An Efficient Approach to Understanding and Predicting the Effects of Multiple Task Characteristics on Performance.

Miles Richardson

University of Derby, Kedleston Road, Derby, DE22 1GB, UK.

Correspondence details: Dr Miles Richardson, m.richardson@derby.ac.uk, +44 1332 593056.

Post-review print of: Richardson, M. (2016). An Efficient Approach to Understanding and Predicting the Effects of Multiple Task Characteristics on Performance. *Ergonomics*, (just-accepted), 1-19.

An Efficient Approach to Understanding and Predicting the Effects of Multiple Task Characteristics on Performance.

In ergonomics there is often a need to identify and predict the separate effects of multiple factors on performance. A cost-effective fractional factorial approach to understanding the relationship between task characteristics and task performance is presented. The method has been shown to provide sufficient independent variability to reveal and predict the effects of task characteristics on performance in two domains. The five steps outlined are: selection of performance measure, task characteristic identification, task design for user trials, data collection, regression model development and task characteristic analysis. The approach can be used for furthering knowledge of task performance, theoretical understanding, experimental control and prediction of task performance.

Keywords: methods; product design; performance; orthogonal; fractional-factorial

Practitioner Summary: A cost-effective method to identify and predict the separate effects of multiple factors on performance is presented. The five steps allow a better understanding of task factors during the design process.

1. Introduction

Understanding the interaction between the task and human performance is fundamental to ergonomics and successful design. The task a user performs usually involves multiple, complex and inter-relating factors (e.g. Layer, Karwowski and Furr 2009), from physical design characteristics to information used for decision-making tasks. These multiple task characteristics have to be understood and manipulated to create the best outcomes for the user. This process can depend upon expert knowledge and intuition, rather than objective evidence (Han and Hong 2003). In order to produce the best outcomes there is a need for methods to understand the relationship between the task characteristics and task performance. This requires methods to create independent variability in order to reveal the separate effects of task characteristics on performance. Full-factorial designs soon become unwieldy when there are many factors, yet although a variety of fractional factorial approaches have been used in various domains, they have not proved popular within ergonomics (Naugraiya and Drury 2009). This paper presents a particularly efficient method that can be applied within various contexts.

To elaborate the problem and the issues presented in understanding multiple factors in task scenarios, two differing contexts are considered. These contexts are also used to illustrate the approach and its successful application below. Firstly, the method presented has been used to identify the key task characteristics involved in self-assembly product and construction play difficulty (names deleted to maintain the integrity of the review process). Within this context the design of self-assembly products and the associated instructions, can be improved if there is an understanding of how each design factor increases difficulty for consumers. For example, one might want to know whether a higher number of components leads to greater difficulty than the level of component symmetry, or the variety of components. Considering several factors in a full-factorial design is not practical, therefore the approach outlined below has answered such issues using an efficient fractional-factorial approach, in seven studies with both adults and children. In addition to identifying the task characteristics that cause assembly complexity and impact on performance, the model has consistently been able to predict performance across differing assemblies, construction materials and task characteristic combinations, with ecological correlations between predicted and actual performance reaching r = 0.977 (name deleted to maintain the integrity of the review process). The systematic approach has provided robust evidence that can provide practical guidelines for handling task characteristics for application during design (e.g. name deleted to maintain the integrity of the review process), or providing a basis for new international standards for self-assembly product instructions (COPOLCO, 2012).

The second context used to elaborate the problem and for illustration, moves away from physical design factors, to the need to understand how task information relates to task decision-making. The design and use of nutrition labels and understanding how they are used to rate the healthiness of foodstuffs requires an understanding of how each nutrient is used (name deleted to maintain the integrity of the review process). For example, do consumers consider saturated fat content more than sugars and salt? The method presented enabled the nutrients used by consumers to be identified, with the relationship model identifying the same key nutrients as eye-movement recording, with the two approaches producing similar results for the relative importance of all the nutrients ($r = 0.88$). The systematic approach brought robust and objective evidence to the domain, providing evidence for policy decisions in the US (Anater et al. 2012; Hersey et al. 2011) and support policy change in New Zealand and Australia (e.g. Royal Australasian College of Physicians, undated; Lyon 2012).

