Proceedings of the 2006 Winter Simulation Conference
L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, eds.

UNDERSTANDING ACCIDENT AND EMERGENCY DEPARTMENT PERFORMANCE USING SIMULATION

Murat M. Gunal
Michael Pidd

Department of Management Science
The Management School
Lancaster University
Lancaster, LA1 4YX, UNITED KINGDOM

ABSTRACT

As part of a larger project examining the effect of performance targets on UK hospitals, we present a simulation of an Accident and Emergency (A&E) Department. Performance targets are an important part of the National Health Service (NHS) performance assessment regime in the UK. Pressures on A&Es force the medical staff to take actions meeting these targets with limited resources. We used simulation modelling to help understand the factors affecting this performance. We utilized real data from patient admission system of an A&E and presented some data analysis. Our particular focuses are the multitasking behaviour and experience level of medical staff, both of which affect A&E performance. This performance affects, in turn, the overall performance of the hospital of which it is part.

1 INTRODUCTION

District General Hospital Performance Simulation (DGHPSim) is a collaborative study involving three British universities that aims to develop generic simulation models of entire acute hospitals so as to understand how hospital performance can be improved. The idea is that the models will be of use to policy makers in setting healthcare priorities and to hospital managers in managing their services effectively and efficiently. In the first stage of DGHPSim we focused our efforts on the modelling of Accident & Emergency (A&E) Departments, which are found in most large acute hospitals. Later stages of DGHPSim, not described here, link the A&E modelling to the care of inpatients so as to model the performance of the whole hospital.

The U.K. National Health Service (NHS) has a performance measurement framework which forms part of an improvement regime. Every year, hospitals in England are assessed against performance targets, many of which are based on patient waiting times. The performance regime was established because of over-long waiting times and seems to have been a success in reducing these (Bevan & Hood, 2006).

A&E Departments of hospitals are intended to deal with critical or life threatening incidents rather than minor injuries or illnesses. They must meet an uncertain demand from patients, some of whom can be treated within A&E, others of whom are admitted to the hospital for further treatment as inpatients. Their performance affects the patients they serve and also the rest of the hospital, since they generate new inpatients. Therefore, understanding A&Es is important not only for effective use of limited resources in the A&Es but also in the wider hospital setting.

Currently, there are two key performance targets for A&Es: “Total time in A&E: four hours or less” and “12 hour waits for emergency admission via A&E post decision to admit”. The second target involves the availability of ward beds therefore it is beyond the control of A&E. Here we focus on the percentage of patients who are seen within four hours of arrival at A&E.

1.1 A&E waiting time targets

The UK Department of Health introduced performance measures for A&Es in 1997 and these have taken two forms.

1. Waiting time of patients from arrival until seen by a doctor or a trained nurse. The target was set at 15 minutes when introduced in 1997 (Department of Health, 1997). As might be expected, the introduction of this target led to gaming and some A&Es employed a ’welcome nurse’, given the task of seeing each patient within 15 minutes of their arrival, but doing little or nothing to treat them.
2. Total time of patients in A&E, measured from arrival to discharge or admission. This was intro-

> duced in 2002 and replaced the 1997 measure. A&E Departments are required to measure the % of patients whose total time in A&E exceeds 4 hours. Currently, these breaches must not exceed 2 %. Though this target presents fewer opportunities for gaming, there is the risk that patients will be discharged too soon or admitted prematurely as inpatients. The former could lead to poor quality of care and the latter could transfer the pressure elsewhere in the hospital.

When introduced in 2002, there were to be no breaches of the 4-hour A&E target,. Thankfully, the folly of this was recognised and the current 2% breach level was set in 2004/5. The relaxation was introduced following recognition that some patients will need extended care in emergency departments for good medical reasons (Department of Health, 2003). However, the 2% relaxation on the target percentage is not sufficient to take the pressure from A&E departments. Locker and Mason (2005) analyses performance data from 83 A&E departments' in England and reports that total times of patients peaks just before the "4 hour total time target" (see figure 3 for an example). They also observe that 1 in 8 patients who are subsequently admitted to hospital are moved out of A&E just before their stay would breach the 4 hour target.

