AN EXAMINATION OF THE EXPERIMENTAL DESIGN PROCESS WITH REGARD TO TOLERANCE REQUIREMENT GENERAL THEORY

David R Moore

De Montfort University, Department of Building Studies, The Gateway, Leicester, LE1 9BH, UK

The paper considers the development process for experimental work intended to provide an initial evaluation of the role of tolerance requirement general theory as a factor in the automated assessment of task difficulty as a component of buildability. The nature of general tolerance requirement theory is discussed. This initial research proposes six individual tolerance requirements, each with a different function within the general theory. Analysis of data from the completed experiments may cause this number to be revised as the hypothesis of a link existing between tolerance requirements and time taken to complete a task is tested.

Keywords: Buildability, buildability attributes, generic task theory, requirement theory, task difficulty, tolerance .

INTRODUCTION

This paper is divided into the following sections: introduction; discussion of what is to be tested experimentally; method of testing; constraints on the experimental process; the nature of the product which may result from test data; conclusions; acknowledgements. A glossary of terms with a specific meaning within this paper is provided after the conclusions. Proposals by the author regarding the automated assessment of task difficulty as a component of buildability have been covered elsewhere [Moore and Tunnicliffe (1994), Moore, Tunnicliffe (1995), Moore (1996)].

FERGUSON'S HIERARCHY	**MOORE'S SUGGESTED BUILDABILITY ATTRIBUTES**
1. Assembly impossible	Closed insertion / Tolerance Requirements / Sequence / Access / Interfacing / Range
2. Assembly only possible with extreme difficulty	Sequence / Access / Tolerance Requirements (TR) / Interfacing /Range
3. Assembly possible but difficult	TR / Interfacing / Access / Range
4. Assembly straight forward but perverse	Range / Interfacing / TR
5. Assembly easy	TR

Figure 1. A possible relationship between Ferguson's hierarchy (summarised from Ferguson's (1989) and Moore's (1997) buildability attributes).

Moore, D R (1997) An examination of the experimental design process with regard to tolerance requirement general theory. *In:* Stephenson, P (Ed.), *13th Annual ARCOM Conference*, 15-17 September 1997, King's College, Cambridge. Association of Researchers in Construction Management, Vol. 1, 154-63.

INVESTIGATION OF TOLERANCE REQUIREMENTS GENERAL THEORY

Moore (1997) suggests that there are a number of individual attributes to be considered within the concept of buildability (Figure 1), with each attribute having specific properties, as stated in Figure 2, within the author's proposed strategy for the assessment of buildability. The buildability attribute of tolerance requirements (TR) in particular is suggested as being an important factor within the proposed assessment process, given the fact that it is involved in Moore's (1997) proposals for the describing of each of Ferguson's hierarchy levels (Figure 1). Because of the TR attribute's seemingly important role, the emphasis within this initial experimental work has been placed upon investigating its significance.

ATTRIBUTE	PROPERTIES
1. Tolerance Requirements	Spatial rules governing completion of each high level task (HLT) eg. bricklaying.
2. Range	Number of different HLT / Times each occurs
3. Interfacing	Fixing requirements at each change of HLT
4. Sequence	Order of HLT completion / Installation precedence
5. Access	Space available to HLT / Space required by HLT
6. Closed insertion	Installation precedence

Figure 2. Suggested properties of individual buildability attributes [Moore (1997)].

THE NATURE OF TOLERANCE REQUIREMENTS.

General Tolerance Requirement theory is not yet fully developed. However, it is important to establish at this stage that the word "tolerance" within this work is not used to infer a plus-or-minus value within the production process. The research to date does not consider there to be any margins of error when placing a construction component or subassembly: it has an ideal location in three dimensional space which relates to the locations of other components or subassemblies. It is this relationship which tolerance requirement general theory considers, and an example of some aspects of this relationship is given as Figure 3, in which various tolerance requirements (TR) for a section of brickwork are illustrated. These aspects are discussed further in the section dealing with the TR calculation algorithm.

