ASSESSING SIMULATION LEARNING IN HIGHER EDUCATION

Brian Hollocks,

Bournemouth University Business School,
Christchurch Road,
Bournemouth, BH1 3LH, U.K.

ABSTRACT

System simulation may be used as a valuable exemplar of modeling in degree programs across a range of disciplines. This paper describes and discusses an approach to assessing learning in that domain, relating it to Bloom's and Biggs' taxonomies. It uses a mini-project, scenario concept, evaluated through oral examination. The scenario text and grading plan are included in the paper. The approach is found to usefully differentiate across the full range of student performance.

1 INTRODUCTION

Arguably, system simulation is the form of management science modeling which has produced the most business impact. It provides a representation of the real world in such relevant respects that experiments carried out using the model are predictors of what will actually happen. Hence, interactive use of a simulation model by or on behalf of a decision-maker provides information from which they can make more informed judgments. Application has been widespread across manufacturing, service, transport and other industry sectors, addressing investment, process restructuring, performance improvement and other problems. The track record of system simulation is impressive.

Naturally, therefore, simulation becomes a strong candidate as an exemplar of modeling for use in University degree courses concerning quantitative methods. In this role it has valuable characteristics; it has clear practical application, it requires quantitative underpinning (for example, in random variate generation and in experimentation), it offers a strong structure-type, it has a formidable supporting literature, it involves the use of supporting software, and it requires a high level of conceptual thinking. This in turn offers interesting challenges to teaching such a subject and to assessing it.

This paper describes and discusses one assessment approach which has been successfully used for a number of years in the context of an undergraduate course at Bournemouth University Business School. The approach is based on a fictional case study, derived from manufacturing industry. The case, as presented to students, is included verbatim, together with the protocol and grading framework.

2 CONTEXT

Bournemouth University is one of the newer UK universities, being originally established (at that time as an Institute of Higher Education) in 1976. It has around 9000 full time undergraduate students, and 14,500 students in total including post-graduate, research and part-time degrees. It recruits nationally and internationally. There are six academic Schools in the University and they offer a variety of vocationally relevant undergraduate and postgraduate programs, of which many were the first of their kind in the UK. The aim of these courses is not only to provide students with the knowledge and skills concerned with the subject matter of the courses, but also to instill in them the qualities of professionalism and enterprise which will lead to a successful career in whatever field they subsequently choose. The Business School includes, inter alia, Bachelors programs in Business Studies and in Business Information Systems Management, Masters programs in International Marketing Management and Information Systems Management, and an MBA, as well as MPhil/PhD and DBA (Doctor of Business Administration) research degrees. The School places strong emphasis on the development of the individual and on their ability to contribute within the management of an organization.

Simulation has formed part of teaching within Decision Support type courses for many years, being featured across Business Studies, Logistics, Information Systems, and Decision Management under-graduate or post-graduate degrees. Such courses have concerned the use of models in support of decision-making, including structure, formulation, identification, application and implementation. In most programs, the courses have been split into two parts. Part A of the course has covered general principles and a selection of modeling approaches; Part B explored one modeling approach, namely simulation, in some depth including model construction, methodology, software support, and the contribution to understanding and insight in problem-solving

situations. The software used for teaching system simulation in the University was originally Simscript II.5 (Kiviat et al. 1973) and SimFactory (Kleine 1986). For this course, teaching initially adopted System Dynamics as the exemplar, using STELLA II (Richmond 1994) but switched Discrete-event Simulation some eight years ago. The standard tool has been Micro Saint (Barnes and Laughrey 1997), its presentation of model logic separately from animated graphics being helpful in teaching. Alongside Micro Saint use, the simulation sub-course has made use of a purpose-written workbook (in turn identifying further reading), engaging the student in managing their own learning.

3 ASSESSMENT APPROACH

Given the nature of system simulation, assessment must address a range of characteristics. Students need to learn a simulation package's syntax, but major on its problem-oriented application; they need to know the basis of pseudo-random number generation, but understand the implications in real-world studies; they must be able to build/deploy appropriate information collection processes, but use these in methodical experiments; and so on. This leads to reflection, inevitably, on Bloom's Taxonomy for educational objectives (Bloom 1956). Another consideration is the creative dimension of model building - what Tocher (1963) described as the *art* of simulation.