Hence, it seems clear that the 4-hour target and the small number of breaches allowed are affecting the performance of A&E departments in England. Waiting times are lower than before the performance targets were introduced and much of this may be due to improved management. However, there is a risk that clinical standards are compromised or problems are just squeezed out of A&E and into the rest of the hospital.

2 AN A&E CASE

We worked with a mid-sized A&E department in the UK, which sees approximately 45000 patients annually and that has met the 2%, 4 hour total time target since its introduction. Using Micro Saint Sharp, we developed a discrete event simulation of this department's activities, with the intention that this serve as a generic model of A&E departments that can be parameterised to fit a range of such departments in different hospitals. The outline process flow in a typical A&E department is as follows:

1. Patient arrives and is registered.
2. Triage to determine severity of condition.
3. Patient waits for a doctor
4. Patient seen by doctor & nurse, who may complete the treatment and discharge or admit.
5. Some patients need tests and X-rays and these then need a second session with a doctor & nurse before discharge or admission.

There are 2 significant complications. First, some patients arrive by ambulance and may bypass registration and triage (registration can be done en route to the hospital). Second, doctors and nurses multi-task: that is, they see more than one patient simultaneously.

To parameterise and validate our model, we analyzed the electronic patient admission data for 2004/05. Even though the data is sufficiently detailed and captures every patient's time in various stages in the A&E, there were some missing elements in it such as doctor's total contact time with a patient and requests for X-Ray and tests. Although these are not available electronically, they are recorded on paper-based patient cards which are completed by medical staff and record every detail of patient treatment. How accurate these records are is unclear, especially when staff are highly pressured, but it was the best available. We collected data from approximately 600 patient cards which were selected randomly over a two month period.

2.1 Triage system and doctor time

At some time or other, most A&E departments in the U.K. have used a 5-colour triage system: Blue, Green, Yellow, Orange and Red. In this, Blue are the least severe cases and Red are real emergencies in which life is at risk. Since we needed to model the triage process in our simulation, we examined the performance of the 5-colour triage system in our client A&E. Table 1 shows the mean and standard deviation of doctor times of the sample population (of size N) by triage colours and percentages of X-Ray and any type of test requests for each of the 5 categories. "Doctor Time" is the elapsed time between the first seen by a doctor and end of doctor treatment and includes the waiting and process time for X-ray and tests.

Table 1: "Doctor Time" and percentages of X-Ray and Tests statistics

Triage Colours	**N**	**Doctor Time**		**% of X-Ray Requests**	**% of Test Requests**
		Mean	**Std. Dev.**		
Blue	44	0:32:25	0:33:20	31%	3%
Green	284	0:32:08	0:33:52	37%	6%
Yellow	122	1:04:18	0:47:10	53%	55%
Orange	79	1:02:28	0:44:21	62%	74%
Red	22	0:54:30	0:30:00	74%	70%

It is clear from Table 1 that the mean and standard deviation of doctor times for Blue & Green and Yellow & Orange patients are very similar. A similar observation can be made for X-ray and test request percentages. This suggests that the 5 colour triage system is not working well. We speculate that patients are actually triaged into 3 categories that we label as Minor, Major and Life Threatening.

This was later confirmed by the lead clinician in the department and seems likely to be the case in other A&Es.

This is important, since triage category is an important attribute of patients in our model. Rather than using a diagnostic code to distinguish patients we use triage category as a parameter in treatment time distributions. Using three is simpler than using five.

2.2 Demand patterns

Patients arrive by either ambulance (25%) or as 'walk-in' cases (75%). Each mode of arrival has different demand patterns. Figure 1 shows that ambulance arrivals do not change much by hour of day whereas 'walk-in' arrivals do. The sharp peak at 9am for walk-in patients shows an anomaly, dissected in Figure 2. On checking the anomaly, it seems that some patients return to A&E around 9am for dressing and fracture clinics which take a very short time and consume little in the way of A&E resources. Thus, the clinic has two modes of arrival.