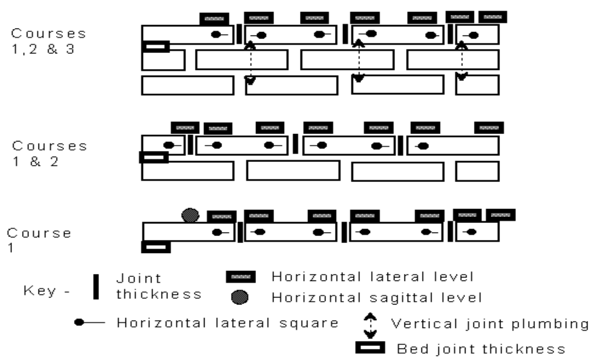

Figure 3. Examples of suggested Tolerance Requirements: brickwork panel [Moore(1997)].

At this stage, the key points of TR theory can be identified in terms of spatial rules which govern the completion of individual high level tasks (HLT) [Chandrasekaran (1988)]. Whilst these rules will vary with each HLT, Moore (1997) suggests there are four possible common production situations which spatial rules will have to deal with: accommodating discrete components (or subassemblies) within; a linear production space defined in three dimensions (LIN); a non-linear production space (changes of direction - COD); a confined production space defined in terms of adjacent production (CND); an isolated production space within a larger production space (ISO). The experimental method is intended to test the relevance of these situations to the process of assessing the on-site task difficulty inherent in producing a designed construction artefact. The nature and extent of any relevance for these situations is seen as being an important consideration in the development of an algorithm for use in the automated calculation of the number of each TR type within a given design.

EXPERIMENT FACTORS AND CONTROLS

Various study techniques are possible within the experimental paradigm, each having their particular advantages and disadvantages (Figure 4). The experimental technique allows both high factor control and level of measurement detail; advantages which cannot be achieved by any other recognised technique. Given the need to achieve a high level of measurement detail within this stage of the research, the experimental technique was selected as being the most viable choice of approach.

	CRITERIA			
TECHNIQUE	**Reactivity**	**Face Validity**	**Control**	**Measure-ment Detail**
Task Analysis	zero	high	-	-
Observation/ Record	low	high	zero	low
Questionnaire/ Ratings	medium	medium	low	medium
Experiments	high	low	high	high

Figure 4. Comparison of experiment techniques by different criteria [Wilson and Corlett (1995)]

Adoption of the experimental technique requires the identification of a suitable dependant variable to be measured within the experiment, along with all the factors which could affect that variable. The dependent variable selected for the proposed experimental work was that of time, selected on the basis of the following:

1. validity; experimental work to support construct validity.
2. reliability; using split-half reliability generally ≥ 0.8 is desirable.
3. sensitivity; reacts sufficiently well to changes in independent variable to allow easy measurement. [Kerlinger (1986)]

These criteria follow on from the basic nature of the construct being put forward for testing: as the assessed level of difficulty within a given artefact design increases, the production time for that artefact will also increase. Factors which could cause time taken to vary within the context of an onsite HLT relevant to the production process for a given artefact can be categorised under the four headings of task, operator, tools, and environment. Factors identified under each of these headings, and the response

('R') deemed most appropriate to each within the proposed experimental work, are given in Figure 5. A positive approach to response selection is to be adopted wherever possible; a factor should only be ignored if there is no other appropriate response.

TASK	**'R'**	**OPERATOR**	**'R'**	**TOOLS**	**'R'**	**ENVIRON-MENT**	**'R'**
Materials	3	Age	1	Defective ?	4	Noise	4
Method	4	Gender	3			Light	3
Size	3	Experience	1			Heat	3
Shape	3	Dexterity	1			Space	3
Location	3	Visual Acuity	3				
		Health	3				
		Size	3				
		Strength	5				
		Mental State*	5				
Response Key	1 2	Build in at multiple levels. Treat as co-variate.	3 4	Fix at a single level. Randomise	5 5* n/a	Ignore (if trauma free.) Not applicable	

Figure 5. Factors affecting dependent variable and control response 'R' [Moore (1997)].

The selection of the measure to be used also considered the speed-accuracy trade-off (SATO) phenomenon, which occurs in resource limited tasks (the more time of effort expended, the more accurate the results) such as the HLT of brickwork. Three alternative response to SATO with regard to the implementation of the measurement scheme exist:

1. Fix speed and measure accuracy.
2. Fix accuracy and measure speed.
3. Let speed and accuracy be chosen by the operator and sort out the effects during analysis. [Wilson, Corlett (1995)]

The approach adopted was to fix accuracy (within standard National Vocational Qualification criteria) and measure speed (time taken to complete production). Perfect accuracy is not required within the context of a study such as this, as it leads into diminishing marginal returns to scale. Two experiments were then developed.

