QUANTITATIVE COMPUTATIONAL EXPERIMENT (QCE): AN ALTERNATIVE POST-POSITIVIST EXPERIMENTAL RESEARCH STRATEGY OF INQUIRY FOR DESIGN STUDIES

[1]ALI GHAFFARIANHOSEINI, [2]RAHINAH IBRAHIM, [3]JOHN TOOKEY, [4]AMIRHOSEIN GHAFFARIANHOSEINI

[1,3]School of Engineering, Faculty of Design and Creative Technologies, AUT University, Auckland, New Zealand
[2]Faculty of Design and Architecture, University of Putra Malaysia (UPM), Malaysia
[4]Department of Geography, Faculty of Arts and Social Sciences, University of Malaya (UM), Malaysia
E-mail: alighh@aut.ac.nz

Abstract- One of the most important aspects for development of valid results in academic research is selection of the appropriate research methodology. Contemporary research in design studies confronts the rapid expansion of emerging high-tech trends. Correspondingly, an adapted research methodology is required to meet the needs of current circumstances. This paper articulates employment of a quantitative computational research methodology for design studies. The research framework, implementation and validations are expressed in details. Computational charrette test method and computational emulation reasoning and representation are incorporated in order to validate the discussed research methodology outputs. In summary, the developed research methodology is articulated in details to enable further exploitations in academic research and practices.

Index terms- Quantitative Computation Experiment (QCE); Social Science Research; Design Cognition; Research Methodology; Design Studies

I. INTRODUCTION

One of the most important aspects for development of valid results in academic research is selection of the appropriate research methodology [1-6]. This paper elaborates on an alternative method for putting into practice the quantitative research based dubbed as Quantitative Computational Experimentation (QCE) [7]. Researchers have proposed quantitative; qualitative and mixed-method research methods for practical research in various fields [1, 4, 8-15]. Moreover; recent researchers have recommended administration of quantitative analysis together with qualitative approaches [16].

Due to the rapid conquest of computer and technology; quantitative and measurement-based research have become increasingly popular during contemporary studies. This article focuses on the post-positivism approach of quantitative experimental method for design studies.

[1] identifies four main philosophical world views of Post-positivism; Social Constructivism; Advocacy/Participatory and Pragmatism in academic research. As a result, Quantitative; Qualitative and Mixed methods research designs and strategies of inquiry are outlined. Post-positivism is applied towards verification of theories however constructivism is used to generate theories. Worldviews of advocacy/participatory and pragmatism are employed for change-orientated and real-world practice oriented research.

QCE is initially developed for design studies however it can be generalized for use in diverse social science research. Various attempts have been made in order to develop innovative computational tools and artificial intelligent-based expert systems to better understand/perform the design process [17-19]. Generally, QCE is applied in order to advance design computing and cognition [20] core progress by understanding the way researcher/designer thinks/acts [21]. Computing and cognition were studied by various field researchers' such as scholars at the Key Centre of Design Computing and Cognition (KCDCC).The design computing and cognition was initiated at the University of Sydney (KCDCC) [22]. For the last three decades, KCDCC has been a pioneer in this area. Development of innovative architectural software and examining their effectiveness was focused on as the main field of research. Experimental test methods must be reliable and valid [4, 23-26]. In regards to reliability, the experiment is examined whether it can produce the same results if used in different circumstances or through different users. Alternatively, reliability is a question on the objectivity of the experiment results. It can be partially considered the extent that the results can be generalized to. On the other hand, validity discusses if the results indicate the correct measurements that the researcher had intended to evaluate. Congruently, any accredited experimental test method should be both reliable and valid. A good experimental test method should also be practical, cost effective and within the skills of the evaluator [27]. Therefore, QCE is considered as a practical and relatively cost-effective approach towards post-positivist research due to its flexibility.

II. RATIONALIZATION

QCE is commonly used to assess the effectiveness of an innovative approach in design studies. The process is then compared to the conventional methodologies based on a consistently developed computational instrument. QCE is limited to preparing the computational instrumentation according to prototype development protocols [28, 29].

Effectiveness testing is done according to Charrette test method principles. [30-35]. Charrette test method is executed when an innovative procedure is proposed to be used as an alternative substitution of the conventional procedure for the same task [30, 31, 34, 36-38]. In this research software in QCE means the new approach/instrument developed for design studies.

Initially, an effective process must be general. The innovative procedure must be simply usable not only through the researcher but, also through different other typical users of this new software. Subsequently, not only the innovative process should be capable of being used for a certain project but, also for other versatile projects. In other words, the innovative process should be a comprehensive methodology to encompass application of different types of projects within the respective field. This process must be faster in comparison with the conventional process. Furthermore, the innovative process should be capable of being utilized with a higher level of accuracy resulting in higher quality outcomes. Consequently, a Pre-Test-Post-Test experimental methodology [39, 40] is chosen to be done according to a within group experiment basis for general QCE. Consequently, the developed test set-up is required to be done based on the hypothesis of research.

QCE should be followed based on proper research component set-ups. In other words, QCE is developed according to a systematic integration of different research components. [41] has developed the Eagle research design framework to represent the research question constructs; research questions; research objectives; and results of the study. The main reason of employing an Eagle research design framework table is to clearly articulate the composition of research components. QCE research questions must encompass development of a new workflow from integration of two processes to develop a new knowledge contribution. It is recommended to organize the research constituents of a QCE based on [41]'s eagle view research design framework as represented during the Table 1.

