2013

NextGen Operational Improvements: Will They Improve Human Performance

Bettina L. Beard

James C. Johnston

Jon Holbrook

NEXTGEN OPERATIONAL IMPROVEMENTS: WILL THEY IMPROVE HUMAN PERFORMANCE

Bettina L. Beard and James C. Johnston
NASA Ames Research Center
Moffett Field, CA, USA
Jon Holbrook
San Jose State University
Moffett Field, CA, USA

Modernization of the National Airspace System depends critically on the development of advanced technology, including cutting-edge automation, controller decision-support tools and integrated on-demand information. The Next Generation Air Transportation System national plan envisions air traffic control tower automation that proposes solutions for seven problems: 1) departure metering, 2) taxi routing, 3) taxi and runway scheduling, 4) departure runway assignments, 5) departure flow management, 6) integrated arrival and departure scheduling and 7) runway configuration management. Government, academia and industry are simultaneously pursuing the development of these capabilities. For each capability, the development process typically begins by assessing its potential benefits, and then progresses to designing preliminary versions of the tool, followed by testing the tool's strengths and weaknesses using computational modeling, human-in-the-loop simulation and/or field tests. We compiled research studies of the tools, assessed the methodological rigor of the studies and served as referee for partisan conclusions that were sometimes overly optimistic. Here we provide the results of this review.

The FAA's Next Generation Air Transportation System (NextGen) proposes to modernize the U.S air traffic system by deploying advanced technology with the aims of streamlining equipment, consolidating common operational tasks and facilitating human management of traffic operations. The FAA has identified specific capabilities, referred to as Operational Improvements (OIs) that will be gradually phased into operations. The OIs of interest here are those classified as Decision-Support Tools (DSTs) and procedures for airport tower personnel. They are:

- Departure Metering
- Taxi Routing & Scheduling
- Departure Runway Assignment
- Runway Scheduling
- Departure Flow Management
- Integrated Arrival/Departure, and
- Runway Configuration Management.

In 2010, the Government Accountability Office (GAO) recommended that the FAA "identify clear goals for the performance of these capabilities or … settle on a set of metrics for measuring their performance relative to any goals." In response, the FAA identified 5 metrics: capacity, efficiency, predictability, safety and environment (NextGen Implementation Plan, 2012). The FAA's Human Factors Research and Engineering Group would like for this set of metrics to also include measurements of human performance. NASA's Human Factors Research & Technology Division at Ames Research Center has entered into an Interagency Agreement with the FAA group to help define a human performance metric(s). We refer to the project as the Human Performance Budget.

NASA has approached the development of a Human Performance Budget in several ways. First, we asked a panel of human factors experts to rate whether the proposed OIs could have a positive or negative impact on 23 human performance metrics (Beard, Parke, Holbrook & Oyung, in preparation). The experts' ratings indicated that overall the OIs could, if implemented appropriately, positively influence team situation awareness and coordination. One of the most useful capabilities to offer the controllers would be automated support for conformance monitoring. According to the expert ratings the greatests risks were the potential for decay of controller skills and knowledge, potential difficulties with tools requring controllers to continuously monitor for the occurrence of a specified target or event and potential problems with meeting training standards using the new tools.

NASA then asked how thoroughly the OIs address current tower safety problems. We extracted and analyzed over 200 Aviation Safety Reporting System (ASRS) reports submitted over a 5-year period (Holbrook, Puentes, Stasio, Jobe, McDonnell & Beard, 2011). We found that the majority of the reports dealt with potential safety threats to runway operations, a concern that could potentially be addressed by OIs aimed at improving controller situation awareness during runway operations. Two issues extracted from the ASRS, oganizational climate and inadequate supervision, are not addressed by any of the current OIs.

