3rd USA/Europe Air Traffic Management R&D Seminar Napoli, June 2000

Analytical Identification of Airport and Airspace Capacity Constraints

William R. Voss
Federal Aviation Administration, AUA-1
800 Independence Avenue S.W.
Washington, DC 20590 USA

Jonathan Hoffman
Center for Advanced Aviation System Development
The MITRE Corporation, M/S W196
1820 Dolley Madison Blvd, McLean, VA 22102-3481 USA

ABSTRACT

Identification of airport and airspace capacity constraints is complicated by the presence of Air Traffic Flow Management programs that protect congested airspace from becoming a safety problem. There are two ways to proceed: one may use a simulation to reproduce the operational environment without the ATFM, or one may analytically deduce the presence of the capacity constraint from the evaluation of performance metrics based on operational data. This paper describes a method by which the latter has been accomplished, using multi-dimensional user-oriented performance metrics to guide airspace redesign efforts in the eastern United States. Throughput, delay, predictability, and flexibility metrics contribute to the analysis. The analytical method is shown to have advantages in resources required and geographical scope. The broader scope can also be used to guide the choice of simulation parameters, when the greater precision of simulation is required in the later stages of an airspace or airport redesign.

INTRODUCTION

Identification of airport congestion problems is well understood. As aircraft line up to use runways, standard queuing theory applies, to a good approximation. The idea of capacity is well defined, albeit dependent upon details of the arriving and departing traffic. Away from airports, the airspace also can become congested. In this case, though, the great volume of space (in the dynamical sense, meaning both position and velocity) through which aircraft can move, and the sensitive dependence on the controller's workload, makes a simple capacity calculation less applicable. When capacity calculation is problematic, identification of a congestion problem lacks a sound analytical basis. This paper demonstrates a theoretical framework for extending the idea of identifying congestion problems from airports to airspace.

Origin of Airspace Problems.

Traffic patterns change, so over time, airspace designs gradually become less efficient. Inefficient airspace design leads to congestion. But airspace congestion, in the sense of too many aircraft in too small a volume, is not permitted to exist for long -- it overworks controllers and threatens safety. To alleviate this sense of congestion, Air Traffic Flow Management (ATFM) was introduced. ATFM is generally successful at managing airspace congestion in its most literal sense. The result is the kind of airspace congestion problems that we see in today's air traffic control systems. An inefficient design is not cured by ATFM -- the congestion is turned into ground delays.

In addition, it has recently been recognized[1] that other inefficiencies than delay cause penalties to the users of the airspace. These are the secondary effects of congestion. Reducing ground delays requires correcting the airport or airspace problem that led to them. This provides the motive force for resolving airspace problems. Unfortunately, institutional memory of the cause of a particular sector regulation (Europe), restriction (USA) or routing (both) quickly fades, and is frequently not well communicated outside the original facility. Therefore, an analytical method for identifying airspace problems from their secondary effects is essential to guide any attempt to correct the problems with new procedures, sectorization, or technology. Thus the technical question that led to this research: Given an airspace where inefficiencies are suspected, how can the true problems be identified analytically?

BACKGROUND

The way that airspace studies are typically initiated and conducted shows the importance of an analytical process to identify airspace problems. When contemplating an airspace study, one is confronted with at least two difficult questions:

- How large a study is needed to find a solution to the problem at hand?

---

[1] Federal Aviation Administration, *ATS Performance Plan*, Washington, DC, 1999.

- How can I solve the problem without creating another one?

In practice, the scope of an airspace study is more likely to be defined by the sphere of influence of the organization running the study rather than by the nature of the actual problem at hand. For example, if a local group is chartered to conduct a study of arrival delays at a specific airport, then the scope of inquiry will normally be limited to arrival routes in the terminal airspace and adjacent en-route centers. This will lead to a local solution even though a subtle adjustment in regional traffic flow may have produced a far more satisfactory result. The same problem also occurs in reverse. If a national authority addresses system-wide inefficiencies, then its inquiry will likely focus on large-scale models and identify strategic solutions when a number of local changes may in fact be far more effective. Decision-makers need analytical tools that ensure the scope and granularity of a study is appropriate to the problem at hand.