Following the 1948 Convention of the American Psychological Association, Benjamin Bloom lead the formulation of a classification of "the goals of the educational process" and, with his co-workers, established a hierarchy which sought to classify cognitive subdivisions ranging from the simplest behaviour to the most complex (Figure 1).
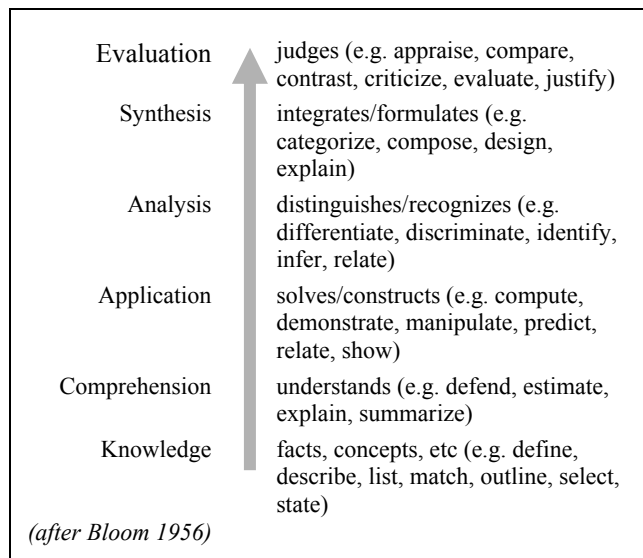
Evaluation	judges (e.g. appraise, compare, contrast, criticize, evaluate, justify)
Synthesis	integrates/formulates (e.g. categorize, compose, design, explain)
Analysis	distinguishes/recognizes (e.g. differentiate, discriminate, identify, infer, relate)
Application	solves/constructs (e.g. compute, demonstrate, manipulate, predict, relate, show)
Comprehension	understands (e.g. defend, estimate, explain, summarize)
Knowledge	facts, concepts, etc (e.g. define, describe, list, match, outline, select, state)

(after Bloom 1956)

Figure 1: Bloom's Taxonomy

With this in mind, an attractive notion to support simulation learning assessment in a course is by a case study (or "mini-project", as it is locally termed to distinguish it from real-world based case studies and research projects also used within programs). Such a mini-project challenges the students at the many levels of the above taxonomy – from necessary knowledge, through informed application and analysis/synthesis, to evaluation/judgement. It can also expose their ability to think creatively in modeling.

The mini-project is therefore conceived in order to test key components of discrete-event simulation in a simulated-study form. The assessment is handed out immediately at the end of the course's formal teaching (which is based on lectures, PC-laboratory work and student-managed learning). Noting Ramsden's (1992) argument that assessment defines the curriculum for students, the assessment is launched with a briefing regarding the goals of the assessment, relating them to the course curriculum. Care is taken to ensure as far as reasonable that the students properly understand the scenario, including a launch briefing and use of an Intranet bulletin-board (part of a Bournemouth University e-learning environment) to facilitate on-going questions as the students work on the assignment. Students are also specifically reminded of Bloom's Taxonomy in the context of assessment expectations. As Brown and Knight (1995) observe, assessment techniques work best when learning outcomes are articulated in advance and, as Biggs (1995) reflects, assessment criteria should be constructed so that students taking a deep approach to learning (Ramsden 1992) are rewarded.

The earliest versions of this assessment approach adopted marking based on specific questions concerning the model/results. However, the weaknesses of that process in evaluating the higher levels of Bloom's structure were rapidly evident. This was replaced by an individual oral examination (viva voce, or simply *viva*) approach, in which each student was subject to examination for 20-30 minutes based on their work as a series of artefacts (i.e. models), demonstrated and discussed. This proved very effective and has remained the practice ever since. The oral examinations are conducted by the course tutor (the Author) with, normally, an academic colleague also present acting as observer and cross-questioner. Since the principal questioning in the oral examinations is by the course tutor, an authorised alternative is to use one assessor only but audio or video record each viva for review by a colleague (second marker) and monitoring by the course's external examiner(s) should they wish. Barton et al. (1994) note the success of team-based oral examinations of case study projects in a graduate course on simulation in engineering design.

Individual oral examinations would be an expensive process were the student numbers other than modest in this course, which is in practice an option for most of the students. An unseen written examination format, based on the

same mini-project, has been considered but would require formal modification of the Course specification through the University's academic quality process – and, with the effectiveness of the oral examination approach, it has not been pursued. As Habeshaw, Gibbs and Habeshaw (1993) note, "in even a short viva it is possible to gain a rich impression of the candidate".