NASA also asked the user community what they need. In a survey involving over 125 tower controllers, we asked to what extent the NextGen capabilities could help them in their job or to improve capacity, efficiency, flexibility, predictability and safety at their airport. Controllers indicated that departure metering from the ramp could be the most helpful tool but that enhanced information would be the most helpful enabler (Holbrook, Parke, Oyung, Collins, Gonter & Beard, 2013).

All three approaches to the Human Performance Budget point to the importance of keeping the controller in the loop and providing them with information to help them do their job. But it remains to be seen whether the advanced automation tools being developed will (1) improve controller and team situation awareness and coordination to aid in decision-making and (2) be mature enough for implementation by 2018.

Methods

Research articles that evaluated tower controllers' performance using NextGen systems were gathered and those articles judged to be of the greatest value were reviewed in considerable depth based on experimental design and relevance to NextGen. The goal was to evaluate whether the research and data analyses were experimentally rigorous and the conclusions valid. In a report to the FAA, Beard (2012) provides a more in-depth discussion of the literature review.

Research studies were classified based on their relevance to the seven capabilities listed in the Introduction (identified from NextGen OIs) and key enabling technologies required to realize the full potential of those capabilities. This report provides a brief description of the meta-analysis of Human-in-the-Loop (HITL) publications classified within the seven capabilities of NextGen. Research related to the key enabling technologies is evaluated elsewhere (Beard, Galeon & Parke, in preparation).

Results

Capability 1: Departure Metering at the Ramp

The need to reduce airport surface delays has been approached from different angles. Departure metering, prior to release from the gate or from the spot, is one approach that has considerable promise and has been implemented at several airports. The basic concept of departure metering is very simple: aircraft destined for the same runway are provided either gate release times or taxi times from the spot. This provides a way to sequence departing aircraft without the many drawbacks of a long physical queues (which impose costly time delay and fuel consumption penalties). This idea has been computationally shown to be advantageous both operationally and environmentally (e.g., Brinton & Lent, 2012; Simaiakis, Sandberg, Balakrishnan and Hansman, 2012).

The Surface Management System (SMS) is a system developed jointly by the FAA and NASA. An early field test focused on controller judgement (Atkins, Brinton, Walton, Arkind, Moertl & Carniol, 2003). Six years later, the tool was still found to provide unreliable pushback predictions (Monroe, 2009) because it did not, at the time, incorporate airport surveillance data.

At NASA Ames, an airport surface testbed based on SMS is being used to test the efficacy of tools for taxi and runway scheduling (Jung et al., 2010). With a predetermined sequence and reliable estimate of taxi times, each flight can be assigned pushback times calculated to allow unimpeded taxiing to the runway, followed by an immediate clearance for takeoff. Even though this concept has been tested in high fidelity simulation twice, it is still an immature concept. Results indicate that the concept does not appropriately take into account the variability in system operation, including variability in actual taxi times. In both HITL simulations, the controller was required to implement the "advice" of the automation. The initial and second HITLs involved two and six controller participants, respectively. Both simulations manipulated traffic level, but did not include off-nominal events. Hoang, Jung, Holbrook and Malik (2011) found that the spot release-time recommendations decreased controller situation awareness. Verbal reports from the controllers indicated that attending to the automation interfered with their own planning. Underlying the observed problems was a mismatch between the goals of the automated tool and the goals of controllers. The effects on situation awareness of a more recent version of the advisor that includes electronic flight data are still being assessed (Hayashi, 2012).

A European tool, Airport Collaborative Decision Making, aims to reduce gate turnaround time by coordinating the actions of airlines, airports, air traffic control and pilots. Each aircraft has a predetermined sequence and knowledge of the taxi time of a given flight. Pushback times can be given to every flight so that it can travel unimpeded to the runway and immediately take off. Successful implementation would require adoption of the automation by a large number of airports, and on accurate data supplied by all parties.