Ensuring that a local solution does not cause problems elsewhere is even more difficult. In practice, each entity involved in a study tends to solve problems in a way that is "good" for its own people. For a controller, getting aircraft out of their airspace (and into somebody else's) in the most efficient way is desirable, even though this may cause a worse problem somewhere else. Solutions obtained in this way inevitably lead to disagreements among ATM facilities. Therefore, implementation of airspace changes is often dominated by negotiation rather than analysis.

To solve this problem, decision-makers need a quantitative, objective definition of what they are trying to achieve with the study. The FAA's *Airspace Management Handbook*[2] describes important steps to follow, in order to accomplish a practical reconfiguration of ATM resources. The first step in the process is analytical problem identification. The work described here is an example of that first step.

HOW TO FIND AIRSPACE PROBLEMS

There are two ways to identify airspace problems: the predictive and the deductive. Each has its advantages and disadvantages. Emphasis has traditionally been given to the predictive method.

Predictive Method

The predictive method uses simulation to remove ATFM ground delay programs and fly the aircraft as they would prefer. The simulated volume of airspace must be large, since the objective is to find problems. For airport problems, look for large taxiway delays on departure, or excessive airborne holding of arrivals. (45 minutes, the amount of holding fuel required in the USA, is a convenient standard for "excessive".) For airspace problems, using some workload metric, identify "red sectors", where the density of traffic is unacceptably high. This is unbeatable in principle, but developing the simulation means balancing the need for a wide-scale survey of the airspace with the need for detailed modeling to accurately reproduce the operating environment of each single sector.

Efforts to develop such a large-scale simulation are under way in various places around the world. NavCanada, at their headquarters in Ottawa, have simulated the entire Canadian airspace using the Total Airport and Airspace Modeller (TAAM). Simulating this enormous area is possible because of the relatively low total number of flights in that airspace, about 9,700 all told. At the Center for Advanced Aviation Systems Development in McLean, Virginia, USA, six air route traffic control centers are being simulated, a volume 800 by 700 nautical miles, containing some 45,000 flights per day. At the Eurocontrol Experimental Centre in Bretigny-sur-Orge, France, the airspace from Britain to Poland, north to Scandinavia and south to Italy, about 1000 nautical miles square, is being simulated using the Reorganized ATC Mathematical Simulator (RAMS).

The biggest advantages to this approach are that it gives the analyst nearly complete control over variables in the system, it can predict future congestion, and it can identify limits to growth. It can be a tool for active management of airspace, not just reactive adaptation once problems are identified. The biggest disadvantage is that it is slow and expensive – the steady increase we have seen in computer power is not accompanied by a corresponding increase in the ease of modeling the unique details of each ATM facility.

Deductive Method

The deductive method of identifying airport and airspace problems is data-intensive, but less work than validating a large simulation. The precise definition of metrics is determined partially by available data, as long as the categories defined above are covered. The ATFM system is monitored on an appropriate scale: airlines change their schedules in the USA about once a month, so long-term trends are based on monthly averages. The time series are inspected for patterns in the evolution of the metrics that are symptomatic of airport or airspace congestion.

After identifying patterns, consult with operational personnel in both the ATM facilities and the primary users' dispatch offices. This step is essential: since the ATM system serves many needs, it must be ascertained

[2] Federal Aviation Administration, *Airspace Management Handbook*, Washington, DC, 1999.

whether an identified phenomenon is a problem, or a solution put in place to cure something worse (See below). This method is relatively fast and inexpensive; our work covered flows to 14 airports in 6 months.

In practice, of course, a combination of the two methods offers best power and speed. Deductive problem identification can identify parts of the ATM system that do not cause problems for users (e.g., because excess demand is decreasing). Simulation offers a confidence in assigning causes to observed effects that lends credence to multi-facility airspace problems, which are not completely verifiable by a single domain expert.

DEFINITION OF AN AIRSPACE PROBLEM

Our operating principle is that *runways should be the limiting factor in an efficient ATC system*.

Runways are expensive to build and difficult to alter, once built. They are the least flexible part of the ATFM system. Airspace design, by contrast, is a matter of agreement between ATFM facilities. Apart from navigation aids, whose location is becoming less constraining as area navigation becomes more common, there is nothing physically fixed about airspace. Therefore, airspace should adapt to the runways. If it does not, that is, if runways are not being used to capacity, you have the archetypal airspace problem, measured in terms of throughput. Problems of this sort are so obvious that ATFM personnel are constantly making adjustments to the airspace to prevent them. The US Air Traffic Control System Command Center has recently begun efforts[3] to compare the agreed-upon flow rates at a number of US airports to the throughput actually achieved. These comparisons, which will be available to managers on a next-day basis, will institutionalize the use of throughput as a diagnostic performance metric.