Table 1 Eagle View Research Design Framework Adapted from [41] [Sample]

Main Research Question (RQ)			
RQ CONSTRUCT	**DESCRIPTION OF SUB RESEARCH QUESTION (SUB RQ)**	**RESEARCH OUTPUTS**	**KNOWLEDGE CONTRIBUTIONS**
	Towards Research Objectives (RO)		
RQ Construct 1	Sub RQ 1 Research Objective (Ro 1)	Output 1	Knowledge Contribution 1
.....	......	......	......
Last RQ Construct	Last Sub RQ Last Research Objective (Ro X)	Last Output	Last Knowledge Contribution

III. IMPLEMENTATIONS

QCE comprises of four main steps of 1) computational system design and development (instrumentation) [42, 43] 2) computational system verification [44]; 3) assessment of effectiveness (Charrette test) [32, 34] and 4) computational simulation/emulation validation [45-47].

QCE compares an innovative system versus a conventional system. The innovative system is the newly developed technique of carrying out a certain task while the conventional system is the regular method of executing the same task.

1) During instrumentation, the computational system to measure QCE parameters is developed. Training sessions must be conducted to familiarize the participants with the innovative system to reduce the risk of data biasness. This may occur as a result of potential unfamiliarity of the participants with the innovative system. The developed instrument must be compatible with computational system development criteria in accordance with the respective research.

2) The developed instrument must be verified in order to confirm its usability. The confirmation must be in accordance with computational system design and development principles [44, 48].

3) After verification, QCE undertakes the effectiveness testing. Testing is run in pilot tests and main tests stages. During testing, participants will be asked to run a certain set of tasks using the conventional method. Furthermore, they are asked to do the same task using the computational method. Each participant will be provided with the complete to-do task breakdown list. The participants are required to provide minimum three alternative solutions for the proposed problem. In other words, each trial must be run based a minimum of three replicates.

During a QCE; parameters (Variables) of "Speed" and "Accuracy" are measured. These variables are measured for each task of each participant at each replicate for each trial during each process. This is done in order to examine the effectiveness (Higher speed and higher accuracy for the selected trial) of the innovative process. Similarly, accuracy level of each task for each task of each participant at each replicate for each trial during each process is measured. Performance speed is divided into three main parameters: 1) "Finish interpretation time: Thinking time for accomplishment of each individual task per individual participant"; 2) "Finish Critique Time: time for accomplishment of each individual task per individual participant" and 3) "Overall finishing time: Summation of the previous two".

Since similar test cases must be selected during the conventional and the innovative trials, a 14 days delay should be allocated between the two trials. This is to mitigate the risk of biasness for the collected data [30, 38]. Moreover, conducting the innovative process later than the conventional process would affect the reliability level of the collected data. This would happen in case of familiarity of the participants with the completed test case as completed during the conventional trial. As a result, QCE proposes to carry out the innovative process prior to the conventional process ensuring no risk of biasness.

Quantitative experimental statistical tests [7, 25, 26, 49, 50] must be carried out in order to inferentially analyze the collected data. This is to ascertaining the effectiveness of the innovative system.

4) Subsequently, computational simulation/emulation validation must be carried out in order to validate the final findings of the study. The overall progress is visually described in Figure 1.

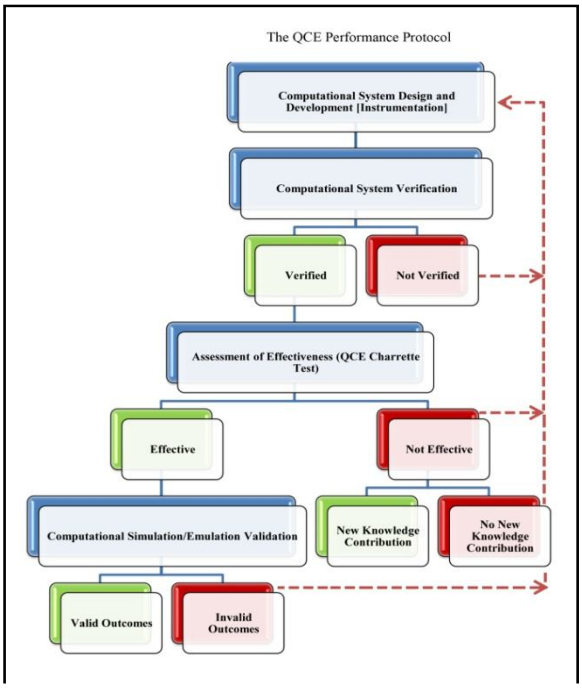

Figure 1 QCE Performance Protocol

Computational System Design and Development [Instrumentation]

Experimental research methods require collection of measurements using instruments [1]. As a result, the researcher must elaborate on the applied instrument e.g. certain software. Issues regarding the validity and reliability of the employed software are essential to be considered. Thus, QCE engages a system design and development procedure [44, 48] in order to prepare the research instrument as the initial step.

According to [44], software development process is a structure planned for the preparation of a software product. Several models have been designed for the overall progress of software development activities. Each individual model encompasses various tasks. One of the most famous models of system development procedure is the waterfall model [44]. According to this approach, all stages of the system development must be incrementally categorized.

One of the most important activities during software development process is prototyping. Prototyping is generally known as the preparation and programming of a scaled down but yet functional version of the main software. Prototypes can be developed with any computer language [44, 48]. [44] proposes development of a prototype for improvement of the eventual software. Software prototyping procedure encompasses; a) identification of the problem; b) development of the prototype; c) implementation of the prototype and eventually d) testing the prototype.

Since various prototyping methodologies are available, QCE recommends the evolutionary prototyping method [48, 51-53]. In this regards, the QCE researcher initially identifies the research problem leading toward feeling the need for preparation of the proof of concept computational software. The researcher may hire a computational programmer in case of extreme complexity. The researcher must hold several discussion sessions with the hired programmer. Different ideas of improvement must be discussed between the researcher and the programmer. This is to form the evolutionary essence of the proof of concept computational model.