In the U.S. there is also a research effort focused on the enhancement of continuous, real-time collaboration between ramp and air traffic controllers (Fernandes, Smith, Spencer, Wiley & Johnson, 2011). Note that this collaboration adds a task to the traffic management coordinator's (TMC) repertoire. The Collaborative Airport Traffic System (CATS) is the current test-bed used to examine this concept, which has been successfully used in operations at JFK airport.

Capability 2: Taxi Routing and Scheduling

Several laboratories in the U.S. and Europe are testing systems that provide automatic generation of taxi routes (Cheng and Foyle, 2002; Stelzer; Morgan, McGarry, Klein and Kerns, 2011; Simaiakis et al., 2011) based on route data from all aircraft in the system, and data from real-time surface surveillance.

Testing of the Ground-operation Situation Awareness and Flow Efficiency tool (Go-SAFE) system (Verma et al., 2010) continues to elicit functional and interface problems (e.g. sub-optimal screen size and resolution, data-tag clutter issues, inappropriate color usage). Automated generation and delivery of clearances failed to lower controller workload, in part because of a mismatch between the clearance plans generated by automation and by controllers. Automation tended to generate unusual ad hoc taxi-route assignments while controllers preferred consistent familiar assignments. This is an example of a much more general clash between ad hoc flexibility, typically championed by operations research algorithms, and simplicity and consistency, typically championed by human controllers.

The Go-SAFE software did not allow a sufficiently flexible partnership between the controller and automation. As events unfolded, the automated agent often issued revised routes to the aircraft without controller agreement, and provided inadequate notification of the changes to controllers. Also the software did not permit the controller to modify clearances already issued, or to issue conditional clearances.

Mitre CAASD also has an airport surface automation testbed (Klein, Stelzer, Nelson, Brinton and Lent, 2010). HITL simulation results showed that controllers trained to have the same goals as the software always chose to follow the taxi routes suggested by software. The reluctance of controllers to issue route modifications was originally due to an unfriendly keyboard interface, but in a later study that added a map-based method for altering routes, controllers remained reluctant to amend automated taxi routes.

Capability 3: Departure Runway Assignment

No published HITL simulations were found for the Departure Runway Assignment capability, although preliminary discussions of the concept have been reported (Morgan, 2010).

Subject Matter Expert (SME) knowledge elicitation performed by NASA revealed that traffic management coordinators may consider up to 26 pieces of information to make a departure runway assignment. Because it is difficult for any human to effectively consider all these factors, particularly under periods of high workload, there is a clear need for automation designers to provide decision support tools. The OIs describe tools that suggest optimal runway assignments to the controller who can then accept or modify the assignment. However, automation designers must guard against the temptation to overweight factors that might help the operations-research goal of increasing capacity while underweighting the concerns of controllers. Where possible automation should be built to collect and integrate information that controllers would assemble manually. It is also important that there be a clear delineation between factors that automation has and has not factored into its recommendations so controllers know when omitted factors justify modifying automated advice.

Runway balancing is a function performed by the TMC, although individual runway assignments can be changed by controllers based on the immediate tactical situation. Atkins and Walton (2002) studied how well departure runways were currently balanced. They reported that traffic managers do not currently have accurate information about the future departure demand, nor the ability to predict how the surface situation will evolve, both of which are needed for effective TMC and controller decision-making.

Capability 4: Runway Scheduling

Jung et al. (2010) investigated an automated runway scheduler used in conjunction with the spot-release planner mentioned earlier. Together they have the potential for impressive reductions in the number and duration of stops in the queue. Surprisingly, however, the expected improvements in human performance (workload and situation awareness) did not materialize in practice (Hoang et al., 2011). Local controller workload and situation awareness were unchanged whether the runway scheduler was present or absent. However, controllers reported that they were mentally performing the runway scheduling task even when the advisor was present.