Consulting with users of the airspace, however, shows that there are other concerns than these, which sometimes do not get addressed so efficiently. Other metrics than throughput[4] can be used to reveal other kinds of inefficiencies in the airspace. Combinations of trends in these metrics may point to airspace design problems, to flow management procedures mismatched to the traffic situation, or to other problems in the ATFM system.

Delay signature of airspace problems. A certain amount of delay is necessary to the proper functioning of any transportation system[5]. If an airspace route is desirable to the flying public, more aircraft will fly on it until it becomes congested and delays result. However, if delay appears without a trend of increasing traffic, it is a sign that something is wrong. “Delay” in this context may be either a departure delay imposed on the ground by ATFM, an increase over several years in the time to complete an operation, or a significant change in the number of an airline’s flights that arrive on time with respect to schedule. The notion of “excess demand” (see below) provides a measure of whether runways are the limiting factor, or something else.

Predictability signatures. Predictability of schedule is just as important to a commercial carrier as the reduction of delays (though less important to other classes of users.) Predictability is usually measured as the variation in some kind of movement time. In this work, the predictability of en-route time (time from wheels up to wheels down) is of interest[6]. At times of low traffic, this time will be short; at times of high traffic, speed controls, offloading of traffic, and *in extremis* holding will add to this time. . To minimize the effect of winds, a single day is analyzed at a time.

Since airspace designs are created with a particular flow in mind, the en-route time from each connecting airport to a common destination are collected for statistical analysis. The interquartile range of en-route times from each origin is calculated. For the purpose of summarizing flows to the airport, a histogram has been most useful[7]. It should be remembered, though, that the variability associated with an airport might not be due to the airport, but to a congestion point upstream. For this reason, variabilities are evaluated separately by direction of flow. In no case evaluated in our initial efforts, however, was there significant variation with direction.

Flexibility signatures. Just as ATFM turns a very bad thing (excess workload) into a less-bad thing (ground delays), stratification of traffic turns congestion into inflexibility. This represents a loss of flexibility in choosing altitudes for the users, which leads to inefficient fuel usage, increased wear on aircraft, and increased

---

[3] Abrahamsen, T and Cherniavsky, E., *National Airspace System Report Card*, MITRE Corp., MP99W106, June 1999

4 Hoffman, J. & W. Voss, *Overview of the ASC System Performance Management Project*, Proceedings of the 1st US/Eurocontrol ATM Seminar, Saclay, June 1997.

5 W. Zimmerli, “Future Traffic – A Challenge to Human Intelligence and Social Values”, *in Advanced Technologies for Air Traffic Management*, Bonn, Springer-Verlag, 1994.

[6] This metric is due to Thomas Bock of the FAA Eastern Regional Office (private communication).

[7] Hoffman, et al. *Initial Eastern United States Airspace Problem Identification*, McLean, the MITRE Corporation, MTR99W32, April 1999.

pollution. While these can be serious, these are of secondary importance to users (compared to arriving on schedule). However, if a certain kind of flow is continually given preference, it can lead to inequitable treatment of different classes of users. Measuring altitude times from top of descent to touchdown, and times from wheels up to top of climb, can give an idea of which users are benefiting from a particular airspace organization, and which are bearing the penalties.

Things that are not Airspace Problems.

Excess Demand. If demand increases beyond the capacity of runways, delays will result that are only treatable by capital improvements (add runways) or reduced separation standards (e.g. via active wake vortex detection). The excess demand metric compares the scheduled traffic at an airport to a theoretical measure of capacity, and adds up the flights that are scheduled above the capacity of the runways to handle them. It is recognized that the "true" capacity is difficult to capture; however, since the objective is to identify changes in excess demand over several years, the exact value of the theoretical capacity is not very important. In this work, simulation results were used where available, Engineered Performance Standards (EPS) were used where not.

The US Air Traffic Control System Command Center's next-day throughput assessment, described above, will yield as a by-product a record of the flow rates airports agree to accept, including airport configuration and meteorological conditions. As the numbers from this evaluation become available, they may replace the EPS, providing a substantial improvement of this metric.