TOLERANCE REQUIREMENTS: EXPERIMENT NO. 1.

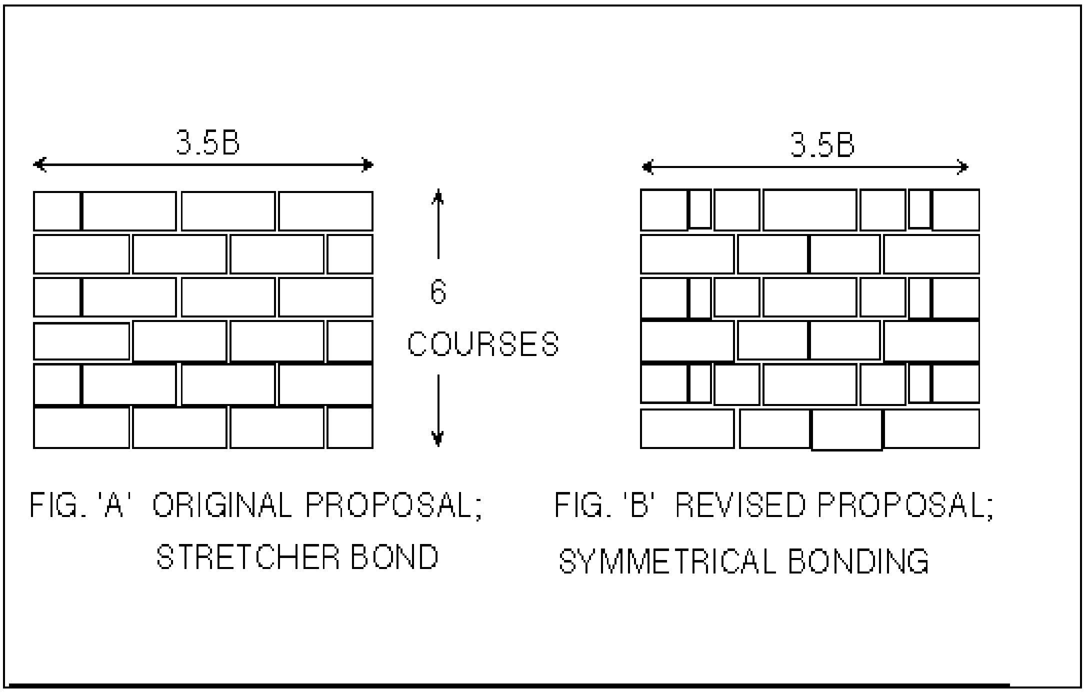

Figure 6. Proposed brickwork panels for use in experiment no. 1 [Moore, Tunnicliffe (1994)].

BENEFIT	NATURE OF BENEFIT
Constant overall dimensions	Panels of equal dimensions, shape and overall area (perimeter X, Y, Z co-ordinates are equal). No variation in operative work performance resulting from differing dimensions etc.
Elimination of technical aids other than level and tape	Panels are not sufficiently large to allow operatives the opportunity of setting up corners and using a line. Tolerances have to be judged initially through the operatives ability to process visual data such as the bed thickness on each course.
One source of varying TR.	Panels only vary with respect to bonding requirements. These bonding requirements are the only source of varying tolerance requirements (TR).

Figure 7. Suggested benefits of using designs illustrated in Figure 6 [Moore (1997)].

Experiment 1 utilised the brickwork panels illustrated in Figure 6, which are suggested as having the benefits listed in Figure 7 regarding isolation of effects which could be expected to result from varying TR within the situation of a linear production space. A possible disbenefit of the panels was identified as being that their similarities could result in a student gaining sufficient relevant expertise on panel 'A' to improve their performance on panel 'B' (positive knowledge transfer [Poulton (1974)]). The possible existence within the experiments of positive knowledge transfer is largely countered by consideration of the following points, along with consideration of Figure 8:

(i) students selected for the experiment were of an expertise level which should allow them to complete panel 'A' without any 'new' learning taking place;

(ii) Panel 'B' is suggested as representing a manual tracking task of greater complexity than panel 'A' due to the greater number of x,y,z co-ordinates occurring within the perimeter of the panel.