Research into understanding the factors involved in a task often studies a single or small number of variables and their various levels, but as the number of factors increases this approach becomes impractical. Despite there often being multiple factors involved in task design issues, fractional factorial designs are uncommon in ergonomics practice, and there has been a variety of approaches in the small number of studies that do appear (Naugraiya and Drury 2009). A likely issue in the poor uptake is the efficiency of the approach. While fractional-factorial methods present a cost-saving through reducing the level of data collection and design points, they can still require investment in time and resources. For example, Camasso and Jagannathan (2001) present a fractional-factorial design that requires 32 design points to be created to study eight factors. The method presented below would require nine. Response Surface Methodology (RSM) has been used in other disciplines and is well suited to virtual and simulation based design approaches as it reduces the number of data-points required markedly. However, there is complexity in the approach, for example it requires multiple studies for variable screening with Han, Williges and Williges (1997) using 75 design points over three experiments. The method presented below would require 17 design points within a single repeated-measures study, as significant reduction in area that can be resource intensive.

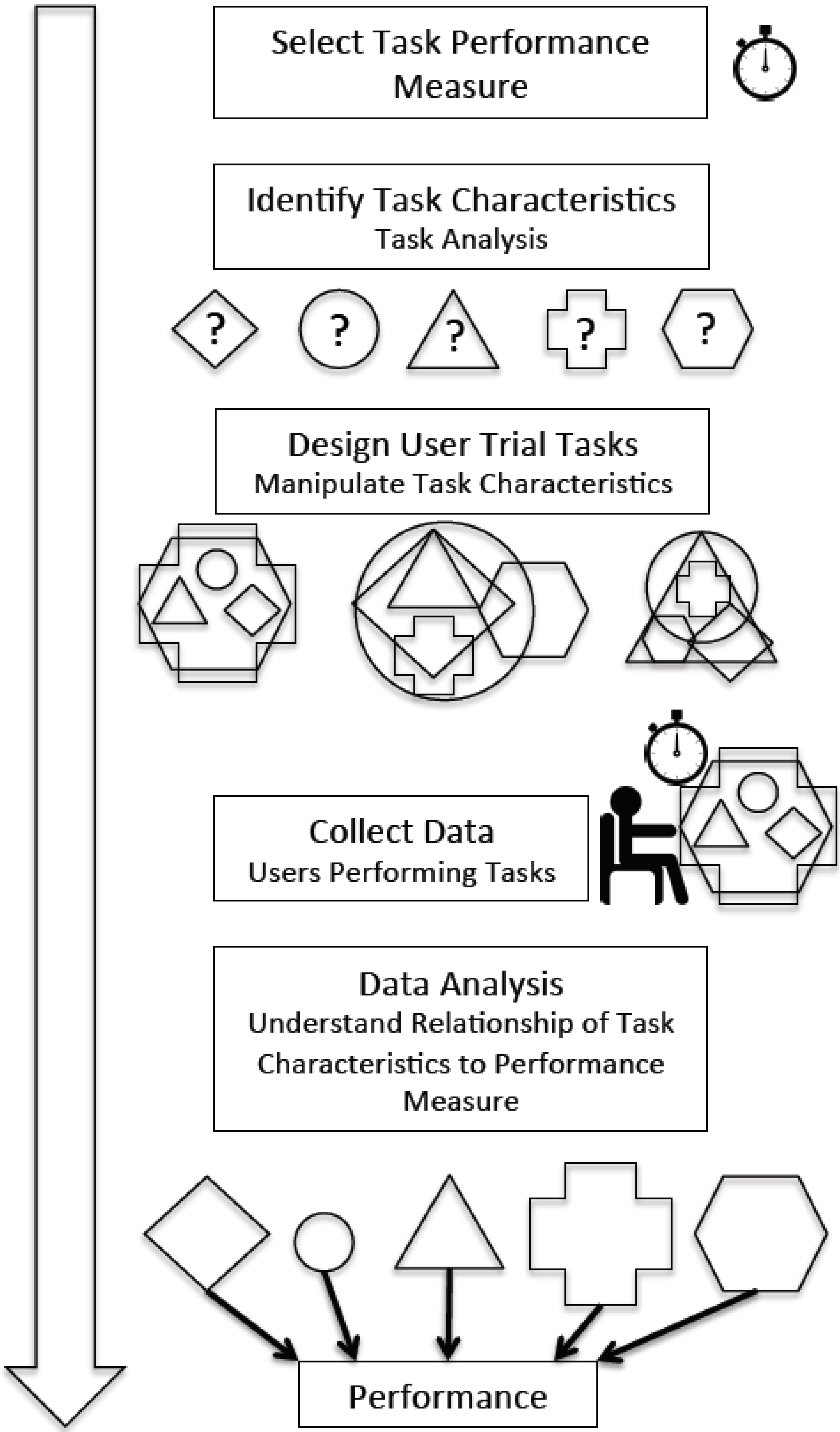
The efficient approach presented involves manipulating identified task characteristics in a balanced, fractional factorial and orthogonal design. The fractional factorial design of all the possible combinations provides sufficient variability to reveal the separate effects of each characteristic. While the orthogonal aspect controls the relationship between the characteristics, providing sufficient independent variability so that the relative importance of the design features is reliable (c.f. names deleted to maintain the integrity of the review process).

