. Lawrence Erlbaum, Mahwah, N.J., 1997, pp. 271–285.

Rasmussen, J.: Human Errors: A Taxonomy for Describing Human Malfunction in Industrial Installations. J. Occupational Accidents, vol. 4, 1982, pp. 311–335.

Reason, J.: Human Error, Cambridge University Press, New York, N.Y., 1990.

Reason, J.; Hollnagel, E.; and Paries, J.: Revisiting the Swiss Cheese Model of Accidents. EEC Note 13/06. Eurocontrol, Brussels, Belgium, 2006.

Senders, J. W. and Moray, N. P.: Human Error: Cause, Prediction and Reduction. Lawrence Erlbaum, Mahwah, N.J., 1991.

Sheridan, T. B. and Simpson, R. W.: Toward the Definition and Measurement of the Mental Workload of Transport Pilots. DoT-MIT Program of University Research Report DOT-OS-70055, 1979.

Sheridan, T. B.: Humans and Automation. Wiley and Sons, New York, N.Y., 2002.

Sheridan, T. B.: Driver Distraction from a Control Theory Perspective. Human Factors, 2006.

Sheridan, T. B. and Parasuraman, R.: Human-Automation Interaction. Nickerson, R. (Ed.), Review of Human Factors and Ergonomics, vol. 1, ch. 2, 2006.

Simon, H.: Sciences of the Artificial. MIT Press, Cambridge, Mass., 1969.

6. APPENDICES

APPENDIX A-1. ANALYSIS OF COMMUNICATION FLOWS BETWEEN INTELLIGENT AGENTS

Analysis of the communication flows (frequency, contingencies, and relative importance) between the various human and computer "intelligent agents" should be performed, comparing NextGen plans with a baseline of current procedures. Figure 1 illustrates such a network.

a) Controller-pilot verbal communication
b) Controller-pilot/aircraft data-link communication
c) ATM-based automatic control of aircraft
d) Controller interaction with ATM decision aid
e) Decision aid interaction with master ground computer
f) Controller direct interaction with master ATM computer
g) Pilot interaction with aircraft decision aid
h) Aircraft decision aid interaction with autopilot
i) Autopilot control of aircraft
j) Pilot direct interaction with autopilot
k) Pilot manual control of aircraft
l) Other controller to OAC communication
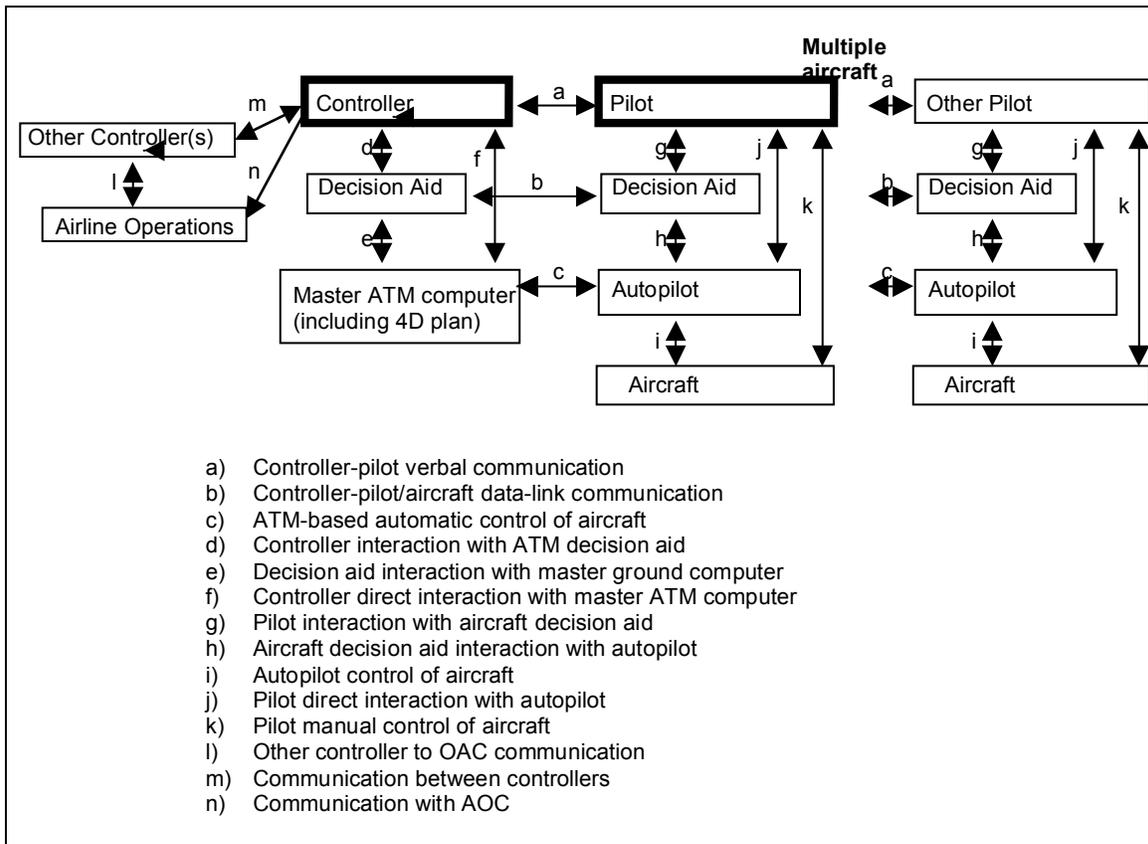m) Communication between controllers
n) Communication with AOC