IV. COMPUTATIONAL SYSTEM VERIFICATION

Computational system must be verified according to 1) software validation and 2) verification "V&V". There are two approaches to software "V&V": 1) software inspection and 2) software testing [44, 54-58]. Software inspection takes place during all different phases of the software development life-cycle inspecting requirements, documents, design diagrams and program codes. Software testing runs the target software to check if it is produced correctly and as intended. Since introduced in early 1990s, the concept of validation and verification has been interpreted in a number of contexts. The followings are some supporting examples for QCE. All following requirements must be met in order to call the QCE instrument verified.

1) In IEEE standard 1012–1998, validation is described as the process of evaluating a system or component during or at the end of the development process to determine whether it satisfies requirements. Verification is the process of evaluating a system or component to determine whether the products of a given development phase satisfy the conditions imposed at the start of that phase [59]

2) ISO (17025E) describes validation as the confirmation by examination and the provision of objective evidence that the particular requirements for a specific use are fulfilled [60]

3) [61], from the software engineering point of view, defines validation and verification as: "validation: Are we building the right product?" and "verification: Are we building the product right? [61]". Correspondingly, the developed software is verified if the software is functioning properly according to the required principles. The software is validated if it is of importance to its respective field.

4) Software validation in the EE discipline is described by Working Group on Digital Evidence [62] as an evaluation to determine if a tool, technique or procedure functions correctly and as intended [62].

Validation procedure is to determine the extent to which the developed software is an accurate representation of the real world from the perspective of the intended use. However, verification is the process of deciding whether computational software correctly represents the developers fundamentals concepts and specifications [56].

To be as general as possible and to take into account the diversity of existing approaches concerning measurement validation, a list of criteria was proposed to be checked throughout the QCE life cycle for verification progress. This criteria is classified into two categories: 1) relative and 2) absolute [44].

Relative verification criteria are criteria which are dependent on the measurement design goals, the precision needed, the maturity of our knowledge about the attribute, etc. In such cases, it would make sense to talk about a criterion that is sufficiently satisfied.

Absolute verification criteria are criteria which, if any of these criteria are not satisfied, we cannot be sure we are measuring what we purport to be measuring [44].

Software verification is often confused with software validation. According to [44, 48, 61] the difference between verification and validation is described;

- Software verification asks the question, "Are we building the product right?" that is, does the software conform to its specification.
- Software validation asks the question, "Are we building the right product?" that is, is the software doing what the user really requires.

Since verification discusses if the product has been developed correctly, the proposed computational system must assure the verification progress based on correct responds to the instrumentation requirements. On the other hand, this system must be validated to corroborate whether it was essential for the system to be built. Correspondingly, to ensure the level of respond to the developed instrument, QCE suggests engagement of the following data collection; analysis and representation table as general raw implementation frameworks. Table 2 projects verification percentage of assigned tasks.

Table 2 Verification Parameter Measurements [Sample]

Percentage of Participants' Success	Task	Verification
Percentage + Interpretations	The first task + Matching Trials + Descriptions	Validated or Invalid
……..	……..	……..
Percentage + Interpretations	The last task + Matching Trials + Descriptions	Validated or Invalid

Assessment of Effectiveness (QCE Charrette Test)

After verification of the developed QCE instrument, the effectiveness assessment is executed. This test is carried out based on Charrette test principles. The term "charrette" itself is a French terminology meaning "cart". In the old days, when architecture students were supposed to do design assignments and wanted to take advantage of their maximum time, they would hire a cart to carry the submission materials to the institute. The student would work "en charrette" "on the cart" in order to finalize the project even until the last moment while the cart was running. Currently, charrette is applied for short but intense problems to be solved. In this regards, charrette makes use of researchers to sort out a certain task through different procedures while comparing the procedures and results of each replicate accordingly. In charrette test, the newly developed methodology is called the "innovative process". The normally used procedure is considered as the "conventional process". Correspondingly, the main aim of research in these types of studies is to identify whether the innovative process is more effective than the conventional process [30, 31, 37].

QCE discusses about preparation of innovative approaches for conducting conventional tasks. Development of new methodologies could be tested in terms of worked examples, demonstrations and trials [31, 34, 37, 38, 63]. The worked example is to prove that the demonstrated concept is actually applicable. Generally, a worked example is referred to a "hypothetical sessions imagined by the researchers or actual sessions undertaken by the researchers"[34]. A worked example is usually used to illustrate and visualize the innovative concept. Subsequently, demonstrations are the methodologies in which a newly developed research software would be provided to the outside in order to receive feedbacks. Demonstration is a stronger evidence than the worked example since a large sample of participants is considered as a higher reliability level. Finally, individuals are asked to conduct the innovative process in a trial method. Since the innovative process is used through normal people and not only the researchers in a trial methodology, the generality of the results is much stronger than the worked example and demonstration methods. However, given that the new methodology is not tested within different projects and through various users in a single trial or small number of trial methods, lack of reliability and validity is highly considerable. Other methodologies such as protocol studies are also engaged in order to test a newly developed process. Correspondingly, participants would be provided with the new set of tools while being observed by the researcher. Activities would subsequently be translated in terms of tasks. This is to help participants identify the structure, function and behaviour of the innovative process [30, 34, 37, 38, 63].

Different methodologies for testing innovative processes as worked examples, demonstration, trial and protocol studies were discussed. However, all the mentioned methodologies have shortage in measuring the effectiveness of the innovative process. Since the worked example is generally used through the researcher, it lacks in objectivity, reliability and validity. Worked examples are not facilitated with measurements to identify to what extent the new process is successful. However, the demonstration method has engaged some outside audience to examine the methodology. However, in order for better objectivity, normal people were neglected during conducting of the innovative process. Therefore; although this method is acceptably reliable but, it is still not appropriately valid. The trial methodology has achieved better generality and objectivity while normal people from outside the project could test the new process rather than researchers. Hence, the issue of effectiveness is not measured throughout trial methodologies while reducing the level of validity [36]. As a result, charrette test method has proposed that new methodologies for evaluating the innovative processes should be developed while providing;

- Higher level of reliability through executing multiple trials (Multiple examples)
- Higher level of reliability through replicates of the test by different testers
- Higher level of validity through decreased biasness level and increased reliance on objective measurements
- Higher level of validity through assessing the effectiveness in comparison with the conventional processes

Consequently, the charrette test method overcomes all the aforementioned shortages. It provides improvements in terms of objectivity, generality, reliability and validity. Charrette test method is highly capable of being engaged in design cognition and computing studies.