2. Orthogonal Relationship Model

The approach is based upon building an orthogonal relationship model in order to analyse the relationship of task characteristics to a performance metric. Broadly similar approaches building relationship models can be found, particularly in affective product design. For example Han and Hong (2003) introduce a systematic approach for coupling user satisfaction with product design. However, the method presented below serves a wider purpose and includes an orthogonal design to control inter-relationships between task variables, and thereby reveal their effects, rather than excluding variables that correlate highly with each other. Orthogonal arrays are important in all areas of human investigation (Hedayat, Sloane and Stufken 1999) and used in experimental design within allied disciplines such as systems engineering (e.g. Huynh 2011).

The orthogonal relationship model consists of the individual task characteristics and a performance metric, for example time or other measure of task success. A successful model will explain the relative importance of the task characteristics, predict performance and help determine how performance might be improved. The approach can inform our quest for fundamental understanding and considerations of use (Karwowski 2005). Therefore, the method presented can be applied in a range of scenarios, from understanding the impact of physical task characteristics on performance, to gaining insight to how a user weights certain task characteristics in decision-making tasks. The method presented and mathematical model also allows the development and control of materials during experimental design, being used to systematically and incrementally increase task difficulty (Nath and Szücs 2014).

Once the research question has been defined, the approach consists of five steps (illustrated in Figure 1): selection of performance measure, identification of task characteristics, task design for user trials, data collection, development of regression model and analysis of task characteristics.


Select Task Performance Measure
Identify Task Characteristics
Task Analysis
?
?
?
?
?
Design User Trial Tasks
Manipulate Task Characteristics
Collect Data
Users Performing Tasks
Data Analysis
Understand Relationship of Task Characteristics to Performance Measure
Performance

3. Selection of the Task Performance Measure

The research aim informs the metric selection and definition, which is an important part in the validity of research and needs careful consideration (Pedhazur and Pedhazur Schmelkin 2013). Performance is a complex phenomenon with many facets and many causes; it has several sources of variation (Pedhazur 1997). The metric could be a measure of task performance, such as task difficulty and complexity, which relate to the qualitative state of the task (Stankov 2000). Difficulty, often measured in terms of time or errors, is a common metric as it refers to a quality of the task and how hard or complex that task is. This is driven by the characteristics of the task that define its inherent complexity, subsequent processing time and the likelihood of errors. This is a likely scenario in the present context, although see Walker et al. (2010) for a wider consideration of complexity in ergonomics.

As performance differences can be related to basic elements of human capability, a wide range of measures could be relevant from objective measures of specific aspects of human performance (e.g. physiological measures) to subjective measures (e.g. mental workload). Tapping into the relationships between human and task characteristics can be used to gain insight into other aspects, such as task cognition.