The project detail has evolved over time (the first versions took the form of a story rather than a problem statement). The most recent text for the mini-project, as issued in the assignment brief to the students (Hollocks 2005), follows verbatim as section 4.

4   THE MINI-PROJECT

The scenario is fictitious and any resemblance of company, data, or circumstances to any organization past or present is entirely coincidental.

4.1 Talbot Heath Computers

Talbot Heath Computers (THC) has recently acquired a business which owns a small factory manufacturing very-specialized computer-based measuring equipment (known in the company as "sensors", and selling for £200 each). The Sales team believes that there is the potential for greater market share for their product. However, the new Operations manager, brought in by THC, sees the central sensor production activity, the Unit Assembly Department (UAD), as disorganized and a constraint on the overall production performance of the factory.

THC have commissioned a simulation-based study of the Unit Assembly Department to contribute information to discussions and decisions regarding the business problem of finding, from the possible options, the most effective way of increasing production output.

The whole factory works continuously for five days per week, Monday 0600 until Saturday 0600, then closes for the weekend. The labor and equipment is used in such a way that there are no interruptions in activity at shift change, lunch periods, or coffee/tea breaks (involving stand-by staff brought in, for those short periods, from elsewhere). Maintenance and equipment adjustment are carried out over the weekend (i.e. Saturday 0600 until Monday 0600) and do not prevent or disturb any department or item of equipment from continuing exactly where it left off at the end of the previous week.

The Unit Assembly Department receives pre-sorted kits of sensor parts (valued at £100) directly from the Warehouse (where the kits are put together) at regular intervals of 9 minutes (this time is governed by the design and resourcing of the Warehouse). Individual kits contain all the components required to make one sensor and, when they arrive in the UAD, the kits are stacked in an area of the Department designated as the Input Storage Point. This

is the boundary from which the UAD takes responsibility.

There are currently three Assembly Stations, each equipped with a complex special-purpose device (costing £100,000) designed to facilitate sensor assembly from the kits, and each such station requires one Operator to function. Analysis of shop-floor data indicates that the time for the Assembly task can be represented by a normal distribution with a mean of 30 minutes and standard deviation of 5 minutes (this includes an Operator taking a kit from the Input Storage Point and placing the assembled sensor on a roller table, known as the Holding Point, to wait for the next operation).

After Assembly, sensors are taken by a Technician (in order of Assembly completion) from the Holding Point to an Inspection Area for an activity known as Proving. This is carried out on one of series of purpose-built Test Rigs (costing £20,000 each), of which there are four. There are currently two Technicians. Overall, Proving consists of three steps:

1. a Technician setting up the sensor on a Test Rig in the Inspection Area ready for the Proving process and initiating that process by switching on the Test Rig (setting up and initiating consistently takes 4 minutes, including transfer from the Holding Point);
2. the Proving process itself (which must last for a minimum of 30 minutes, but which does not require a Technician present once started);
3. a Technician evaluating the results of the Proving process (the evaluation time being a constant 6 minutes), without moving the sensor from the Test Rig.

Assembly is a delicate task and, after that process, only 80% of sensors are placed directly on an Output Conveyor to go to Packing & Dispatch, the next department, taking 2 minutes. (UAD's responsibility for each sensor ends when it is on the Output Conveyor.) Of the remainder, 2% of sensors are scrapped after Proving and placed by the Technician in a Rejects Chute (taking them out of the Department and the process, with no scrap value), again taking 2 minutes. A further 18% fail the Proving but are considered reclaimable and must be subjected to a further process in the Inspection Area, known as Diagnosis, to confirm the rectification action to be taken. The 2% and 18% failure rates are random and exhibit no trends or patterns.

Diagnosis is carried out by the Technician immediately following evaluation – with the sensor still on the Test Rig - and taking a time which is (negative) exponentially distributed with a mean of 7 minutes. By the end of Diagnosis the sensor has been removed from the rig.

The concluding step of Diagnosis (and taking a time so short it can be ignored in this analysis) is stacking the

sensor by the Assembly Stations ready to be Rectified by an Operator at one of those stations.