Figure 1. Network of intelligent agents that communicate with one another.

APPENDIX A-2. TYPES, LEVELS, AND FAILURE MODES OF CONTROL

Types of Control

Figure 2 illustrates the distinctions between direct manual control, control of a decision support tool that allows "what would happen if…" simulation, and human supervisory control of automation. Note that control of the decision support system (middle diagram) by itself has no effect on the actual system, but only provides a basis to revert to manual control (top) or supervisory programming of an automatic controller (bottom) to make use of what was learned.
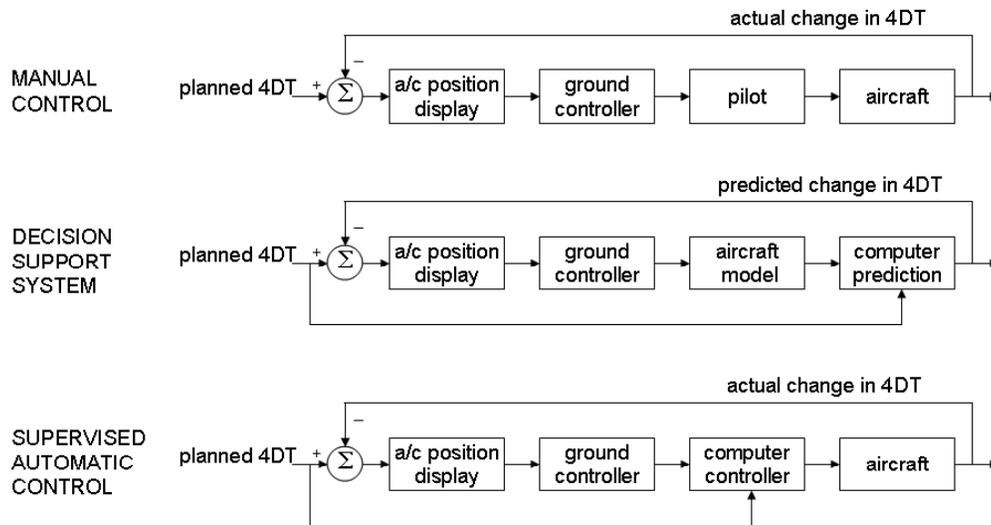

Figure 2. Distinctions between manual control, control of a decision-support tool, and supervised automation in air traffic control.

Levels of Control

Table 1 lists eight "levels" of automation (from none to "full") commonly discussed as a basis for system design (Sheridan 2002).