4. Identification of Task Characteristics

The investigation into the relationship between task characteristics and the chosen performance metric requires carefully selected, and depending on the research aims, theoretically justified task characteristics to be identified. These are operationalised as independent variables in the user trial task design and regression model. In some circumstances the task characteristics will be clear and pre-defined, for example, the nutrients displayed in nutrition labels are fixed. In other cases task description will be required in order to identify task characteristics that relate to task operations and performance. These task characteristics will be hypothesised to affect the chosen metric (dependent variable in the regression model).

Task description suggests the use of Task Analysis (TA) methodologies. The various TA methodologies provide a framework for describing tasks in sufficient detail to identify individual decisions and actions required by the human. Equally TA can be used to identify which details of the task relate to physical characteristics of the task or cognition. Task analysis may be seen as a specific and meticulous method or a less detailed guiding framework for neutral examination of tasks (Shepherd 1998). In the present context the latter is more relevant as the task characteristics identified, and results, will often need to be generalisable and applied to a range of task applications. In order to identify generic task characteristics there is a need for a generic task analysis that describes a variety of tasks within the domain. The purpose of a generic TA is to identify the fundamental steps (task sub-operations) required during a task and provide a guiding framework that can then be used to identify and define task characteristics related to the chosen metric.

For example, during an assembly task a generic sub-operation is to orientate the selected components. Orientation of the component is critical to their positioning, fastening and task completion. There is also theoretical justification for a task characteristic related to component orientation as spatial orientation and mental rotation is cognitively demanding. Once a task sub-operation is identified there is a need to propose a definable task characteristic hypothesised to relate to the chosen metric. In the current example, the process of orientation of the component to allow positioning is affected by the characteristics of the component, for example a symmetrical component can be placed in more than one orientation. Therefore, decisions relating to orientation of components to allow positioning in the assembly are related to the number of symmetrical planes of the component. With a higher level of symmetrical planes fewer rotations are required until the correct orientation is found. Therefore the number of symmetrical planes per component will impact on cognition and task complexity. Within the example of self-assembly tasks, a task characteristic that relates to the sub-operation of orientation can be identified as the level of 'component symmetry', which can be operationalised as the mean number of 'symmetrical planes' within the assembly task.

5. Design Tasks for Data Collection in User Trials

This section explains the method used to develop the materials used during data collection. The method requires the task characteristics to be manipulated in a series of tasks that will enable the relationship between the characteristics and metric to be studied. Existing tasks may be used, but correlation between characteristics is likely to be a problem as there will be no control of the interactions between the task characteristics causing issues during the analysis phase. Further, the task characteristics identified could be distinct for theoretical reasons, but related in actual task designs. An experiment-based approach where the task characteristics are manipulated using bespoke tasks offers greater control and there can be greater confidence in interpreting the regression coefficients.

To establish an accurate model of the relationship between the task characteristics and chosen metric the tasks must be rigorously controlled so that the correlation between the task characteristics is kept to a minimum; the task characteristics levels should be mutually independent or orthogonal.

The cost of this control can be the ability to use realistic tasks, because the tasks created from orthogonal task variables levels will be artificial or abstract. However, once the relationship of the task variables to assembly complexity has been established a return to investigating real-world tasks is possible. Although, the work with self-assembly tasks has shown the key factors found in abstract assemblies to be repeated in real-world assembly tasks (name deleted to maintain the integrity of the review process)

A further disadvantage with this approach is that exercising control of the task variables may distort relations among the variables of interest, this should be considered before applying this approach. The potential benefits, as indicated in the referenced work in nutrition and self-assembly, are identifying the key task characteristics and developing a powerful relationship model.

5.1 Task Characteristic Levels

A balanced fractional factorial design is a common method for defining a subset of stimuli for evaluation as it provides an effective use of resources and the statistical properties of the study are known in advance (Gunst and Mason 2009). The minimum number of stimuli to ensure reliability of the results can be calculated by subtracting the number of task characteristics (TCs) from the total number of levels (L) across all characteristics and adding one (((TCs x L) - TCs) + 1), (Hair et al. 2006).