Non-ATFM factors. Low throughput may be mandated by statutory limits on airport traffic, as in Washington, New York, and Chicago. If these limits are based on considerations other than delay reduction, these metrics will show signs of a capacity problem, usually in the form of arrival and departure throughput below runway capacity.

DATA SOURCES

For long-term tracking, the Airline Service Quality Performance (ASQP database) was used. Times of pushback, wheels up, wheels down and arrival at gate are available for all domestic flights by the ten largest airlines in the US. For the flows of interest in the Eastern US, more than half of the traffic is represented. ASQP data were used when averages over traffic were sufficient. Data are available from January 1994 to the present.

When trajectory data are needed, the Enhanced Traffic Management System (ETMS) provides radar data with a four-minute update rate for all flights under US or Canadian control. The version used here was the Aircraft Situation Display to Industry, version 4.2.

This analysis was primarily looking for performance shortfalls under the best possible operating conditions, so days of good weather were chosen. Thursday, October 1, 1998 was a heavy traffic day of generally good weather (only 1 hour of fog at Boston). Friday, October 23, 1998 was a day of clear weather across the entire continental US. No major equipment problems were reported on either day. When total traffic numbers are necessary, as in the case of the actual throughputs, these two days are shown.

RESULTS

Overview

Generally, across the country, increases in delay were seen from 1994 to 1997. The delay performance of the system improved significantly in 1998. A survey of OAG schedules showed that a large part of this improvement was due to changes in aircraft schedules, reducing the excess demand. At several of the biggest airports, which were examined here, there were several cases in which the delays did not improve, even given reduced excess demand.

Low arrival throughputs are seldom seen. As the most obvious airspace problem, low arrival rates tend to be fixed as soon as they appear.

Taxi out time is increasing, often in conjunction with low departure throughputs. This is a conspicuous sign that traffic flow management programs are working.

Examination of the throughputs at various points in the airspace is facilitated by a table that includes a positive/negative indicator for each stage of aircraft flight (Table 1.) The table properly has 64 rows, corresponding to all possibilities. Rows that do not correspond to observed behavior in the Eastern US were omitted for brevity.

Below, a number of interesting phenomena observed at individual airports are described in detail.

Table 1. Throughput and Delay Summary

Airports	Delay increase w/o excess scheduling?	Increasing taxi-out time?	Departure throughput below EPS?	Large en-route time variability	Center-to-TRACON throughput below EPS?	Possible Airport or Airspace Problem?	Comments
ATL				X		A/S	look for inefficient routing
BOS	X	X	X			maybe	early stages of airspace problem
CLT, MCO			X				not much traffic, watch out if demand increases
CVG	X	X	X		X	A/S	TFM over-resolving airspace congestion
DCA, JFK		X	X		X	A/S	airspace congestion over-solved (both slot controlled)
EWR		X	X	X		maybe	could be airport arrival capacity or TRACON problem
LGA	X	X		X		A/S	insufficient en-route capacity
MIA		X				A/P	(a) departure capacity limits, or (b) TFM decongesting airspace
ORD	X	X	X	X	X	A/P, A/S	airport and airspace problems combined
PHL	X	X	X	X		A/S	TRACON departure capacity problem or en-route congestion
PIT, DTW							no sign of problems from these metrics

ORD

O'Hare International Airport in Chicago is a singularly busy airport, serving as a hub for two of the largest air carriers in the US. Demand for space at ORD is so great that neither the airport nor the airspace can handle it. The airport is slot controlled, but this did not prevent all of the metrics in this study from showing performance shortfalls. Arrival and departure delays have been rising since 1997, even though the excess scheduling has decreased (figure 1.) Seasonal fluctuations in average delay are large. The most pronounced performance shortfall is in average taxi-out time (figure 2). Since 1995 there has been a steady trend toward longer taxi times, from 16 to 19 minutes per flight on average. Given that ORD sees about 1100 departures per day, this excess taxi-out time represents about 50 hours of extra time on taxiways per day.