TABLE 1. A SCALE OF LEVELS OF AUTOMATION THAT MUST BE DECIDED UPON
(independent for each of the four functions)

1	The computer offers no assistance; the human must do it all.
2	The computer suggests alternative ways to do the task.
3	The computer selects one way to do the task and
4	executes that suggestion if the human approves, or
5	allows the human a restricted time to veto before automatic execution, or
6	executes the suggestion automatically, then necessarily informs the human, or
7	executes the suggestion automatically, then informs the human only if asked.
8	The computer selects the method, executes the task, and ignores the human.

Modes of Automation Failure

Automation can fail in various ways. Below are 1) a taxonomy with regard to causal circumstances, and 2) a classification with regard to temporal aspects of automation failure.

Causal Circumstances of Automation Failure

- Hardware equipment failure (prevention by design, redundancy)
- Software equipment failure (prevention by testing under variety of input circumstances)
- Apparent failure because user not operating it correctly (prevention by training, interaction with designer to elaborate contingencies (e.g., use of fault trees))
- Apparent failure because user expected it to provide proper answer/action under circumstances it was not designed for (prevention by HITL and mental model testing)

Time Characteristics of Failure and Recovery

- Sudden – What alarms? What means to buy time? Where time is not available, what automatic backup action? Effect of time delay in human response?
- Gradual – What warnings? What displays to diagnose level of competency of remaining function? What means to consult relevant information (diagnose failure causality, call up playbook of recovery options, relevant people to communicate with)? What recovery information stored where in SWIM? What possibility to restore before complete failure? How long will such information access take?

APPENDIX A3. A NOTE ON A SATISFICING APPROACH TO 4D TRAJECTORY NEGOTIATION

Satisficing is a term generally taken to mean "To accept a choice or judgment as one that is good enough, one that satisfies, in consideration of the mental effort, time or other resources necessary to further refine one's choice." Formal satisficing techiques have been used in system design, control, and many other technical arenas. It can be contrasted with optimization, finding the unique maximum of some explicit objective function (the latter being beyond attainability in most real-world situations). The formal idea of satisficing is usually attributed to economist and artificial intelligence pioneer Prof. Herbert Simon of Carnegie-Mellon University (Simon, 1969).

The Next Generation Air Transportation System (NextGen) calls for participating (sufficiently equipped) aircraft to follow four-dimensional (three in space, one in time), or "4D" trajectories. These trajectories would be negotiated with the FAA by airline operations offices well before day of flight (from months to days?), possibly modified one or more times prior to actual flight time, and possibly modified again during flight due to weather constraints, equipment failures, medical emergencies, etc.

The FAA Operational Evolution Partnership ("Smart Sheet" documentation) stipulates that "The overall philosophy driving the delivery of Cooperative Air Traffic Management services in the NextGen is to accommodate the flight operator (airline) "to the maximum extent possible, and to impose restrictions only when a real operational need exists to meet capacity, safety, security or

environmental constraints. If constraints are required, the goal is to maximize the operator's opportunities to resolve them based on their own preferences."

This philosophy sets the goal of any system designed to allow 4D trajectory negotiation to take place. Presumably it would start by trying to accommodate all the requested trajectories that are allowable by preset criteria of corridor restrictions based on trip origin and destination, contractual relations with airport authorities, expected loading and airspeed for type of aircraft, etc. This negotiation would most likely involve interaction between AOCs and other airline personnel with FAA and airport personnel to set nominal regular schedules (for scheduled airlines) much as they are set today, including "horse trading" of schedules, and possibly including monetary side payments. This would be a continual process, with schedule updates occurring periodically as occur now. At this stage, although computer simulations would be used, little or no formal satisficing algorithms might be used (though any negotiation among humans will necessarily involve informal satisficing).

As is well known, since deregulation airlines tend to schedule to their own advantage, typically scheduling on the basis of VFR conditions at peak travel demand periods. This is often frustrating to controllers who realize full well that VFR schedules cannot be maintained in poor weather and are forced to impose ground delays and other means to maintain safety and make traffic control tractable. The traveling public is caught in the middle.

Having a nominal schedule in place (presumably weeks ahead of any subject flight), formal satisficing can begin to occur. Inherent in that nominal schedule are the constraints the computer must begin work with in order to execute a satisficing operation. At this point, as each new request comes in from an operator the computer can engage personnel representing the operator in a satisficing negotiation to fulfill the NextGen philosophy described above.