A set of tasks that systematically vary the levels of each task characteristic are required so that the influence of each task variable can be examined independently of the others. To create these tasks, each task characteristic is assigned a value, either high, medium or low; or two levels can be used (high or low). For example, using two levels for seven task characteristics (2^7) will produce a total of 128 possible combinations. Statistical software (e.g. SPSS) can then be used to produce an orthogonal design based on a random sample of at least the minimum sample for the fractional factorial design. For example, using (((TCs x L) - TCs) + 1), the minimum sample is 8 (((7 x 2) -7) + 1).

As an example, Table 1 shows the task characteristic levels used to design test assembly tasks in previous research. For example, assembly three contained: A low number of component parts; a high level of symmetrical planes, a variety of components to create a high number of novel assemblies, a high number of fastenings, a low level of fastening points and a low number of component groups. Adding unneeded items to the components increased the selections variable.

	Components		Symmetrical Planes		Novel Assemblies		Fastenings		Fastening Points		Component Groups		Selections	
Assembly	Level	Actual Score	Level	Actual Score	Level	Actual Score	Level	Actual Score	Level	Actual Score	Level	Actual Score	Level	Actual Score
1	Hi	9	Lo	1.00	Lo	2	Hi	12	Hi	6.22	Hi	5	Hi	17
2	Hi	9	Hi	2.44	Lo	2	Hi	12	Lo	2.70	Hi	5	Hi	17
3	Lo	5	Hi	2.40	Hi	4	Hi	10	Lo	4.00	Lo	2	Hi	13
4	Lo	5	Hi	2.40	Lo	2	Lo	4	Hi	9.00	Hi	3	Hi	13
5	Lo	4	Lo	1.00	Lo	2	Lo	4	Lo	2.00	Hi	4	Hi	12
6	Hi	8	Lo	1.25	Lo	2	Hi	12	Hi	7.25	Lo	2	Lo	8
7	Lo	5	Lo	0.20	Hi	4	Hi	8	Hi	8.60	Hi	4	Lo	5
8	Lo	5	Lo	1.00	Lo	1	Lo	4	Lo	2.40	Lo	1	Lo	5
9	Hi	8	Hi	2.75	Hi	5	Lo	4	Hi	5.75	Lo	2	Hi	17
10	Lo	5	Lo	1.20	Hi	4	Hi	11	Hi	9.00	Lo	1	Hi	13
11	Lo	5	Hi	1.80	Hi	4	Hi	9	Lo	3.80	Hi	4	Lo	5
12	Lo	4	Hi	2.75	Lo	1	Lo	3	Hi	8.00	Lo	1	Lo	4
13	Hi	9	Lo	0.67	Hi	11	Lo	4	Lo	1.70	Hi	4	Lo	9
14	Hi	10	Hi	2.10	Lo	1	Hi	9	Lo	1.80	Lo	1	Lo	10
15	Hi	9	Hi	2.78	Hi	9	Lo	4	Hi	6.44	Hi	4	Lo	9
16	Hi	9	Lo	0.30	Hi	8	Lo	3	Lo	1.70	Lo	1	Hi	18

Table 1. Task characteristic levels used to design assembly tasks (from name deleted to maintain the integrity of the review process).

To ensure an orthogonal design the task characteristics levels should be modified as follows: calculation of the actual task characteristic levels for each task; analysis of the correlation between task characteristics; modification of tasks ensuring a range of construction task characteristic levels (e.g. high and low based on the mean) and reduction of any high correlations. Ensuring correlations between the task characteristics are low will help avoid multicollinearity causing issues in the subsequent regression analysis. Although higher levels of correlation are deemed acceptable (e.g. r = .80, Field 2009), the work of [name deleted to maintain the integrity of the review process] worked to a maximum correlation of r = .40.

As noted, depending on the domain, the task characteristics can be pre-defined. For example for nutrition label research, [name deleted to maintain the integrity of the review process] assigned values of either high, medium or low for eight nutrients across eighteen orthogonal label designs. The design was balanced so that there were six instances of high, medium and low for each nutrient. Further, to ensure nutrient levels varied across labels, the nutrient values were randomly assigned within a 12.5% band at each of the low/medium/high levels. This helped ensure the labels were realistic while maintaining orthogonality.