Capability 5: Departure Flow Management

The goal of the Departure Flow Management (DFM) tool is to use automation to improve the present manual process for releasing takeoffs, which requires a tower controller to make a phone call to Air Route Traffic Control Center (ARTCC). In the field trial, the tool was used in shadow-mode (Spencer, Carniol, Pepper & Smith, 2009). Controllers reported benefitting from two features of the tool: a record of what was done, and the seamless integration of new actions into the current operational picture. Several interface issues were identified: font size was too small to read the screen while standing, and an auditory cue was needed to signal when new information had arrived. A possible advantage of the DFM tool is that it provides the means to automatically communicate traffic management restrictions to the Traffic Flow Data Manager (TFDM). This kind of automatic sharing of information across tools and across installations is an important potential benefit from improved NextGen tools.

Doble, Timmerman, Carniol, Klopfenstein, Tanino & Sud (2009) reported the results of a field trial of a DFM tool. Tower controllers judged the tool to be useful, easy to use, and provided good access to needed information. Controllers also reported that the tool actually opened up more of their time for managing other issues, an important “figure of merit” that is rarely achieved.

So far, the Integrated Departure Route Planning (IDRP) tool has only been assessed via Subject Matter Expert surveys (Masalonis et al., 2008) therefore no conclusions will be drawn here.

Capability 6: Integrated Arrival/Departure Scheduling

There were no published HITL simulations found for this capability.

Arrival/Departure Management Tool (A/DMT) is actually a set of tools that will be a critical part of the tower controller’s workstation. A/DMT will integrate information from surveillance (stored in an easily accessible database) and other DSTs to characterize arrival/departure demand and surface and airspace constraints. The vision is that it will integrate traffic flow constraints provided by DFM.

Capability 7: Runway Configuration Management

There were no published HITL simulations found for this capability.

The Runway Configuration Management (RCM) problem is to determine which runways should be used for arrivals or for departures. NASA Langley is developing System Oriented Runway Management (SORM) tools. SORM is a composite of two subsystems: Runway Configuration Management (RCM) and the Combined Arrival/Departure Runway Scheduling (CADRS) which assigns flights to runways in real time, accomplishing goals such as runway balancing (Lohr, Brown, Stough, Atkins, Eisenhawer & Long, 2011). A very small capacity increase was seen in computational simulations. In addition, it appears that even these small capacity gains are achievable with more frequent dynamic runway changes likely to increase complexity and controller workload. It is unclear whether this capability should be included in the FAA’s mid-term plans.

Discussion

A great deal of NextGen resources have been channeled toward the difficult job of algorithm development and the identification of the NAS benefits (e.g., increased capacity, increased efficiency, reduced environmental impact) that can be expected if the algorithms are deployed. The operational and human performance results and tool maturity show some successes, such as the departure flow management tool that automates the manual process of ATC tower phone calls to ARTCC to release take-offs. Unfortunately, there are numerous tools that require considerable progress before the NextGen vision can be realized. Four of the seven capabilities have reached a level of maturity where they have been tested with humans in the loop; Departure Metering, Taxi Routing and Scheduling, Runway Scheduling and Departure Flow Management.

For all capabilities current operations have been explored, the proposed concept has been detailed and its application discussed, parameters of interest have been identified, and operational benefits that are likely to be realized highlighted. Weaknesses in the capabilities development include:

- Algorithm heuristics were rarely developed from knowledge about the human user’s mental model,
- Algorithms often did not properly incorporate uncertainty,

- Initial validation using computational models of the operational system typically did not include sufficient (or sometimes any) representation of the human operator as a sub-system,
- High fidelity simulations were conducted with tools of inadequate maturity,
- HITL simulations neglected to collect both objective and subjective human performance measures,
- HITL simulations disregarded basic experimental design principles, and
- As the capabilities mature, testing continues in isolation of other capabilities.