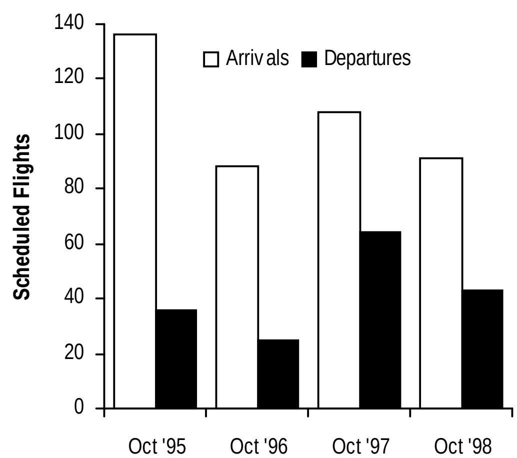

Figure 1. Excess scheduling at ORD

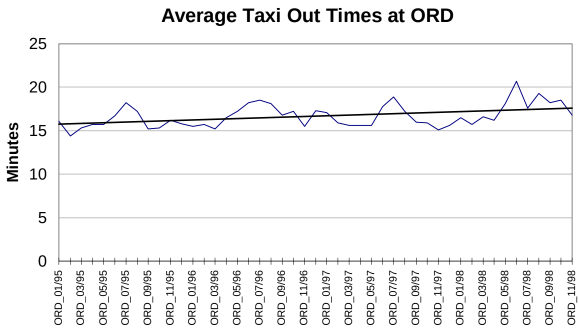

Figure 2. Taxi-out time at ORD

Chicago sees less departure throughput than its runways can handle, even on a day when there is no inclement weather at its connecting airports. This is a sign of aggressive traffic flow management, avoiding capacity problems.

EWR

The Newark International Airport was one of the 27 airports that experienced more than 20,000 hours of annual delay in fiscal year 1997. Almost 6% of the EWR operations experienced delays of 15 minutes or more, making EWR the worst airport in the nation for percentage of operations experiencing these delays. A number of factors combine to produce the high delays at EWR. The most obvious is demand. The New York metropolitan area is served primarily by EWR, LGA, and JFK airports. Of these, LGA and JFK are slot-controlled, so traffic can not increase there. Any airline wishing to add flights to New York must fly either to EWR, or begin from zero at one of the smaller airports in the vicinity.

A second, and perhaps more important, factor, is the highly congested airspace around New York City. The presence of two other major hubs deprives EWR of the airspace it would ordinarily use to organize and separate arrival and departure flows. The result, as can be seen in Figure 3, is restrictions on departures. There are three lines on this chart. The heavy horizontal line is the departure runway capacity in its preferred departure configuration, according to the Engineered Performance Standards. The dotted line (“OAG”) is the scheduled departure demand, according to the Official Airline Guide for October 1998. The solid line (“Actual”) is the departures recorded by the Enhanced Traffic Management System for Thursday, October 1, 1998, a day of generally good weather.

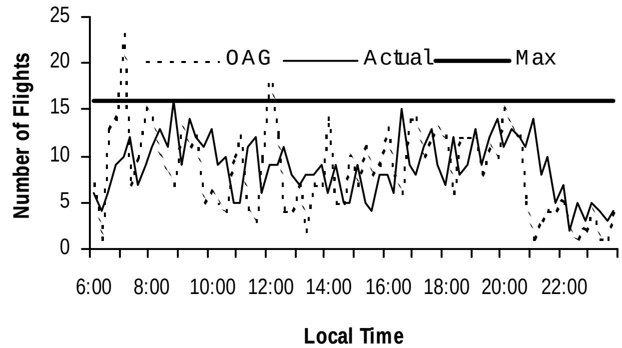

Figure 3. Departures from EWR.

The runway capacity is not fully utilized, even on this day of good weather. Demand is there: the scheduled arrival line shows several times where demand exceeds capacity (*e.g.* near 0700 and 1200). Even worse, between 2100 and 2200 local time, there is little scheduled demand, but significant departure throughput. These are flights with an hour or more of departure delay. Some part of the ATM system is limiting throughput, and it is not the runway. This is the clearest sign of airspace capacity limits in the current system. Arrivals at EWR do not show unused capacity, primarily because holding arrivals would aggravate airspace congestion.

ATL

In Atlanta, the various throughput and delay metrics show no deviation from theoretically expected behavior. However, the predictability of en-route times is poor compared to the national average (Figure 4). Variabilities in en-route time of 6-12 minutes are common. The variability is due to extensive airborne holding, as can be seen by inspection of flight trajectories. Taken by itself, this would be a sign of airspace congestion. However, the lack of delay means that further investigation is called for.