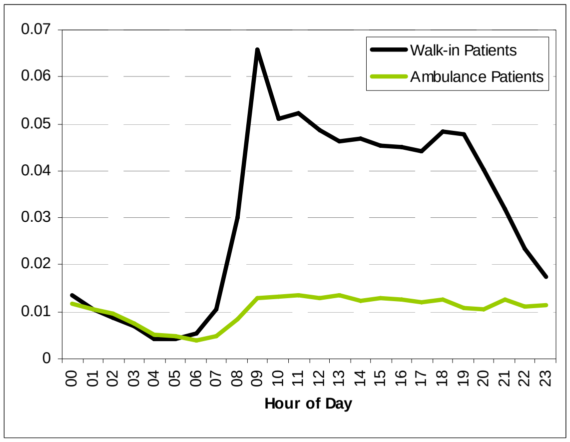

Figure 1: Patient arrival pattern by mode of arrival

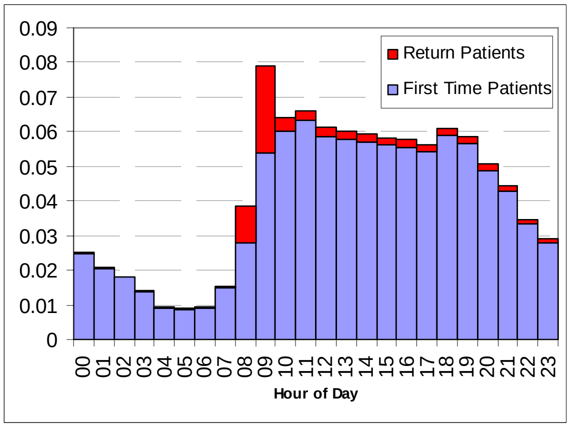

Figure 2: Patient arrival pattern by hour of day

Since the return patients indicate that all arrivals are not independent, we examined first arrivals in the data for Poisson attributes. K-S tests revealed that hourly inter arrival times from the data follow Negative Exponential distributions for both modes of arrival. Because of time varying arrival rates we employed thinning to represent the non-stationary process to sample inter arrival times of patients in the simulation model (Lewis and Shedler, 1979).

2.3 Total time of patients in the A&E

As one would expect, more severely injured patients spend more time in the A&E than less severe patients. Figure 3 shows total time spent in A&E by Major and Minor patients. As observed by Locker and Mason (op cit) there is a peak waiting time just before 4 hours, presumably caused by the 4-hour target, and this is especially severe in Major patients. Since most Major patients are subsequently admitted, this suggests that it takes time to decide whether or not to admit some of these as inpatients; or that there is a delay whilst waiting for a bed.

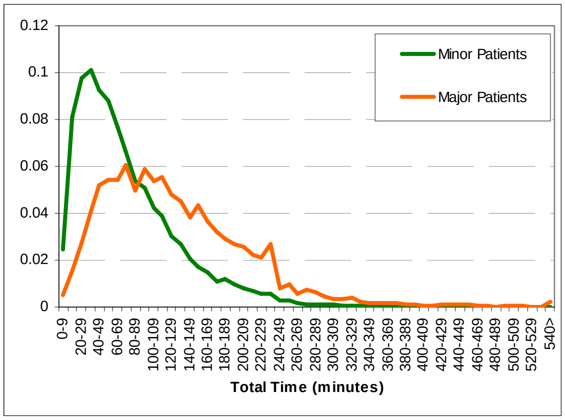

Figure 3: Total time of patients by triage category in the A&E histogram

3 SIMULATION MODEL

3.1 Introduction

We developed a discrete event simulation model of the A&E department by using Micro Saint Sharp, based on a task network representing the process flow of patients. For example a walk-in patient is first registered, triaged, treated, sent to X-Ray, re-evaluated and discharged each of which has different service time distributions which depend on patients' triage category. We used triangular distributions as the service time distributions of registration, triage and treatment processes. For re-evaluation and X-Ray service times, a log-normal distribution is used. Patients consume different resources (staff and room) depending on their triage category. Inputs to the model are:

- Patient arrival volumes and patterns for walk-in and ambulance patients,
- Staff by hour by role (senior doctor, junior doctor and nurse),
- Physical bed (cubicle) capacity,
- Service time distributions parameters by triage categories and by doctor type,
- Test and X-ray percentages by triage category
- Patient population's triage category distribution

The model gives the total times of patients in A&E and percentages of patients who breach the 4 hour target. Instead of using a warm-up period, we started the model in empty state. Historical figures reveals that on Thursdays between 4 and 5 am, there was almost no one in the system. Therefore the model is started at simulation clock "Thursday 4 am".

Since the model is intended for generic use by data parameterisation, the A&E processes and parameters are as general as possible so that it can be tailored for any other A&E department.

3.2 Mini Doctors for Modelling Multitasking

Doctors and nurses are scarce resources in A&Es and, most of the time, they treat multiple patients concurrently. Whilst a patient is waiting for test results or X-ray, a doctor may go and see another patient in another cubicle; that is, they multi-task. Multi-tasking has been studied in other domains, (see Elfving and Tommelein (2003), Spink et al (2002) and Wild et al (2004)) but very few analytical studies have been conducted in health care. Carter (2002) addresses the challenge of simulation modelling healthcare and stresses the difficulties in data collection and in determining how to model staff time.

Empirical work includes Gibson et al (2005) and Chisholm et al (2000). The latter reports a time & motion study to determine the number and types of interruptions in Emergency Departments (EDs) in the USA. One person shadowed emergency physicians (EPs) for a 3 hour period, every day for a month. They defined 8 possible "tasks" for EPs, such as patient care, viewing diagnostic test results etc. and "interruption" as any event that briefly required the attention of the subject but did not result in switching to a new task. If the subject decides to switch a task then this is recorded as a "break-in-task". The results of this study revealed that the number of patients simultaneously managed per 3 hour period is 5.1± 2.1 and number of break-in-task is 20.7 ± 6.3. Also they observed a statistically significant positive relation between the number of patients who visited the EDs and the number of break-in-tasks.

There are 2 obvious ways to model the multitasking behaviour of medical staff in a discrete event simulation;

1. Fragment the process of doctor's interaction with a patient into "S" numbers,
2. Fragment a doctor into "M" parts (or say "Mini doctors").

"S" and "M" determine how many patients doctors can handle simultaneously.

Both are artificial but practical solutions to tackle this problem and each has advantages and disadvantages. Option 1, fragmenting the doctors' and nurses' interactions with patients, requires reasonable estimates for each interaction time, which requires data that is very difficult to collect. However this is a more realistic representation. Option 2, fragmenting a doctor or nurse into mini doctors, is less realistic but is easier to implement. Its main disadvantage is that it could underestimate doctor and nurse utilisation figures. However, since we focus on patient waiting times in different stages in A&E rather than staff utilisation, then using mini doctors & nurses seems a sensible choice.

Hence we model multi-tasking by fragmenting each doctor and nurse into M parts. Based on Chisholm (op cit) we estimated this parameter as 6 for senior (and experienced) doctors, 4 for junior (and inexperienced) doctors and 2 for nurses. We called these numbers "Multitasking Factors (MTF)".

3.3 Validation

Our main objective in the model is to estimate the percentage of breaches of the 4 hour total time target. Therefore the model generates simulated patients total times in the A&E, from which we find the percentage of breaches. As part of our validation process, we used one year's data that records all A&E attendances. The model is designed to work for on a weekly basis for two reasons. First, the staff roster is organised weekly and secondly, there is little or no seasonality in arrival patterns, though there is much in-week variation. The model is run for 52 weeks and 50 times in each experiment.