Rectification is not the same activity as Assembly. The time for Rectification follows a normal distribution with a mean of 45 minutes and standard deviation of 8 minutes. Current practice is for Rectification of a defective unit to be given priority over Assembly of a new unit. Rectified sensors are put at the same place as newly-assembled sensors, i.e. the Holding Point, ready for a further Proving process.

Scope for expansion on the site is very limited, with space for two further Assembly stations, or two further Test rigs, or one of each. All employees are paid at the same rate of £400 per week. THC has no problem borrowing capital for investment, but charges it to subsidiaries, such as this one, at 10% p.a.

4.2 The Requirement

You are required to study the above scenario and address the following:

4.2.1 Part 1

Construct a Micro Saint model of the Unit Assembly Department, incorporating appropriate information collection, as a basis for the inquiries in Part 2 below. (A graphics mimic diagram is not compulsory.)

4.2.2 Part 2

As if a consultant to THC, use the model to explore the business problem faced, giving attention to the following questions (producing further models and incorporating modifications or enhancements as may, in your judgment, be required).

1. What is your performance assessment of the existing Unit Assembly Department configuration (i.e. current equipment and staffing level), including information on output and bottlenecks?
2. How many Test Rigs are optimally required in the Inspection Area?
3. What difference in throughput would result from a Business Process Re-engineering initiative combining the roles of Operator and Technician?
4. How much might it be worth THC spending on quality improvements in Assembly to reduce the 2%/18% failure rates?
5. What other conclusions do you reach regarding the effectiveness of possible UAD resource changes (equipment or staffing), operating rules, or profitability?

4.2.3 Part 3

Produce a written report of no more than 2000 words (not counting any title page, contents list, tables, diagrams or graphs - but the word count must be shown); the report to include:

1. the model's Network Diagram and summaries of Task, Job Queue and Decision-point Descriptions;
2. descriptions of the runs carried out, including your rationale for the selection and design of the runs used;
3. results and discussion concerning Part 2 above.

4.3 Assessment

You must submit the report, enclosing with it a disk or disks (diskette or CD) containing your basic Micro Saint model and any variations used for Part 2 above, plus any further variation(s) you may wish to include. The Project will be assessed primarily through an individual oral examination in which you will be required to demonstrate your model(s) and answer questions regarding its/their operation and regarding your report. This oral examination (or *viva*) will use the model(s)/disc(s) and report submitted as its basis. Consequently, failure to submit a report and model(s)/disc(s) will mean that the viva cannot be carried out and hence mean failure of the Project (with zero marks).

The broad marking allocation for the Project (from 100 marks) is: Part 1 = 20 marks, Part 2 = 70 marks, Part 3 = 10 marks .

A key factor in assessment will be the extent to which you show a grasp of simulation and modeling and, through its use, show insight into the system, problems and issues being investigated. Assessment of the report, of itself, is limited to its meeting the requirements of Part 3 above, and to its clarity. No separate mark will be published for the report.

5 IMPLEMENTATION

The above assignment text is followed by details adding the report hand-in date/time, a warning that it is an individual not a group Project (referring to the University regulations regarding plagiarism), and the arrangements for the oral examinations.

The inclusion of specific cost figures is a recent addition. Previously, students had been expected to discuss value more abstractly - in terms of ratios of performance improvement to system change.

Care is taken to manage the oral examination process carefully, in terms of briefing the cohort on the process beforehand, putting each student at ease within their examination (this is not a stress test!) and following a standard

framework – subject however to paths of cross-questioning that depend on the submission and the student responses.

As a schedule progresses, there has been, at times, some "leakage" of information evident between students regarding the detailed viva content. However, this has not proved a problem to the assessment process, either in the performance observed or the grade attained.

6 GRADING

Reflecting the detail necessary to the elements of assessment discussed earlier and seeking to establish a systematic basis, a pro forma was devised providing a relatively comprehensive check list of aspects.

The pro forma is used by each assessor within the oral examinations and a final grade/mark agreed immediately following each student's session. The balance of criteria and grading is shown in Figure 2, extracted (and summarized) from the form.

Part 1 encompasses the "knowledge" and "comprehension" of Bloom's taxonomy and provides a basis for "application". Part 2 continues "application" in sections (b) to (e), with (a) to (e) requiring both "analysis" and "synthesis" leading to "evaluation" as the student seeks to draw conclusions and propose judgments/courses of action.

The Part 3 report is, to an extent, a means of exploring the students ability to articulate their outcomes and process from Parts 1 and 2 without the stress of the viva voce environment. However, its key role is as a platform for that oral examination.