6. Data Collection Procedures

This section covers the design of the user trial procedures used for data collection. A repeated measures design where participants complete all the tasks is used. Repeated measures designs reduce unsystematic variability in the design (Field 2009) and control for individual differences, which are probably the largest source of variation in most research studies in ergonomics. Using a repeated methods approach it is possible to identify the variance due to individual differences and this leads to more precise analysis (Pedhazur 1982). Disadvantages of repeated measures designs are practice effects and fatigue, although the randomisation or counter balancing can counteract these issues.

The repeated measures design includes the task characteristics as IVs and the DV is the chosen performance metric, e.g. errors or time. The task designs are then presented in a random order for completion by sufficient participants. A sufficient number of participants being based on guidance (e.g. Green 1991) for the minimum number of observations in multiple linear regression analysis, e.g. (20 x TCs) divided by the number of task designs (owing to the repeated measures design). For example, Table 2 shows examples of the number of required tasks for a fractional factorial design (Hair et al. 2006), minimum observations for the regression analysis (Green 1991) and minimum number of participants given various numbers of task characteristics and associated levels.

TCs	Levels	Tasks	Minimum Observations	Participants
4	2	5	80	16
4	3	9	80	9
8	2	9	160	18
8	3	17	160	9
16	2	17	320	19

Table 2. Examples of the number of experimental tasks, repeated measures observations and participants required.

7. Development of Regression Model and Task Characteristic Analysis

Multiple linear regression is eminently suited for analysing effects of independent variables (task characteristics) on a dependent variable (chosen metric). As introduced above, the choice of variables and approach should be determined by a theoretical framework that explains the nature of the relationship between the variables being studied (Pedhazur 1997).

Further, regression is a widely used technique in various studies to model the relationship between system design factors and human performance (Han and Hong 2003).

Multiple regression can be applied in both predictive and explanatory research. Predictive research is aimed at the development of models to predict criteria of interest by utilising information from one or more predictors. Explanatory research is aimed at the testing of hypotheses formulated to explain the phenomena of interest (Pedhazur and Pedhazur Schmelkin 2013). The distinction between explanation and prediction should be borne in mind when setting out the research question, especially if the results of a predictive study are interpreted as explanations (Pedhazur 1982). For example, theoretically driven explanatory approach may not always produce the best predictive model. However, it can be argued that theory is the best guide to selecting independent variables and explanatory research may serve as the most powerful means for prediction, and Pedhazur and Pedhazur Schmelkin note that "If we can predict successfully on the basis of a certain explanation we have good reason, and perhaps the best sort of reason, for accepting the explanation" (Kaplan 1964, 350).

The success with the approach outlined above has been based upon a first-order regression model which only includes the main effects of each independent variable. Such models are more straightforward to interpret and enable the identification of key task characteristics and their relationship to the performance metric. Good practice with the regression analysis should be observed, for example checking for skewed metric data as measures such as time are often skewed towards zero. A log transformation can be a useful approach to restore a linear relationship. Alternative, and more involved statistical approaches, can also be employed depending on the research question.

As there is multiple data for each participant from the repeated measures design, dummy variables to identify each participant should be created and entered in the first block in order to control for variability due to individual differences (Pedhazur 1982). The task characteristics can then be entered as a set into the second block.

The method is designed to reduce multicollinearity affects, so diagnostic checks for multicollinearity (e.g. variance inflation factors) and cases exerting undue influence should also be performed. Owing to the orthogonal design there should not be a need to exclude model items in order to deal with multicollinearity. The R Square Change figure will indicate the variance in the chosen metric accounted for by the task characteristics, independent of individual differences captured by the dummy variables. The significant predictors of the metric will also be revealed through the standardised regression coefficients for each task characteristic. These can be used to identify the significant, and most important task characteristics and produce a regression equation in order to predict performance.