Software productivity, effectiveness and usability testing are the essential tests for ascertaining the validity of a newly developed software. Charrette takes advantage of the aforementioned aspects in order to examine a newly developed approach. In a charrette test, variables of speed and accuracy are compared during innovative and conventional processes [30, 31, 34].

Unlike the Charrette test in architectural experiments, the charrette test in computational experiment is an established methodology to test theoretical findings. Charrette test was developed at CIFE (Centre for Integrated Facility Engineering - Stanford University) where researchers had to test prototypes to examine theoretical findings and to validate them. Generally, the main intent of charrette test method is to prove fulfillment of a certain process through a certain set of tools is superior to conducting the same process through a different set of tools [34, 38]. Correspondingly, charrette calls the issue of “Effectiveness” for measurement of the preferred process in terms of better quality and higher speed.

When a software is used in order to enhance a certain process, the term “software usability” is translated as the “effectiveness” of the software in improving the highlighted process. However, ascertaining the effectiveness of a new process specifically while it is developed at a research prototype extent is a difficult task to be done. QEC uses an empirical comparative experiment methodology to define the validity of newly developed processes in design studies.

V. DEFINITION AND COMPONENTS OF A CHARRETTE TEST

Charrette test method is initiated with development of charrette test components as follows.

1) Process; a set of activities that are fulfilled in order to perform a certain job. Any charrette test method studies at least two processes.
2) Innovative Process; the newly developed experimental process that will be tested
3) Conventional Process; the baseline process for conducting the job
4) Proposition; a declaration about the effectiveness of the innovative process
5) Trial; a session in which the participant fulfill the task while the researcher can collect the corresponding data
6) Task; the job that should be remained constant in both of the innovative and conventional processes.

Number of participants able to finish the task within an acceptable time in comparison with the total number of participants proves or rejects the usability of the software. Subsequently, the time consumed for conducting each trial by the participants is measured. In other words, the innovative process and the conventional process will be measured in terms of time consumption through the participants in order to finish the same task. Average time will be calculated for each trial. Comparison of the measured average times represents the faster process. Moreover, accuracy is best advised to be measured in terms of variations from the correct answer. A certain optimum answer is assumed as the correct answer to the assigned task. Variations from the correct answer will be measured for both tasks and for each participant. Accuracy can be evaluated as the amount of variation from the correct answer displayed through several trials. This accuracy level can be measured through arithmetic calculations. Combination of the both accuracy levels determines the most accurate process.

VI. CHARRETTE TEST GUIDELINES

In order to conduct a charrette test, the following guidelines should be followed.

1) Preparation of clear hypothesis.
2) “Devise two or more processes for performing the same task, one to be designated the innovative process and one to be designated the conventional process. Formalize the two processes to an extent such that clear measurements of performance may be taken.” [34].
3) “Develop clear, quantifiable measurements of participants’ performance.”[34]
4) The task should be defined so that it can be clearly accomplished in a limited period of time.
5) The computational prototype should be clearly developed not to cause any unwanted interface complexities and crashes while causing biasness in the comparison results.
6) A certain previously designed test should be provided to all the participants. Participants should be given a minimum of two trials while the second trial will lessen the risk of biasness through learning about the test. Time delays should be assigned in between the conventional and computational progress completed by the same participants in order to reduce the risk of biasness.
7) Participants should be selected as the representatives of the target group.
8) Measurements which have been previously settled through the hypothesis will be measured after the test conducted.
9) The collected data should be statistically analyzed to prove/reject the hypothesis. Speed and accuracy are the two common parameters which are compared in test cases [27, 31, 34, 37].

VII. LIMITATIONS OF CHARRETTE TEST

A charrette test is conducted to compare only two processes where mainly the conventional and the innovative process are compared accordingly. Charrette test method does not usually compare multiple projects while a single project is basically selected for all trials. In other words, all the participants are assigned to undertake a task based on the same design problem. Charrette test method is advised to be used in intense circumstances. Participants should be asked to conduct the task quickly and accurately. In many cases, the researcher does not expect the participant to completely finish the given task within the limited time duration. The researcher usually measures the same data for both of the conventional and innovative trials for further comparison.

Charrette test principles highlight engagement of minimum three participants for examination of the effectiveness of a newly developed system [34, 37]. On the other hand statistical data analysis is carried out during the QCE test. Since inferential analysis is done in addition to descriptive analysis in order to generalize the findings, minimum of 30 data sets are required to be available. Therefore, using probability sampling [1, 64-68], QCE follows random sampling [69] in order to randomly select minimum three individuals as the participants for test trials. This approach is recommended for ensuring an acceptable probability sampling in order to be able to generalize the findings [68-70].

It is recommended to run the experiment in pilot and main test sessions in order to increase the reliability of eventual outcomes [68]. During the pilot test, participants will be given similar tasks (based on a developed QCE tasks-to-be-fulfilled protocol) to run in a multiplier of minimum three replicates. Finishing time and accuracy level of each task for each participant for each trial for each replicate must be measured. Finishing interpretation time (Thinking time for each task);
finishing critique time (Implementation time); finishing overall time and the accuracy level of each task for each participant for each trial for each replicate must be measured. Substantial improvement of speed and accuracy for the innovative process compared to the conventional process is considered as the positive response to effectiveness testing. In other words, higher speed and accuracy highlights the innovative process to be effective. Table 3 represents a data collections platform for task-based verification measurements of the instrument during the three replicates.