7 EXPERIENCE

Biggs' structure of the observed learning outcome (SOLO) taxonomy (Biggs and Collis 1982) suggests a classification of the (simulation) learner's performance to assist interpret the outcome. It identifies:

- *Pre-structural* responses - involving irrelevant information, not meaningful
- *Uni-structural* responses - where submissions focus on one relevant aspect only
- *Multi-structural* responses - where several relevant features are covered, but are uncoordinated - treated independently and additively
- *Relational* responses - presenting a coherent whole, with details linked to conclusions and meaning evidently understood
- *Extended-abstract* responses - where conceptualization at a higher level is shown, generalizing to a new perspective.

SECTION	CONTENT *General criteria: grasp of modeling and, through it, insight into the system and issues.*	MAX. MARK
Part 1 Construct a Micro Saint model	Scope, adequacy (inc. Network and Task/Queue/Decision descriptions - in particular Assembly and FinishTest); software use, iinformation collection, testing (verification and validation), graphics (if appropriate).	20
Part 2 General Explore the business problem	Experimentation,(inc. run-in, run-length, runs tried, prn seeds, systematic process).	10
Part 2a Performance of existing configuration	Model use, results (e.g. throughput, queues, utilisation), analysis, understanding, conclusions.	10
Part 2b Test rigs required	Model use, changes, results, analysis, understanding, conclusions.	10
Part 2c Combining Operator and Technician roles	Model use, changes, results (e.g. throughput, queues, utilisation), analysis, understanding,, conclusions.	10
Part 2d Reducing 2/18% failure rates	Comparison of output against failure rates, recognition of other constraints, financial assessment.	10
Part 2e Other conclusions?	Model use, changes, results (e.g. throughput, queues, utilisation), analysis, understanding, conclusions.	10
Bonus?	For particularly commendable work in any section.	10
Part 3 Report (2000 words, exc figures) plus disk/CD	Network diagram(s), description of the runs carried out and rationale for run selection, results and discussion, clarity and presentation.	10
Total		100

Figure 2: Grading Framework

Pre-structural responses in the context of this case have been typified by discussion of general issues such as social behavior, not the point of the study nor material within the dynamics of the problem set – and certainly not reflected in the students submitted model! Uni-structural responses have included (indeed been dominated by) a focus on a single inadequate performance metric, most commonly machine, operator and/or tester utilization. The project's specified Requirement tends to deflect the students from focusing on only one dimension of the problem – several are presented in the sub-questions.

A common Multi-structural response has been to address Parts 2a to 2d, but distinctly and with little or no deeper exploration nor material effort in 2e. Relational responses fulfill the requirements of 2a to 2e, drawing them together, understanding the system operation and drawing valid conclusions. Finally, Extended-abstract responses showed a full and insightful grasp of the problem dynamics, abstracting appropriately, adopting a sound experimental framework and exploring wider factors, such as the relationship of the UAD with the supplier warehouse.

Ramsden (1992) argues that there are no error-free tests, but experience with this approach is reassuring. Lesser students, who may still be able (perhaps with help!) to write a seemingly-competent model and/or report, are unable to bluff for long in a face-to-face discussion of the work. The oral/viva voce examination is particularly strong in identifying the Synthesis/Evaluation levels of Bloom's Taxonomy and differentiating the Multi-structural, Relational and Extended-abstract responses of Biggs' SOLO framework. The process has a good discriminatory record in terms of grades - from poor Fail to high First-class.

Common weaknesses that have occurred in past students' assessment submissions have included many issues familiar across assessment, for example:

- Not reading the question carefully and/or not answering (all) the question
- Not really understanding what they were doing
- Not getting to grips with the question - in particular not using simulation to get insight into the problem(s) presented
- Wasting time on unnecessary tasks
- Misusing the software
- Taking too narrow a view of requirements
- Not exploiting the course content, in particular regarding methodology.

Some of these may, of course, mask other failings, such as starting late or planning work badly. (Each cohort of students is told of this accumulated list of weaknesses in the launch briefing – with limited effect!)

It has proved important to use a grading scheme pro forma, summarizing the requirements in such a way as to support mark allocation in a consistent manner as the viva proceeds – summarized promptly as soon as the student leaves the room. Some moderation may be applied when the whole cohort has been seen, but grading is absolute and not relative. The danger of not following this grading discipline is that there is a long follow-through in mark allocation, making the overall process costly in time.