It is important that individual tools be reliable. When an unreliable tool is introduced into tower operations, there is an increased likelihood that it will not be used as intended. Controllers may use the tool's recommendations during nominal low-workload conditions (where little help was needed anyway), but turn them off under off-nominal high workload conditions (precisely those conditions for which it was assumed that help was most needed), because the tool's recommendations for complex situations could not be trusted. Alternatively controllers may use a poorly designed tool for an unintended purpose (such as gaining access to raw information rather than an action recommendation), a purpose for which some other aid would have been more cost-effective. And of course, if using the tool is more trouble than it is worth, controllers may place it under the console and not use it at all.

With the simultaneous introduction of multiple tools anticipated for NextGen, it becomes even more critical to ensure that tools have reached a high readiness level. If unforeseen problems arise, it will be difficult to pinpoint which of the multiple new systems introduced is the source of the problem. Furthermore, joint use of multiple new tools is likely to produce emergent problems due to unforeseen and unstudied interactions among the tools. Of course there is a continuum in the degree to which new tools will interact in usage, but in general there is huge overlap in the information used by different tools, and in the operational impact of tools on traffic. It is implausible to make a "default assumption" that multiple new tools developed independently will play well together. The typical practice of developing each tool in isolation postpones the issue of properly integrating multiple tools into the future. But integration cannot be indefinitely postponed if tools are ever to reach operational status. Research into integrated suites of new tools has begun in the last few years. A further ramp-up of integration research is needed, including more HITLs testing integrated suites of multiple tools, useful for accomplishing multiple task goals. In the long run, integration can become a positive strength of tool-development research, providing benefits greater than the sum of individual tool benefits.

Acknowledgements

This work was funded under Interagency Agreement #DTFAWA-09-X-80020 between NASA and the FAA. The FAA Human Factors Division coordinated the research requirement and its principal representative acquired, funded, and technically coordinated execution of this research.

References

Atkins, S., Brinton, C., Walton, D., Arkind, K., Moertl, P. & Carniol, T. (2003). Results from the initial Surface Management System field tests. In *Fifth Eurocontrol/FAA Conference*, Budapest, Hungary.

Atkins, S. & Walton, D. (2002). Prediction and control of departure runway balancing at Dallas/Fort Worth airport. In *American Control Conference*, Anchorage, AK.

Beard, B. L. (2012). Human-in-the-loop research supporting mid-term NextGen air traffic control tower operational improvements. Downloadable at: *https://www.hf.faa.gov/hfportalnew/Search/SearchReport.aspx.*

Beard, B. L., Galeon, M. & Parke, B. (2013). Are NextGen enablers ready for prime time? Manuscript in preparation.

Beard, B. L., Parke, B., Holbrook, J., & Oyung, R. (2013). Survey of human factors experts on the potential human performance risks and benefits of NextGen capabilities. Manuscript in preparation.

Brinton, C. & Lent, S. (2012). Departure queue management in the presence of traffic management initiatives. Paper presented at the *Integrated Communications, Navigation and Surveillance Conference*, .

Cheng, V. H. L. & Foyle, D. C. (2002). Automation tools for enhancing ground-operation situation awareness and flow efficiency. *Proceedings of the AIAA Guidance, Navigation, and Control Conference, Paper AIAA 2002-4856*, Monterey, CA.

Doble, N. A., Timmerman, J., Carniol, T., Klopfenstein, M., Tanino, M., & Sud, V. (2009). Linking traffic management to the airport surface. *Proceedings of the Eighth USA/Europe Air Traffic Management Research and Development Seminar*, Napa, CA.

Fernandes, A. B., Smith, P. J., Spencer, A., Wiley, E., & Johnson, D. (2011). Collaborative Airport Traffic System (CATS) to evaluate design requirements for an airport surface departure management system. *Proceedings of* *the Human Factors and Ergonomics Society, 55*: 1827.

Government Accounting Office. (2010). NextGen Air Transportation System: FAA's Metrics Can Be Used to Report on Status of Individual Programs, but Not of Overall NextGen Implementation or Outcomes. GAO-10-629.