Table 3 Computational System Verification Measurement - Task-based [Sample]

Task/Time	Rep 1-Time Interpretation	Rep 1-Time Critique	Rep 2-Time Interpretation	Rep 2-Time Critique	Rep 3-Time Interpretation	Rep 3-Time Critique	Total-Time Interpretation	Total-Time Critique
First Task	Spent Time	Spent Time	Spent Time	Spent Time	Spent Time	Spent Time	Summation	Summation
……	Spent Time	Spent Time	Spent Time	Spent Time	Spent Time	Spent Time	Summation	Summation
Last Task	Spent Time	Spent Time	Spent Time	Spent Time	Spent Time	Spent Time	Summation	Summation
Total Sum	Summation	Summation	Summation	Summation	Summation	Summation	Overall Summation	Overall Summation

Figure 2; Figure 3 and Figure 4 represent sample parameter measurement comparisons in graphical formats.

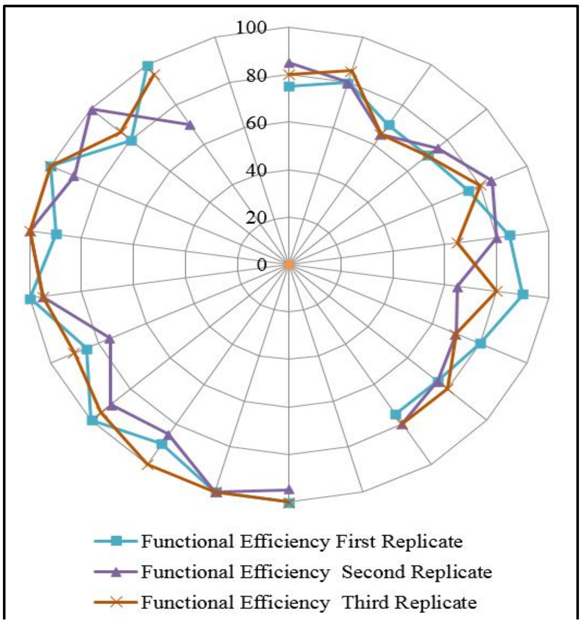

Figure 2 Accuracy Level Comparisons - Main Test/Pilot Test [Sample]

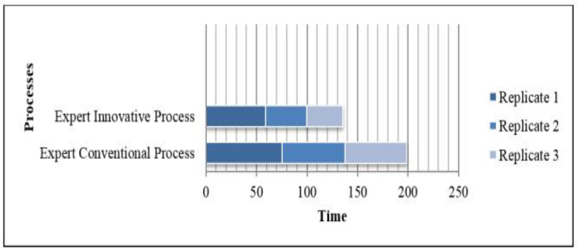

Figure 3 Performance Time Comparisons - Main Test/Pilot Test – Overall Mean – [Sample]

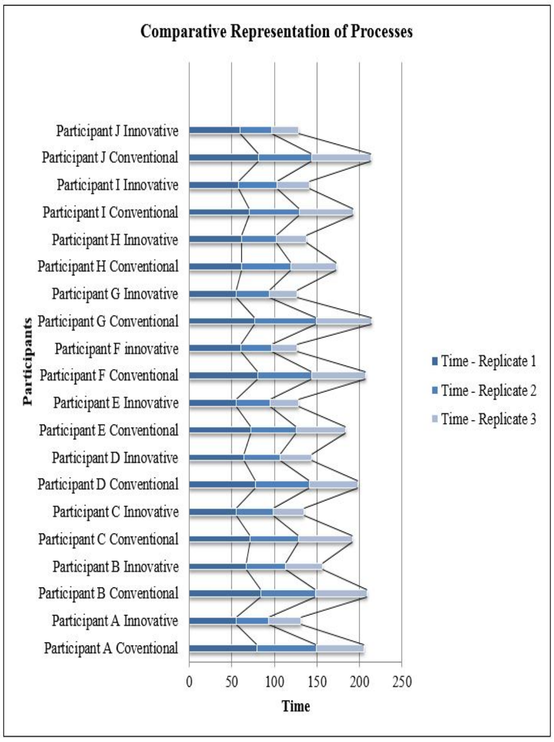

Figure 4 Performance Time Comparisons - Main Test - Separate Participants [Sample]

QCE comprises of various hypothesis regarding each test case for pilot and main test sessions. Eventually, it is recommended to prepare a QCE hypothesis testing result projection table in order to systematically organize and represent the testing outcomes as shown in Table 4.

Table 2 QCE Hypothesis Testing Results - Charrette Test [Sample]

TEST	HYPOTHESIS [1, 4]	TRIALS	CRONBACH'S ALPHA	RESULT – (95%)	REMARKS
PILOT	Hypothesis X (Trial X + Trial X)	*Trial X (The Innovative Process) + Trial X (The Conventional Process)*	x.xxx – Reliable or x.xxx Unreliable	Null Hypothes is rejected at x.xxx Or Null Hypothes is accepted at x.xxx	
PILOT	Hypothesis X (Trial X + Trial X)	*Trial X (The Innovative Process) + Trial X (The Conventional Process)*	x.xxx – Reliable or x.xxx Unreliable	Null Hypothes is rejected at x.xxx Or Null Hypothes is accepted at x.xxx	