8. Discussion

Although, a variety of broadly similar, and little used, approaches are available (Naugraiya & Drury, 2009), the method presented offers an efficient method for establishing the relationship between task characteristics and performance using minimal design points. The method has also been shown to be effective in differing contexts: identifying the task characteristics that cause assembly complexity and predict performance; and enabling the nutrients used by foodstuff consumers to be identified, delivering robust and objective evidence that have contributed to case studies rated as having very considerable impact in terms of reach and significance (name deleted to maintain the integrity of the review process). The relationship model allows the prediction of performance and the relative influence of task characteristics to be identified. Such findings inform understanding of tasks and ultimately can inform design. Further, the control of variables is fundamental to any experimental work, and the method presented includes and enables that, with the possibility of creating incremental increases in task complexity for further research or training.

The tightly controlled approach presented is justified by the need for objectivity and to reveal the effects of multiple variables. However, this systematic approach does bring disadvantages. It requires research time and resources, with the task design phase bringing some complexity, although the minimal fractional-factorial design is particularly efficient, while also delivering results comparable to those obtained by resource intensive methodologies such as eye-tracking. Further, the task design aspect can lead to somewhat contrived tasks in the user trials because the tasks created from orthogonal task characteristic levels by nature can appear artificial. The extent of this as an issue depends on the domain of methods application, and it is an issue for the researcher to consider when considering using this approach. It is also the case that once the relationship of the task variables to performance has been established with an orthogonal design, a return to investigating real-world assemblies is possible. An alternative approach for those wishing to forgo the orthogonal design is to employ other forms of analysis, for example principal component analysis; although the results may be more difficult to interpret into clear guidelines.

One reason why the approach to analysis presented is more accessible is it is based on a main effect model. Interaction effects are not accounted for with regression analysis, although interaction terms can be created and included in the analysis. However, the addition of these interaction terms means there is a danger of overfitting the model to chance variations in the data. Therefore, interaction terms should only be added when a theoretical rationale exists. Interaction effects from non-theory based interaction terms are unlikely to occur in other datasets. Further, independent variables in the regression equation can be highly correlated with cross product interaction terms, making it difficult to assess the relative importance of interaction effects and main effects. It should be noted that [name deleted to maintain the integrity of the review process] found results achieved when adding cross product interaction terms were not replicated, suggesting that interaction effects were artifacts of overfitting. The consistency of the main effects model over several studies also reduces the possibility of higher-order interactions and supports the focus on a main effects model, however the model cannot explain the complex relationship between task characteristics and performance completely.

9. Conclusion

The ergonomics approach often requires understanding of the separate effects of multiple factors on performance. Through systematically varying these factors in an orthogonal and fractional factorial design their separate effects on task performance can be better understood. As with any method, before proceeding with its use, there is a need to ensure it is appropriate for the given research question. A key factor being whether it can be applied in the domain area; that is can the task characteristics be manipulated within tasks that users can perform. The strength of the method is the systematic creation of independent variability in order to reveal the separate effects of task characteristics. The repeated measures design provides an efficient use of resources and can be used in both research and design settings. The method creates a relationship model that can improve task knowledge and understanding by identifying key task variables, which can also be utilised as a basis for successful prediction. There are many potential applications for this approach, from understanding the relationship between the physical characteristics of a task and the difficulty of performing that task to inform product design, to understanding how task information relates to decision-making, or used within experimental paradigms for added control and rigour. Further, including a theoretical basis during task characteristics selection can also inform theoretical understanding in the chosen area of application.

References

Anater, A.S., Wohlgenant, K., Cates, S., Hersey, J., Muth, M.K., Zaccaro, D., and Zhen, C. 2012. *Evaluation Planning and Tools for Front of Package Nutrition Labeling: Final Report* (RTI Project Number 0212050.020.000.001). Prepared for the Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, Office of Science and Data Policy, Washington DC.

Camasso, M. J., & Jagannathan, R. 2001. Flying personal planes: modeling the airport choices of general aviation pilots using stated preference methodology. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 43(3): 392-404.

COPOLCO. 2012. *Consumer Priorities in International Standardization Work Priority Programme, Annual Report 2012*. http://www.iso.org/iso/copolco_priority-programme_annual-report_2012.pdf

Field, A. 2009. *Discovering statistics using SPSS*. Sage publications.