The choice of software does not seem material to the assessment process, given that it provides a transparent basis for model building, application and, particularly, experimentation. Given up-coming Course redesign, a change in software is being considered, for example to Simul8 (Major 1998).

There is no evidence, for example in grade trends, that use of (basically) the same assessment year-on-year has led to any issues in validity, consistency or equitability.

8 CONCLUSION

The paper has described a mini-project, scenario-based approach to simulation assessment utilizing oral examinations for evaluation. It has proved successful in discriminating performance, particularly in addressing the higher levels of Bloom's and Biggs' taxonomies. This includes such aspects as insight and understanding.

The significance of the manpower cost of oral examinations is a function of circumstances (such as student numbers and staff disposition) and the alternatives. It can be a cost worth paying for the relative reliability of the process. The approach also has the advantage of containment - that is, the oral examinations are held within a limited (uninterrupted) time-span and, when over, there is no further marking to be done. It is also popular with the students as it disposes of one assessment prior to their main round of end-of-program written examinations.

The approach may also contribute to the actual learning from the course given the integrative nature of any project application of a technique. Hence it will meet the goals for an assessment method: of being valid, reliable, efficient and beneficial.

ACKNOWLEDGMENTS

REFERENCES

Barnes, C. D., and K. R. Laughrey. 1997. Advanced uses for Micro Saint simulation software. In *Proceedings of the 1997 Winter Simulation Conference*, ed S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson, 680-686, Institute of Electrical and Electronic Engineers. Available via <http://www.in-forms-sim.org/wsc97papers/0680.PDF> (accessed June 27, 2005)

Barton, R. R., L. W. Schruben, J. C. Ford, D. Hopkins, D. Goldsman, and K. J. Healy. 1994. USMED: Broadening the impact of simulation analysis methodology (panel). In *Proceedings of the 1994 Winter Simulation Conference,* ed. J. Tew, M. S. Manivannan, D. A. Sadowski, and A. F. Seil, 1382-1386, Institute of Electrical and Electronic Engineers.

Biggs, J., and K. Collis. 1982. *Evaluating the quality of learning: the SOLO taxonomy.* New York: The Academic Press.

Biggs, J. B. 1995. Assessing for learning: some dimensions underlying new approaches to educational assessment. *Alberta Journal of Educational Research* 41: 1-8.

Bloom, B. S. 1956. *Taxonomy of educational objectives: the classification of educational goals: handbook I, cognitive domain.* New York: David McKay Co Inc.

Brown, S., and S. Knight. 1995. *Assessing learning in higher education.* London: Kogan Page

Habeshaw, S., G. Gibbs, and T. Habeshaw. 1993. *53 interesting ways to assess your students.* Bristol : Technical and Educational Services.

Hollocks, B. W. 2005. *Decision Management (part B) project,* B.Sc. Business Decision Management - year 4, 2004-5. Bournemouth: The Business School, Bournemouth University Institute of Business & Law.

Richmond, B. 1994. *STELLA II: An introduction to systems thinking.* Lebanon NH: isee systems inc.

Kiviat, P., R. Villanueva, and H. Markowitz. 1973. *SIMSCRIPT II.5 programming language.* La Jolla CA: CACI.

Kleine, B. 1986. A SimFactory tutorial. In *Proceedings of the 1986 Winter Simulation Conference*, ed. J. R. Wilson, J. O. Henriksen, and S. D. Roberts, 193-196, Institute of Electrical and Electronic Engineers.

Major, R. L. 1998. Simul8: a visual approach to modeling, simulation and analysis. *OR/MS Today* 25: October. Available via <http://www.lionhrtpub.com/orms/orms-10-98/swr.html> (accessed on June 27, 2005).

Ramsden, P. 1992. *Learning to teach in higher education.* London, New York: Routledge.

Tocher, K.D. 1963. *The art of simulation*. London: English Universities Press.

AUTHOR BIOGRAPHY

BRIAN HOLLOCKS is Professor of Management Science & Information Systems at Bournemouth University Business School. He is a Chartered (Mechanical) Engineer and holds a Ph.D. in Management Science. His system simulation research interests are primarily in software and methodology. He is a member of the UK Operational Research Society and its Simulation Study Group. His email address is: <bwhollocks@bournemouth.ac.uk>