How might a satisfying negotiation work? Any 4D request implies a series of 4D points in space and time. Let that number of points (dimensions) be N, hopefully a relatively small number. That implies establishing points in a 4N hyperspace to specify a given trajectory. The computer is capable of discovering any other one of the 4N points where there would be a conflict with an already scheduled aircraft. Note that for even a small N, and an expected number of aircraft during that time period and that general routing, it would be impossible for a human being to visualize this number of points in a hyperspace.

The computer can then point out the conflict to the human AOC, pilot, or other person negotiating, and inform him/her of what range of options there might be for coming close to the 4N trajectory points desired. On the computer's part, this is more than a static task for the aircraft is constrained by its own aerodynamic and thrust capabilities as to how it can move from allowable point to allowable point. Thus, in the computer's hyperspace (memory) there are what can be called *constant constraints* (trajectory points for aircraft already scheduled) and *relational constraints* (curves defining relations between points for the new aircraft having to do with the capability to transition between points).

While an example of a satisficing negotiation in a many-dimensional hyperspace cannot be represented easily in this brief paper, it can be done in a simple 2D space (figure 3). In this case only departure time and arrival time are represented. There is a constant departure time constraint (Td = departure time of some other aircraft, plus separation time) represented by a vertical line, a constant arrival time constraint (Ta = arrival of before some other aircraft minus separation time) represented by a horizontal line, and a single relational constraint (diagonal line) representing the speed capability of the aircraft (Ta = Td + travel time). These constraints bound what can be called an *attainability space* (set of points that are allowable, i.e., meet the established constant and relational constraints and the aircraft separation criteria).

Suppose the negotiator's initial request is for point 1. The computer returns a denial, indicating that departure must occur after Td. With this information the human negotiator might tell the computer, "Okay, I can't be at point 1, but how about at point 2?" It can be seen that this satisfies the departure constraint, but not the diagonal relational constraint, so the negotiator is denied again by the computer. At this point the computer can add, "I see in which direction you are backing off, so here is how close you can get, namely point 3, just inside the attainability space." But, noting that point 3 requires high thrust and low fuel efficiency, the negotiator might well say, "Better not waste fuel, so how about arriving just a little later, point 4," to which the computer now says, "Okay, you've got it." In this simple example the negotiator could easily interpret the graph from the outset, observe the attainability space, and go directly to point 4. But the point is that in a hyperspace of many dimensions it would be impossible for the human to visualize the problem—nevertheless the required comparisons and responses would be no problem for the computer (note that "how close" would also take into account standard separation requirements). This negotiation may take several back-and-forth interactions between negotiator and computer.
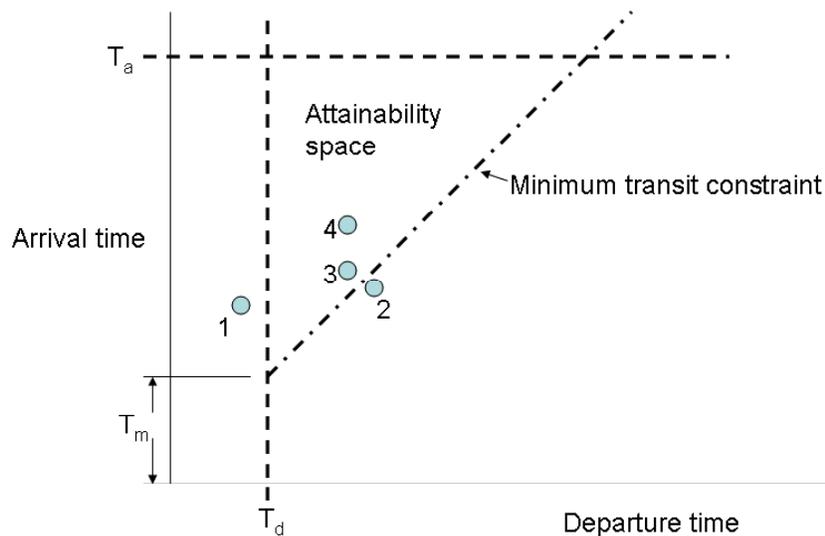

Figure 3. Simple example of satisficing interation with computer to negotiate trajectory.