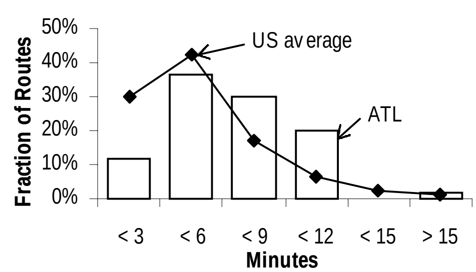

Figure 4. En-route time variabilities for ATL

In this case, interviews with ATC personnel revealed that Delta Air Lines, the primary carrier at ATL, preferred airborne holding to ground delays. Since Atlanta is far from any other large airport, there were no major flows conflicting with the airspace reserved for holding, so it was possible for the en-route Center and approach control to oblige. Therefore, this is not an airspace problem – it is an example of how the system was designed to work. The benefits to users are seen in Figure 5. This chart, similar to Figure 3, shows theoretical, planned, and actual arrival throughputs at ATL. The theoretical capacity was frequently exceeded by the actual arrivals, possibly as a result of the fact that approach controllers have a reservoir of holding flights that they can insert into arrival streams with great efficiency.

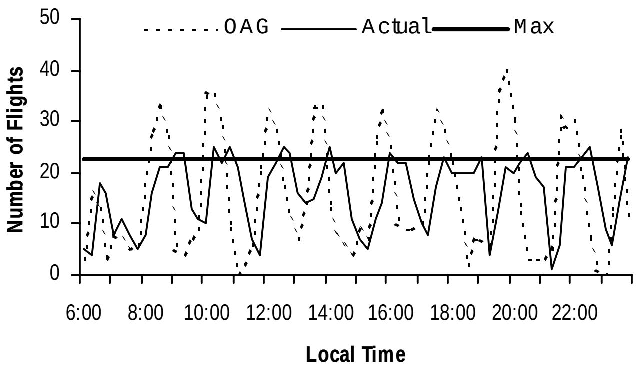

Figure 5. Arrival Throughputs at ATL

CVG

Greater Cincinnati International Airport is a relatively new hub for Delta Air Lines and its regional partner ComAir. Since 1995, traffic has grown considerably, and an additional 73% growth is expected in the next 15 years. From mid-1997 through 1998, delays have decreased, mostly because excess demand has decreased. The airlines have been expanding their scheduled flight times for some reason, presumably a capacity limit. Arrival throughput into the terminal matches runway capacity, but departure throughput is below capacity. As above, this is a sign that en-route airspace is the limit in the system.

CVG is particularly interesting because of its altitude flexibility metric (Table 2). Jets in the northeastern departure flow take, on average, almost 10 minutes longer than the other flows to reach cruising altitude. Upon inspection of climb profiles, it was found that regional jet departures were filing lower altitudes in their flight plans, presumably to avoid the famously congested airspace in the area above 23,000 feet. In this case, the airline has traded one form of penalty for another, accepting increased fuel consumption to reduce delay. The multi-dimensional metric approach to analyzing system performance has exposed the fact that the throughput score at CVG is actually misleadingly high.

Table 2. Median Descent and Climb Times (in minutes)

Arrivals from:	Descent	Departures toward:	Climb
Northeast	32	North	32
Southeast	34	NE	**41**
Southwest	36	SE	32
Northwest	40	South	26
		SW	31
		West	32
mean for Eastern US	34.9	mean for Eastern US	21.4

CONCLUSION

The balance between predictive and deductive analysis of ATFM problems has long been tilted in favor of predictive simulations. This paper has outlined a method by which deductive analysis can use widely available data to identify airport and airspace problems to support Airspace redesign and guide simulations. We have shown examples of how individual impacts can be extracted from the complex network of the ATM system.

This work has been presented to many of the Facility Liaison Teams involved in the National Airspace Redesign. It provides a synoptic view of the system, reaching across facility boundaries to ensure that individual facilities' efforts to ameliorate their own problems can be done with due consideration of the impacts on neighboring airspace. Ultimately, it is hoped that this work can guide simulation studies of proposed redesigns, making it possible to estimate the relative benefits of alternative airspace organizations that includes the second-order benefits and penalties that occur as users adapt to the new system.

ACKNOWLEDGEMENTS

We would like to acknowledge the invaluable contributions of Mr Thor Abrahamsen, Mr Samuel Bowden, Ms Debra Moch-Mooney, and Ms Tess Brothersen, who performed much of the research on which this work is based.