MAIN	Hypothesis X (Trial X + Trial X)	*Trial X (The Innovative Process) + Trial X (The Conventional Process)*	x.xxx – Reliable or x.xxx Unreliable	Null Hypothes is rejected at x.xxx Or Null Hypothes is accepted at x.xxx	
MAIN	Hypothesis X (Trial X + Trial X)	*Trial X (The Innovative Process) + Trial X (The Conventional Process)*	x.xxx – Reliable or x.xxx Unreliable	Null Hypothes is rejected at x.xxx Or Null Hypothes is accepted at x.xxx	
MAIN	Hypothesis X (Trial X + Trial X)	*Trial X (The Innovative Process) + Trial X (The Conventional Process)*	x.xxx – Reliable or x.xxx Unreliable	Null Hypothes is rejected at x.xxx Or Null Hypothes is accepted at x.xxx	
MAIN	Hypothesis X (Trial X + Trial X)	*Trial X (The Innovative Process) + Trial X (The Conventional Process)*	x.xxx – Reliable or x.xxx Unreliable	Null Hypothes is rejected at x.xxx Or Null Hypothes is accepted at x.xxx	
MAIN	Hypothesis X (Trial X + Trial X)	*Trial X (The Innovative Process) + Trial X (The Conventional Process)*	x.xxx – Reliable or x.xxx Unreliable	Null Hypothes is rejected at x.xxx Or Null Hypothes is accepted at x.xxx	

Computational Simulation/Emulation Validation

QCE validation procedure is based on principles for computational simulation/emulation testing [45, 71]. This validation consists of three incremental stages. On the other hand, "researchers interested only in theory would conduct intellective experiments, based on stylized models, and test the sensitivity of the model to changes in input parameters. The results then can be compared to existing theory, or if new behaviors are observed, new macro-level theory discovered using simulation." [71]. This investigation can be validated using experiments that examine reasoning. Similarly, "investigators interested in macro-behaviors can use experiments in reasoning and representation to evaluate whether their theories have captured the essential elements of the computational system and can reason about these representations using computer simulation." [71]. Subsequently, the usefulness of the computational emulation is validated as the final step. QCE researchers will define the level of required validation based on the corresponding research limitations.

Since not all simulation models are developed for the same reason, validation methods should be tailored to reflect the intention of the simulation model [71-77]. Validation of simulation models is not a substance of valid or invalid, but rather of degree [78, 79]. In this regards, QCE recommends [71]'s three main stages for validation. [71] proposes validation for simulation/emulation of organizational structures and their behavior. QCE considers organization as the design studies project being evaluated

1. Validation of Reasoning using Toy Problems and Intellective Experiments (Theoretical Validation – Operationalization of Theories)

1.1. Toy problem (Whether the simulation engine works correctly and the data can be correctly simulated or not)

1.2. Intellective Simulation (To examine hypothetical problems in idealized settings. Variable values will be altered to test the computational software's framework limitations. The output results will be compared with the real world expected results for validation)

2. Validation of Reasoning and Representation (Real World Data Validation)

2.1. Validation of Authenticity ("Can we represent a real organization with our simulation model?" [71] and "Can we emulate quantitatively relevant performance characteristics of the organization?" [71]. This stage is about gathering information from the real-world and translating that information into a symbolic language that a computer simulation can understand.)

2.2. Validation of Generalizability ("To assess whether the model is over fitted to a particular organizational setting") [71]. Generalizability is expected as a major principle for assessing the eminence of the respective research [24, 80, 81]

2.3. Validation of Reproducibility (Validates whether two modelers will get the same results when they model and simulate the same organization.)

3. Validation of Reasoning, Representation and Usefulness

3.1. Retrospective Validation (A retrospective experiment duplicates past performance, using a simulation model, and calibrates the model as needed to reproduce previous experiences.)

3.2. Gedanken Validation of Hypotheses (These validations ask what-if questions of participants, the simulation system, and theory, and then compare the answers. The simulation of this hypothetical scenario is then compared to 1) theory (to make sure it is consistent with theory) and 2) to predictions made by managers in the organizations.

3.3. Natural History Validation (In this validation experiment, an organization is modeled, the future results of the project are predicted, and the organization is observed to see if performance predictions come true.)

3.4. Prospective Validation with Interventions (Not only predicts the future, but also attempts to change the future based on the results of the simulation.)

Table 5 represents the raw data platform for presentation of successfulness percentage of task validation measurements.

Table 5 Validation Parameter Measurements [Sample]

Percentage of Participants' Success	Task	Validation
Percentage + Interpretations	The first task + Matching Trials + Descriptions	Validated or Invalid
……..	……..	……..
Percentage + Interpretations	The last task + Matching Trials + Descriptions	Validated or Invalid

VIII. SUMMARY

This paper elaborated on utilization of QCE for social science research. QCE comprises of four main steps of 1) computational system design and development [instrumentation]; 2) computational system verification; 3) assessment of effectiveness (QCE Charrette test) and 4) computational simulation/emulation validation. QCE is advised to be used while conducting design studies on examining the effectiveness of a newly developed system in comparison with the conventional process. QCE highlights the inevitable rapid conquest of computer and technology within the field of research and professional practice. It promotes the resultant popularity gain of developed quantitative and measurement-based research approaches during contemporary design studies research. Researchers are advised to follow the provided QCE raw data collection; analysis and representation platforms in order to ensure systematic academic appropriateness. Eventually, QCE is expected to be regulated as a systematic platform for those design studies researchers applying innovative implementation techniques using computationally developed inventive systems.

REFERENCES

[1]. Creswell, J.W., Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. 2013: SAGE Publications.

[2]. Creswell, J.W., Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research. 2014: Pearson Education.

[3]. De Vaus, D.A. and P.D. de Vaus, Research Design in Social Research. 2001: Sage Publications (CA).

[4]. Bryman, A., Social research methods. 2012: Oxford university press.

[5]. Gorard, S. and C. Taylor, Combining Methods In Educational And Social Research. 2004: Open University Press.

[6]. Tashakkori, A. and C. Teddlie, SAGE Handbook of Mixed Methods in Social & Behavioral Research. 2010: SAGE Publications.