	NUMBER OF INTERNAL X, Y, Z CO-ORDINATES.
PANEL 'A'	72. (3 joints / course, 4 co-ordinates / joint)
PANEL 'B'	108 (50% greater than panel 'A': 3 courses having 6 joints/course)

Figure 8. Comparison of panels 'A' and 'B' X, Y, Z co-ordinates [Moore (1997)].

TOLERANCE REQUIREMENTS: EXPERIMENT NO. 2

Panels 'A' and 'B' in experiment no. 1 are intended to examine only one of the possible production situations suggested previously by the author: a linear production space defined in three dimensions. Experiment no. 2 was devised so as to examine the remaining three possible production situations. Examination of Figure 9 reveals that the model to be used for experiment 2 contains examples of the change of direction (COD), isolated (ISO), and confined (CND) production situations.

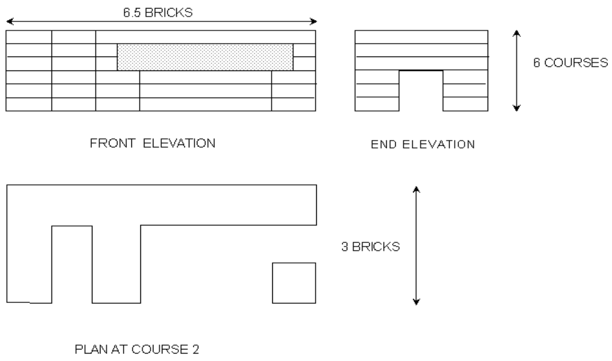

Figure 9. Brickwork model used in experiment no. 2 [Moore (1997)].

All models used within this experimental work were developed in consultation with experienced brickwork staff at Southfields College (Leicester), and are therefore similar to existing models used within the National Vocational Qualification (NVQ) assessment framework. However, it is worth noting that construction work is invariably prototype in nature, and that any task will represent, at some level, the opportunity to further develop expertise. This is not seen as detrimental to the experiments, given that the intention within this research is to examine production situations characteristic of construction work generally, and the controlled nature of the experiments.

INTENDED OPERATION OF THE EXPERIMENTS

Experiment no. 1 requires each of two participants to individually complete panel 'A' and then, after a short break, complete panel 'B'. The participants will be bricklaying students. Before commencing the experiment the students will be asked to complete a short questionnaire designed to produce information on the extent of their bricklaying experience, their perception of each panel's difficulty, and their intended approach (strategy) to the completion of each panel. This information is of relevance in identifying any difference of strategy and perception of difficulty with changing levels of experience/expertise. Each panel will be assessed against typical NVQ level 2 tolerances by the participating students' brickwork lecturer. After completing each of the panels, participants will be given the opportunity to reassess their perception of panel difficulty. Video equipment will used to record the participant's work on both panels 'A' and 'B', combined with the work performance of each student being timed, using a time-study stopwatch.

The model designed for experiment no. 2, given its greater complexity will not be attempted by the participants in experiment no. 1, but will be attempted by a participant of greater experience. In order to provide a common basis for comparison between the participants, the participant undertaking experiment 2 will also attempt panel 'B' prior to attempting the more complex model, and will be asked to complete the standard questionnaire before and after each attempt. Timing and video-recording will be as in experiment no. 1.

CONSTRAINTS

Certain resource constraints had to be considered in the design of the experiments. The most significant was that of limited participant availability, with the possible subsequent criticism of data resulting from the experiments with regard to the Hawthorne effect. A particular problem is the supposed need for control groups so as to identify the extent of the Hawthorne effect on those participants who know they are being observed. However, in Mayo's original studies the series of tests involving the use of a control group resulted in both the control group and the experimental group increasing their performance to a similar extent [Bailey (1982)]. There are three key considerations with regard to the nature of the Hawthorne effect in general:

i) it can occur in any experimental work.

ii) invariably experimental performance is better than real world performance.