Green, S. B. 1991. "How many subjects does it take to do a regression analysis?" *Multivariate Behavioural Research*, 26: 499-510.

Gunst, R. F., and Mason, R. L. 2009. "Fractional factorial design." *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(2): 234-244.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., and Tatham, R. L. 2006. *Multivariate data analysis* (Vol. 6). Upper Saddle River, NJ: Pearson Prentice Hall.

Han, S. H., and Hong, S. W. 2003. "A systematic approach for coupling user satisfaction with product design." *Ergonomics*, 46(13-14): 1441-1461.

Han, S. H., Williges, B. H., & Williges, R. C. 1997. A paradigm for sequential experimentation. *Ergonomics.* 40(7): 737-760.

Hedayat, A. S., Sloane, N. J. A., and Stufken, J. 1999. *Orthogonal arrays: theory and applications*. Springer Science and Business Media.

Hersey, J.E., Wohlgenant, K.C., Kosa, K.M., Arsenault, J.E., and Muth, M.K 2011. *Policy Research for Front of Package Nutrition Labeling: Environmental Scan and Literature Review: Final report (Contract No. HHSP23320095651WC)*. Prepared for the Department of Health and Human Services, Office of Assistant Secretary for Planning and Evaluation, Office of Science and Data Policy, Washington DC. http://aspe.hhs.gov/sp/reports/2011/FOPNutritionLabelingLitRev/

Huynh, T. V. 2011. "Orthogonal array experiment in systems engineering and architecting." *Systems Engineering.* 14(2): 208-222.

Kaplan, A. 1964. *The conduct of inquiry: Methodology for behavioural science*. San

Francisco: Chandler.

Karwowski, W. 2005. Ergonomics and Human Factors: The Paradigms for Science, Engineering, Design, Technology, and Management of Human-Compatible Systems. *Ergonomics*. 48(5): 436 – 463.

Layer, J. K., Karwowski W. and Furr, A. 2009. The Effect of Cognitive Demands And Perceived Quality Of Work Life on Human Performance in Manufacturing Environments. *International Journal of Industrial Ergonomics*. 39(2): 413-421.

Lyon, J. 2012. *Evidence Snapshot: Food labelling*. Agencies for Nutrition Action. http://www.ana.org.nz

Naugraiya, M., & Drury, C. G. 2009. A fractional factorial screening experiment to determine factors affecting discrete part process control. *Theoretical Issues in Ergonomics Science*, 10(1): 1-17.

Nath, S., and Szücs, D. 2014. "Construction play and cognitive skills associated with the development of mathematical abilities in 7-year-old children." *Learning and Instruction.* 32: 73-80.

Pedhazur E.J. 1982. *Multiple Regression in Behavioral Research*. New York, NY: CBS College Publishing.

Pedhazur, E. J. 1997. *Multiple regression in behavioral research*. (3rd Ed.). London: Thomson Learning.

Pedhazur, E. J., and Schmelkin, L. P. 2013. *Measurement, design, and analysis: An integrated approach*. Psychology Press.

Royal Australasian College of Physicians (undated). *Mandatory Front–of-Pack "Traffic Light Labelling" on Food and Beverages A Policy Position Statement by the Royal Australasian College of Physicians*. Royal Australasian College of Physicians. Retrieved May 1, 2015 from http://www.foodlabellingreview.gov.au/internet/foodlabelling/submissions.nsf/lookupSubmissionAttachments/1ATAN-85JVSB20100518094116DMLO/$FILE/448a.pdf

Shepherd, A. 1998. "HTA as a framework for task analysis." *Ergonomics.* 41 (11): 1537-1552.

Stankov, L. 2000. "Complexity, metacognition and fluid intelligence." *Intelligence.* 28 (2): 121–143.

Walker, G. H., Stanton, N. A., Salmon, P. M., Jenkins, D. P., and Rafferty, L. 2010. "Translating concepts of complexity to the field of ergonomics." *Ergonomics.* 53(10):

1175-1186.