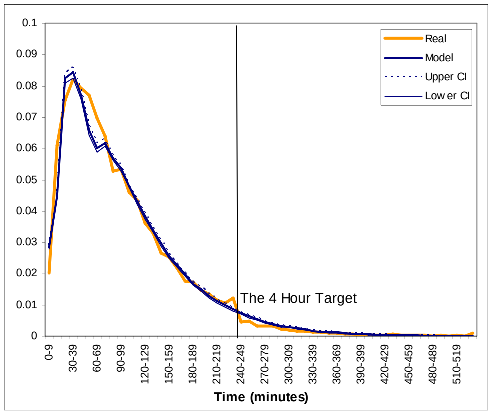

Figure 4: Patient total time in A&E histograms

Figure 4 shows the total time spent in A&E, for the real and simulation outputs. Both lines exhibit a good fit in most parts. However we are more interested with the tail of this histogram that is the part after “the 4-hour target”. As in Figure 3, the blip in the orange line just before the target is caused by ‘panic’ intervention to meet the target. Because it is hard to represent this behaviour in the model, we get slightly higher percentage of 4 hour target breaches than real. Other than that, the model seems satisfactory.

3.4 Experimentation

We focus on two things in experimentation with the model; first, what is the effect of multitasking on the performance and second, what other factors affect performance. To investigate multi-tasking, we ran the model with different MTFs and we varied treatment times, X-Ray service times and percentages and physical cubicle capacities and investigate other factors.

Figure 5 shows how simulation output changes with different MTFs. MTF numbers are shown with the order “Nurse-Junior Doctor-Senior Doctor”. As explained earlier, these numbers show the number of patients that a doctor (or nurse) can treat concurrently. For example in the base model (used in the validation), the meaning of “2-4-6” is that a nurse can treat 2, a junior doctor can treat 4 and a senior doctor can treat 6 patients at a time. One would expect better performance with higher MTF values because more patients can be treated in the same time period.

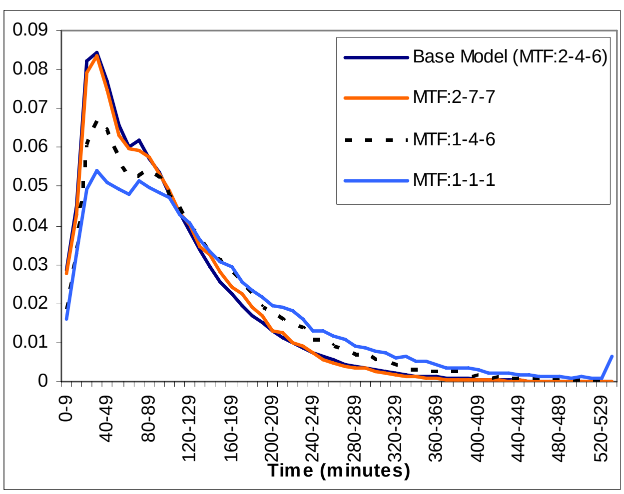

Figure 5: Model outputs with different MTFs

The base model and “MTF 2-7-7” output lines almost converge. This suggests that more multi-tasking of doctors (or having more “mini doctors”) may not increase the performance. However decreasing nurse MTF by 1 leads to worse performance, which suggests that the limiting factor for better performance may be the number of nurses in this A&E as currently staffed. On the other hand “MTF 1-1-1” line shows the worse performance of all suggesting that the multitasking of staff, in general, is a real determinant of the performance.

To understand how performance is affected by other factors, we examined the four scenarios presented in Table 2, in which we varied the service times of doctors, service times for X-Ray and tests, the proportion of X-Ray requests and the number of cubicles.