There are examples in the literature of such negotiation software including user-friendly interface; see Charny and Sheridan (1989), who experimented with satisficing in a five dimensional (5D) negotiation space. Satisficing has been proposed in some specific ATM realms, in particular to trade-off environmental noise standards against schedule and fuel requirements of operators (Clarke and Hansman, 1997).

The problem for NextGen, in this writer's opinion, lies largely with the interface design, and especially the form of a display that allows a negotiator to easily observe the constant constraints, test the effects of choices along each dimension as to the effects in other dimensions (due to relational constraints), receive "help" from the computer in finding an attainable set of points, and understand the implications of a final attainable trajectory choice.

NextGen planning stipulates that "flight operators be accommodated to the maximum extent possible, and that restrictions be imposed only when a real operational need exists to meet capacity, safety, security, or environmental constraints." New surveillance and computational technology should enable this to be accomplished to a greater degree than now, even given increased traffic demand. However, means must be found to perform negotiations between appropriate human parties prior to and during flight, as a function of weather and other unexpected events. Decision support tools based on "satisficing" are an attractive approach to implementing this objective, however development of such tools demands significant research in human-computer interaction and integration, together with depth of knowledge in flight operations.

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) 21/07/2009	2. REPORT TYPE Contractor Report		3. DATES COVERED (From - To)
4. TITLE AND SUBTITLE Human Factors Research Needs for NextGen-Airportal Safety		5a. CONTRACT NUMBER	
		5b. GRANT NUMBER	
		5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Thomas B. Sheridan		5d. PROJECT NUMBER	
		5e. TASK NUMBER	
		5f. WORK UNIT NUMBER WBS 305295.02.07.01.10	

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Department of Transportation John A. Volpe National Transportation Systems Center Cambridge, MA 02142	8. PERFORMING ORGANIZATION REPORT NUMBER A-090009
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, D.C. 20546-0001	10. SPONSORING/MONITOR'S ACRONYM(S) NASA
	11. SPONSORING/MONITORING REPORT NUMBER NASA/CR–2009-215372

12. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified — Unlimited   Distribution: Nonstandard Subject Category: 03 Availability: NASA CASI (301) 621-0390

13. SUPPLEMENTARY NOTES
Point of Contact: Mike D. Madson, Ames Research Center, MS 257-1, Moffett Field, CA 94035-1000 (650) 604-3621

14. ABSTRACT
The purpose of this document is to recommend human factors research needs for NextGen airportal safety. Based upon assumptions of what key features are likely to be in NextGen that will affect human factors research needs, including increased capacity, greater mix of aircraft, new technologies, pilot responsibility for self-separation, controllers assuming more flow management responsibilities, the document first summarizes, from the perspective of human factors engineering, how NextGen is expected to differ from the current National Airspace System. It then lists some caveats regarding known human operator performance limitations that will affect airportal safety and accordingly imply necessary research application. Finally, it discusses human factors research needs specific to major stages of airport operations (surface and terminal airspace). Many new decision aids will be available to both controllers and aircrew, including conflict probes and alarms and enhanced Traffic Collision Avoidance System time-based metering for spacing and flow; descent and landing advisories; moving maps and advisories for taxiing and gate assignment; pushback scheduling, etc. There is a need to anticipate how these decision aids will function, simulate their function, and evaluate their usability. Where commercial aviation safety has been very good compared to other transportation modes, experience has shown that system changes typically precipitate unpredictable human errors and system failures. Policymakers and the public will be especially apprehensive and critical during transition to NextGen. Acceptance of changes will require extensive demonstration by human-in-the-loop simulation and demonstration in actual operations.

15. SUBJECT TERMS
NextGen, Airport, Human Factors, Safety, Controller, Pilot, Decision aids

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Mike D. Madson
a. REPORT	b. ABSTRACT	c. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unclassified	28	19b. TELEPHONE (Include area code) (650) 604-3621

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18