[7]. Bryman, A. and D. Cramer, Quantitative Data Analysis with IBM SPSS 17, 18 & 19: A Guide for Social Scientists. 2012: Taylor & Francis.

[8]. Teddlie, C. and A. Tashakkori, Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences. 2009: Sage Publications Inc.

[9]. Abowitz, D.A. and T.M. Toole, Mixed method research: Fundamental issues of design, validity, and reliability in construction research. Journal of Construction Engineering and Management, 2009. 136(1): p. 108-116.

[10]. Flemming, K., Synthesis of quantitative and qualitative research: an example using Critical Interpretive Synthesis. Journal of advanced nursing, 2010. 66(1): p. 201-217.

[11]. Palinkas, L.A., et al., Mixed method designs in implementation research. Administration and Policy in Mental Health and Mental Health Services Research, 2011. 38(1): p. 44-53.

[12]. Stentz, J.E., V.L.P. Clark, and G.S. Matkin, Applying mixed methods to leadership research: A review of current practices. The Leadership Quarterly, 2012. 23(6): p. 1173-1183.

[13]. Bryman, A., Integrating quantitative and qualitative research: how is it done? Qualitative research, 2006. 6(1): p. 97-113.

[14]. Lewis-Beck, M., A. Bryman, and T.F. Liao, The Sage encyclopedia of social science research methods. 2003: Sage Publications.

[15]. Gorard, S., Research Design: Creating Robust Approaches for the Social Sciences. 2013: SAGE Publications.

[16]. Markusson, N. and H. Chalmers, Characterising CCS learning: The role of quantitative methods and alternative approaches. Technological Forecasting and Social Change, 2013. 80(7): p. 1409-1417.

[17]. Gero, J.S. and M.L. Maher, Modeling Creativity and Knowledge-Based Creative Design. 2013: Taylor & Francis.

[18]. Gero, J.S., Artificial Intelligence in Design '91. 2014: Elsevier Science.

[19]. Coyne, R.D.D., et al., Knowledge-Based Design Systems. 1989: Addison-Wesley Longman Publishing Co., Inc. 416.

[20]. Gero, J.S., Design Computing and Cognition '12. 2014: Springer Netherlands.

[21]. Lawson, B., How Designers Think: The Design Process Demystified. 2014: Elsevier Science.

[22]. Bilda, Z., J.S. Gero, and T. Purcell, To sketch or not to sketch? That is the question. Design studies, 2006. 27(5): p. 587-613.

[23]. Shadish, W.R., T.D. Cook, and D.T. Campbell, Experimental and quasi-experimental designs for generalized causal inference. 2002: Wadsworth Cengage learning.

[24]. Polit, D.F. and C.T. Beck, Generalization in quantitative and qualitative research: Myths and strategies. International journal of nursing studies, 2010. 47(11): p. 1451-1458.

[25]. Hardy, M.A. and A. Bryman, Handbook of Data Analysis. 2009: SAGE Publications.

[26]. Bryman, A. and P.S.R.A. Bryman, Quantity and Quality in Social Research. 2003: Taylor & Francis.

[27]. Koo, B., M. Fischer, and J. Kunz, A formal identification and re-sequencing process for developing sequencing alternatives in CPM schedules. Automation in Construction, 2007. 17(1): p. 75-89.

[28]. Donath, D. and D. Lobos, Plausibility in early stages of architectural design: a new tool for high-rise residential buildings. Tsinghua Science & Technology, 2009. 14(3): p. 327-332.

[29]. Gero, J.S., Design prototypes: a knowledge representation schema for design. AI magazine, 1990. 11(4): p. 26.

[30]. Senescu, R.R. and J.R. Haymaker. Specifications for a social and technical environment for improving design process communication. in W078 26th International Conference, Managing IT in Construction, Istanbul, Turkey. 2009.

[31]. Clevenger, C.M., J.R. Haymaker, and A. Ehrich, Design exploration assessment methodology: testing the guidance of design processes. Journal of Engineering Design, 2013. 24(3): p. 165-184.

[32]. Knox, M.W., et al., Impact of Charrettes and Their Characteristics on Achieved LEED Certification. Journal of Architectural Engineering, 2013. 20(3).

[33]. Clayton, M.J., et al., The Charrette Testing Method for CAD Research. Applied Research in Architecture and Planning, 1996. 2: p. 83-91.

[34]. 34. Clayton, M.J., J.C. Kunz, and M.A. Fischer, The charrette test method. CIFE Technical Rep, 1998. 120.

[35]. Fischer, M., Formalizing construction knowledge for concurrent performance-based design, in Intelligent Computing in Engineering and Architecture. 2006, Springer. p. 186-205.

[36]. Koo, B., M. Fischer, and J. Kunz, Formalization of construction sequencing rationale and classification mechanism to support rapid generation of sequencing alternatives. Journal of Computing in Civil Engineering, 2007. 21(6): p. 423-433.

[37]. Senescu, R.R. and J.R. Haymaker, Communicating design processes effectively and efficiently. Design Process Communication Methodology, 2011: p. 107.

[38]. Garr, M., et al., Workflow and Resources for Facility Operations Analysis. 2012, Citeseer.

[39]. Jennings, E., Models for pretest-posttest data: repeated measures ANOVA revisited. Journal of Educational and Behavioral Statistics, 1988. 13(3): p. 273-280.

[40]. Hennig, C., D. Müllensiefen, and J. Bargmann, Within-Subject Comparison of Changes in a Pretest-Posttest Design. Applied Psychological Measurement, 2010. 34(5): p. 291-309.

[41]. Ibrahim, R., Demystifying the Arduous Doctoral Journey: The Eagle Vision of a Research Proposal. The Electronic Journal of Business Research Methods, 2011. 9(2): p. 130-140.