Table 2: Experimentation scenarios

Scenario Name	Meaning
Experienced Doctors	We assumed that experienced doctors (e.g. senior doctors) can treat patients quicker; that is, they make decisions faster than inexperienced doctors. In this scenario we simulate non-senior doctors working at the same speed as senior doctors, but retain base case multi-tasking.
Longer X-Ray and Test Time	Waiting times and process time for X-Ray and tests are very significant delays in patients’ journey in A&E. In this scenario we increased process times for X-Ray and Tests by 10 minutes above the base case.
Less X-Ray	Not all patients require an X-Ray but some proportion of them do. We assumed in this scenario that this proportion is 10% less than in the base case.
More Cubicles	In this scenario we assume that we have more cubicles (5 Resus, 14 treatment cubicles which is a 2 cubicle-resus, 4 cubicle-treatment cubicle increase).

We used a Two-Sample Kolmogorov Smirnov test to compare the differences between the base model and other scenarios in experiments. The comparison revealed that “MTF 1-1-1” and “Longer Xray and Test time” scenarios statistically differ from the base model scenario at 10% significance level.

Simulation outputs are shown in Figure 6 for these scenarios. Three lines, “Experienced Doctors”, “Less X-Ray” and “More Cubicles”, almost converge. These scenarios demonstrate better performance than in the base model. On the contrary, “Longer X-Ray and Test time” scenario exhibits the worst performance of all. It is difficult to draw any direct conclusions from these results. However these scenarios help us understand how performance changes with different parameter values. For example, if the A&E department had only very experienced and fast-decision maker doctors or the doctors requested 10% less X-Rays from patients, the overall performance would be better. Likewise, having a 10 minutes increase on average, X-Ray and test times will lead to performance drops.

Though these scenarios are for illustration purposes, they show how the model may be used to investigate performance options.

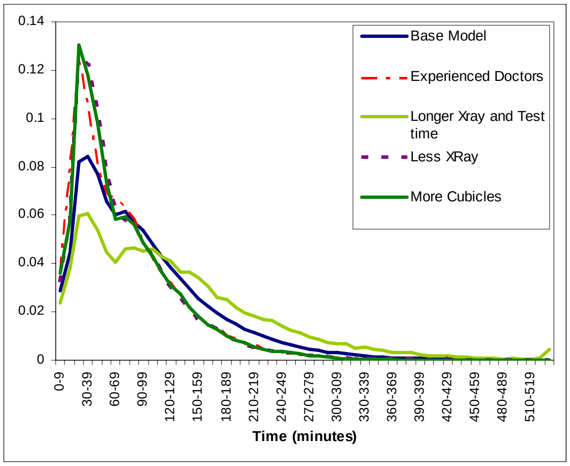

Figure 6: Model outputs with different scenarios

4 CONCLUSION

From our data analysis we observed that the 5 category triage system is not being used in practice. Patients are actually triaged by 3 categories. The aim of triaging is to prioritize patients so that more severe cases are treated before less severe ones. However it is arguable that formal triaging is necessary at all given the fact that the actual triage system tend to categorize patients as "Minor", "Major" and "Life threatening". It is easy to categorize patients by these three, even without medical staff intervention. Triaging patients adds extra time to the patient total times in the department and affects performance.

The simulation model we built is a conventional emergency department simulator which predicts performance under different circumstances. Performance is measured as the percentage of patients who stayed in A&E more than 4 hours. Our aim is to show medical staff how performance is affected by various factors. For example one concern is that a likely change in rotation period of junior doctors from 6 months to 4 months may affect performance. To investigate the effect of this change, we set up two scenarios to experiment in the model. Experienced doctors spend less time with patients and request fewer investigations to make decisions. It is the opposite for inexperienced doctors that is they are slow in decision making. "Experienced doctors" and "Less X-Ray" scenarios revealed that experience level of doctors are determinant in improving performance. As another example, we run the model with the increased X-Ray process time and observed that X-Rays take great time in the patients' length of stay in A&E.

One of the significant characteristics of the A&E environment is that medical staff multi-task. Staff treat more than one patient at a time, especially when the system is congested. We modelled this behaviour by using "mini-doctors", which is a factor of how many patients can a doctor see simultaneously. We experimented with different multitasking factors on our model and observed that multitasking affects performance. We also observed that the binding constraint on the performance, in our client A&E, seems to be the number of nurses.