Hoang, T., Jung, Y., Holbrook, J. B., & Malik, W. A. (2011). Tower controllers' assessment of the spot and runway departure advisor (SARDA) concept. *Proceedings of the Ninth USA/Europe Air Traffic Management Research and Development Seminar*, Berlin, Germany.

Holbrook, J., Parke, B., Oyung, R., Collins, R., Gonter, K. & Beard, B. L. (2013). Perceived usefulness of planned NextGen tools by air traffic control tower controllers. *Proceedings of the 17th International Symposium on Aviation Psychology.* Dayton, OH: The Wright State University.

Holbrook, J., Puentes, A., Stasio, N., Jobe, K., McDonnell, L., & Beard, B. L. (2011). How thoroughly do proposed NextGen mid-term operational improvements address existing threats? *Proceedings of the 16th International Symposium on Aviation Psychology.* Dayton, OH: The Wright State University.

Hayashi, M. (2012). SARDA HITL preliminary human factors measures and analysis. Retrieved on February 27, 2012 from http://ntrs.nasa.gov/search.jsp?R=20120015369.

Jung, Y., Hoang, T., Montoya, J., Gupta, G., Malik, W., & Tobias, L. (2010). A concept and implementation of optimized operations of airport surface traffic. *Proceedings of the 10th AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*. AIAA, Fort Worth, Texas.

Klein, K. A., Stelzer, E. K., Nelson, E. T., Brinton, C. & Lent, S. (2010). Improving efficiency with surface trajectory-based operations and conformance monitoring. *Proceedings of the 29th Digital Avionics Systems Conference*, Salt Lake City, UT.

Lohr, G. W., Brown, S., Stough, H. P. I., Atkins, S., Eisenhawer, S. & Long, D. (2011). System oriented runway management: A research update. *Proceedings of the Ninth USA/Europe Air Traffic Management Research and Development Seminar*, Berlin, Germany.

Masalonis, A., Bateman, H. DeLaura, R., Song, L. Taber, C. & Wanke, C. (2008). Integrated departure route planning. *Proceedings of the 27th Digital Avionics System Conference,* St. Paul, MN.

Monroe, G. A. (2009). Surface Management System departure event analysis. NASA/TM-2009-215383.

Morgan, C. E. (2010). Mid-term Surface Trajectory-Based Operations (STBO) Concepts of Use: Departure Runway Assignment, MTR100269V4.

NextGen Implementation Plan (2012). Retrieved February 27, 2012 from *www.faa.gov/nextgen/implementation/plan/*.

Simaiakis, I., Khadilkar, H., Balakrishnan, H., Reynolds, T. G., Hansman, R. J., Reilly, B., & Urlass, S. (2011). Demonstration of reduced airport congestion through pushback rate control. *Proceedings of the Ninth USA/Europe Air Traffic Management Research and Development Seminar*, Berlin, Germany.

Simaiakis, I., Sandberg, M., Balakrishnan, H. & Hansman, R. J. (2012). Design, Testing and Evaluation of a Pushback Rate Control Strategy. *Proceedings of the Fifth International Conference on Research in Air Transportation*, Berkeley, CA.

Spencer, A., Carniol, T., Pepper, J. & Smith, P.J. (2009). Airport departure flow management (DFM): Findings from field trial testing. *Proceedings of the 15th International Symposium on Aviation Psychology*, Dayton OH.

Stelzer, E. K., Morgan, C. E., McGarry, K. A., Klein, K. A. & Kerns, K. (2011). Human-in-the-loop simulations of surface trajectory-based operations. *Proceedings of the Ninth USA/Europe Air Traffic Management Research and Development Seminar*, Berlin, Germany.

Verma, S., Kozon, T., Lozito, S., Martin, L., Ballinger, D., & Cheng, V. (2010). Human factors of precision taxiing under two levels of automation. *Air Traffic Control Quarterly, 18(2),* 113-141.