iii) it does not mean that the data has no real world significance for productivity. [Fox (1971)]

The Hawthorne effect is particularly important when tests are being carried out to evaluate job design decisions, which are rarely based upon comparative data. The intention within the proposed experiments is to produce comparative data, and the experiment is therefore effectively an initial study in a new area of research. For this reason alone the significance of the Hawthorne effect decreases [Bailey (1982)]. A further consideration is that within this research the objective is <u>not</u> the production of a better working method for the operative. The objective is to seek to provide a mechanism whereby design process workers can be alerted when their designs become difficult to construct within the <u>existing</u> method of carrying out construction processes. They can then choose to evaluate alternative designs or carry on with their existing design, forewarned that it will be difficult to produce on site. This research does not therefore concern itself solely and immediately with the productivity of the construction method. Rather it seeks to determine a basis for such a future examination by studying: if any objective relationship between the demands of the 'as drawn' design and the response of the construction process, at the task level, can be identified; the basis and nature of any such relationship. The case can therefore be argued that the Hawthorne effect, which will inevitably occur to some extent during the experiment, is not detrimental to the resulting data.

The above points are particularly relevant within the context of highly controlled experimental work using exact sampling theory; large samples of a population are not required, and sample sizes of less than thirty are applicable [Spiegel (1962)]. Given the problem of participant availability, exact sampling theory was applied to the experiment design. For example, Student's 't' statistic enables determination of a statistically significant association between two variables. In the case of such an association the resulting data can be accepted as reliable, and it would be expected, with a reasonable level of confidence, that the experiment results could be repeated with a larger sample [Bailey (1982)]. Furthermore, if the data gathered from an experiment can be shown not to exhibit autocorrelation, then any omitted explanatory variables (from any regression equation describing the relationship between identified variables) can be deemed to have no significant impact on the robustness of any relationships identified within that data [Lewis-Beck (1993)]. One technique for identifying the existence of autocorrelation is the Durbin-Watson test (for data containing 15 or more observations). The minimum number of observations achievable within the final form of experiments 1 and 2 is 24 (in experiment 1).

Finally, study design techniques can be used to identify the minimum required number of observations to allow robust relationships to be identified. An equation which can be used to verify the minimum value for N is:

$$N = \left[\frac{t(SD)}{(M1-M2)}\right]^2$$

Where t = critical value of t statistic, SD = standard deviation, M1 = mean for data set 1, and M2 = mean for data set 2. This equation is derived from [Wilson, Corlett (1995)]: $t = \frac{(M1 - M2)N^{1/2}}{SD}$ and can be used to calculate the number of observations required when carrying out comparative work on two data sets. However, the equation presents difficulties with regard to establishing the required level of performance in advance of carrying out an initial study, particularly the standard deviation value, as it is unlikely that similar studies will have been carried out previously. In these circumstances many ergonomists use their expertise and guess. The author, not being an ergonomist, accepted the following values (resulting from a general literature search), on the basis of the proposed study being an iterative process of establishing realistic standards within a new area of research: t at a 5% level of significance = ≥ 1.717; M1 - M2 = ≤ 0.01; SD = ≤ 0.04. From these values a value of N = 48 for comparative analysis between two sets of data (24 observations per set) is calculated. An experiment allowing 24 observations would give sufficiently robust results within the context of a tightly controlled study of the presumed relationship between task difficulty and time taken. However, when using exact sampling approaches an unavoidable disadvantage is a lack of generality in the findings, which cannot be taken as indicators of performance outside each of the three experience categories (< 1 year of experience, >1 but <2 years of experience, >2 but <3 years of experience) identified within the proposed experiments. This is suggested as not being a significant disadvantage given the original nature of the research being undertaken.

THE TR CALCULATION ALGORITHM

The primary objective of this stage of the research programme is to gather data for the production of a robust algorithm for the automated calculation of tolerance requirements. Whilst the proposed initial study is confined to the HLT of bricklaying, it is intended that the possibility of producing a generic algorithm for all high level tasks will at some point be tested. An initial algorithm has been produced and will be tested against the study results in order to determine if any link between the number of tolerance requirements and time taken for a given artefact can be identified.