The method we used for modelling multitasking behaviour of medical staff is a new and requires further research for improvement. Using MTF to model multitasking humans can be applied to other fields such as air traffic controllers (ATC), police and ambulance radio dispatchers.

The model is the first stage of a generic simulator to predict hospital performance and will serve as one of the generators of inpatients for a typical hospital. To do so, it must be parameterised with data from different A&Es in the UK. It will then be linked to models of inpatient care as a component in a total hospital model performance simulator.

ACKNOWLEDGMENTS

This research is funded by the Engineering and Physical Sciences Research Council (EPSRC) in the UK under grant EP/C010752/1. Research collaborators are Gwyn Bevan & Alec Morton (London School of Economics), Peter C Smith (University of York), Morecambe Bay Hospitals Trust, Royal Orthopaedic Hospitals Trust and the Department of Health.

REFERENCES

Bevan R.G. and Hood C. (2006) Have targets improved performance in the English NHS? BMJ. 2006 February 18; 332(7538): 419–422.

Carter M. (2002) Diagnosis: Mismanagement of Resources. OR/MS Today. April 2002, 29/2.

Chisholm C.D., E.K.Collison, D.R. Nelson, W.H. Cordell (2000) Emergency Department Workplace Interruptions: Are Emergency Physicians ''Interrupt-driven'' and ''Multitasking''? Academic Emergency Medicine, November 2000, Volume 7, Number 11, 1239-1243.

Department of Health (1997) Patient's Charter developing a new charter and changes to the immediate assessment in A&E standard, www.dh.gov.uk/assetRoot/04/01/14/32/04011432.pdf [accessed 1 March 2006].

Department of Health. (2003) Clinical exceptions to the 4 hour emergency care target. www.dh.gov.uk/assetRoot/04/07/95/56/04079556.pdf [accessed 1 April 2005].

Elfving J.A., I.D.Tommelein. (2003) Impact of multitasking and merge bias on procurement of complex equipment. Proceedings of the 2003 Winter Simulation

Conference. S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, eds.

Gibson M., K.N. Jenkings, R. Wilson, I. Purves (2005) Multi-tasking in practice: Coordinated activities in the computer supported doctor-patient consultation. International Journal of Medical Informatics, 74, 425-436.

Lewis P.A.W. and Shedler G.S. (1979) Simulation of non-homogeneous Poisson process by thinning. Naval Research Logistics Quarterly, 26, 403-13.

Locker T.E. and S.M. Mason (2005) Analysis of the distribution of time that patients spend in emergency departments, British Medical Journal 330 (2005) 1188-1189

Spink A., H.C. Ozmutlu, S. Ozmutlu (2002) Multitasking information seeking and searching processes. Journal of the American Society for Information Science and Technology; Jun 2002; 53, 8;

Wild P.J., P.Johnson, H.Johnson (2004) Towards A Composite Modelling Approach for Multitasking. Task Models and Diagrams for User Interface Design (TAMODIA'2004) 3rd International Workshop, Prague, Check Republic, November 15-16 2004.

AUTHOR BIOGRAPHIES

MIKE PIDD is Professor of Management Science at Lancaster University where his work spans simulation modelling and the development of improved simulation methods utilising current developments in computing hardware and software. His current application work focuses on simulation modelling for improvement in public services, especially in policing and healthcare. He is known for three books: Computer simulation in management science (in its 5th edition), Tools for thinking: modelling in management science (in its 2nd edition) and Systems modelling: theory and practice – all published by John Wiley. m.pidd@lancaster.ac.uk.

MURAT GUNAL is pursuing his PhD at Lancaster University. He received his MSc degree from the same university in 2000. His main research interest is to investigate the use of simulation on performance measurement in public services such as hospitals. He worked as a simulation analyst in the Turkish Navy where he was interested in military applications of operational research. Email to m.gunal@lancaster.ac.uk.