The initial form of the TR calculation algorithm is rudimentary, in that it represents a formalisation of what the author perceives, on the basis of his own experience, as being the tolerance requirements represented by the rules for bricklaying. Reference to Figure 4 illustrates this perception of tolerance requirements: as the quantity of completed brickwork increases, so the number of TR to be considered for subsequent work also increase. In the example given, six types of tolerance requirement are considered, ranging from mortar joint thickness through to vertical joint plumbing. These TR are suggested by the author as being generic in nature, and therefore possibly of relevance to all construction (and possibly also non-construction) HLT. A further aspect of this proposal for TR is that there may prove to be a relationship in which the importance of individual TR vary with each of the four different production situations suggested within this paper. Such a situation would represent a complex paradigm for the production process.

CONCLUSIONS

Research to date has produced an experimental framework to evaluate key aspects of the author's proposals for the assessment of task difficulty as a component of buildability. These aspects relate to the nature of general tolerance requirement theory, particularly the number and function of individual TR. This initial research proposes six individual TR types.

The evaluation of a prototype TR calculation algorithm which is of importance regarding the proposed automated nature of assessment for buildability is proposed. The logic upon which the prototype algorithm is based has to be capable of being encoded in the form of robust rules. With regard to production situations, this paper identifies four possible production situations which are suggested as possibly being generic to all HLT. The relationship between TR and production situations will be examined when data from the completed experiments is available for analysis.

The author suggests that the experimental work proposed within this paper is an innovative approach to a new area of research which may prove to have a value not just to the construction industry. A truly generic means of assessing task difficulty may result from this research and this is worthy of further investigation.

ACKNOWLEDGEMENTS

The author would like to thank the staff of Southfields College Leicester for their involvement in the experimental stage of this research.

GLOSSARY

High level tasks - Tasks representing the processes carried out within a specific specialism represented by its own rules for completion. *Standardisation* - A design philosophy requiring the designed product to be produced from those materials, components and subassemblies remaining after rationalisation has taken place. *Simplification* - The minimisation of complexity within a design or project to that which is essential. *Tolerance requirements* - The defining of a given productive action in terms of predetermined plumb, level, and square quality criteria, expressed in terms of x, y, z criteria.

REFERENCES

Bailey, R.W. (1982) *Human performance engineering*. New Jersey: Prentice Hall Inc.

Chandrasekaran, B. (1988) Generic tasks as building blocks for knowledge based systems: the diagnosis and routine design examples. *The Knowledge Engineering Review*, **3(**3).

Ferguson, I. (1989) *Buildability in practice*. London: Mitchell Publishing.

Fox, J. G. (1971) Ergonomics in production engineering, In: Singleton, W. T., Fox, J.G., Whitfield, D. (Eds) *Measurement of man at work*, London: Taylor and Francis Ltd.

Kerlinger, F.N. (1986) *Foundations of behavioural research, 3rd ed*., New York: Harcourt Brace.

Lewis-Beck, M.S. (Ed.)(1993) *Regression analysis.* **2,** London: Sage Publishing Ltd.

Moore, D. R. (1993) Buildability and skill concept packages: development of a possible design tool. *Building Research and Information*, **21**(2), 117-121.

Moore, D. R. and Tunnicliffe, A. (1994) AGES: an automated design aid (ADA) for improved buildability**.** *11th International Symposium Automation and Robotics In Construction*, Brighton, May 1994.

Moore, D. R. and Tunnicliffe, A. (1995) An automated design aid (ADA) for constructability. *Second ASCE Congress on Computing in Civil Engineering,* Atlanta.

Moore, D. R. (1996) Buildability assessment and the development of an automated design aid for managing the transfer of construction process knowledge. *Engineering Construction and Architectural Management*, **3**(1 & 2), 29-46.

Moore, D.R. (1997) *Task difficulty assessment: a contribution towards improved buildability through simplification*. PhD Thesis, De Montfort University Leicester, January 1997.

Poulton, E.C. (1974) *Tracking skill and manual control*. London: Academic Press.

Spiegel, M.R. (1962) *Theory and problems of statistics*. London: McGraw-Hill Book Company.

Wilson, J.R. and Corlett, E. N. (eds) (1995) *Evaluation of human work*, 2nd. ed., London: Taylor & Francis Ltd.