[42]. Boehm, B., et al., The Incremental Commitment Spiral Model: Principles and Practices for Successful Systems and Software. 2014: ADDISON WESLEY Publishing Company Incorporated.

[43]. Selby, R.W., Software Engineering: Barry W. Boehm's Lifetime Contributions to Software Development, Management, and Research. 2007: Wiley.

[44]. Valacich, J.S., J.F. George, and J. Hoffer, Essentials of Systems Analysis and Design. 2014: Pearson Education, Limited.

[45]. Thomsen, J., et al., A trajectory for validating computational emulation models of organizations. Computational & Mathematical Organization Theory, 1999. 5(4): p. 385-401.

[46]. GhaffarianHoseini, A., R. Ibrahim, and R. Abdullah, Graphical visualization principles for maintaining functional relativity of spaces during architectural design. International Journal of Sustainable Tropical Design Research and Practice ALAM CIPTA, 2009. 4(1): p. 9-16.

[47]. Levitt, R.E., Computational modeling of organizations comes of age. Computational & Mathematical Organization Theory, 2004. 10(2): p. 127-145.

[48]. Hoffer, J.A., J.F. George, and J.S. Valacich, Modern Systems Analysis and Design (2014,). INSTRUCTOR, 2015.

[49]. Rice, W.R., Analyzing tables of statistical tests. Evolution, 1989: p. 223-225.

[50]. Gorard, S., Quantitative Methods in Educational Research: The Role of Numbers Made Easy. 2001: Bloomsbury Academic.

[51]. Freiberg, M., A. Striffler, and F. Puppe, Extensible Prototyping for pragmatic engineering of knowledge-based systems. Expert Systems with Applications, 2012. 39(11): p. 10177-10190.

[52]. Bäumer, D., et al. User interface prototyping—Concepts, tools, and experience. in CSE '96 proceedings of the 18th international conference on software engineering. 1996.

[53]. Sears, A. and J.A. Jacko, The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, Second Edition. 2007: CRC Press.

[54]. Fisher, M.S., Software Verification and Validation: An Engineering and Scientific Approach. 2007: Springer.

[55]. Fisher, M.S., Software Verification and Validation: An Engineering and Scientific Approach. 2010: Springer US.

[56]. Guo, Y., J. Slay, and J. Beckett, Validation and verification of computer forensic software tools—Searching Function. digital investigation, 2009. 6: p. S12-S22.

[57]. Debbabi, M., et al., Verification and Validation in Systems Engineering: Assessing UML/SysML Design Models. 2010: Springer Science & Business Media.

[58]. Sargent, R.G. Verification and validation of simulation models. in Proceedings of the 37th conference on Winter simulation. 2005. Winter Simulation Conference.

[59]. IEEE Standard for System and Software Verification and Validation - Redline. IEEE Std 1012-2012 (Revision of IEEE Std 1012-2004) - Redline, 2012: p. 1-223.

[60]. ISO, I., General requirements for the competence of testing and calibration laboratories, ISO/IEC 17025: 2005. International Organization for Standardization, Geneva, Switzerland, 2005.

[61]. Boehm, B., Software Engineering; R&D Trends and Defense Needs. In R. Wegner, ed. Research. Directions in Software Technology. 1979, Cambridge, MA: MIT Press.

[62]. SWGDE, SWGDE Recommended Guidelines for Validation Testing Version: 2.0 (September 5, 2014). 2014.

[63]. Clayton, M.J., M.A. Fischer, and J.C. Kunz. CAD prototype testing: worked examples, demonstrations, trials, and charrettes. in Computing in Civil Engineering (1998). 1998. ASCE.

[64]. Earl, B., The Practice of Social Research. 2015: Cengage Learning.

[65]. Daniel, J., Sampling Essentials: Practical Guidelines for Making Sampling Choices. 2011: SAGE Publications.

[66]. Sapsford, R. and V. Jupp, Data Collection and Analysis. 2006: SAGE Publications.

[67]. Lim, W.M. and D.H. Ting, Research Methodology: A Toolkit of Sampling and Data Analysis Techniques for Quantitative Research. 2013: GRIN Verlag.

[68]. Fraenkel, J.R., N.E. Wallen, and H. Hyun, How to Design and Evaluate Research in Education. 2014: McGraw-Hill Education.

[69]. Levy, P.S. and S. Lemeshow, Sampling of Populations: Methods and Applications. 2013: Wiley.

[70]. Wallen, N.E. and J.R. Fraenkel, Educational Research: A Guide To the Process. 2013: Taylor & Francis.

[71]. Thomsen, J., et al., A trajectory for validating computational emulation models of organizations. 2003, CIFE Working Paper.

[72]. Burton, R.M. and B. Obel, Strategic Organizational Diagnosis and Design: The Dynamics of Fit. 2004: Springer.

[73]. Burton, R.M., B. Obel, and G. DeSanctis, Organizational Design: A Step-by-Step Approach. 2011: Cambridge University Press.

[74]. Burton, R.M., et al., Organization Design: The Evolving State-of-the-Art. 2011: Springer US.

[75]. Burton, R.M., et al., Designing Organizations: 21st Century Approaches. 2010: Springer US.

[76]. Kleindorfer, G.B., L. O'Neill, and R. Ganeshan, Validation in simulation: various positions in the philosophy of science. Management Science, 1998. 44(8): p. 1087-1099.

[77]. Sargent, R.G., Verification and validation of simulation models. Journal of Simulation, 2013. 7(1): p. 12-24.

[78]. Law, A., Simulation Modeling and Analysis. 2014: McGraw-Hill Education.

[79]. Law, A.M. How to build valid and credible simulation models. in Simulation Conference (WSC), Proceedings of the 2009 Winter. 2009. IEEE.

[80]. Lee, H.B., Foundations of Behavioral Research. 2008: Wadsworth.

[81]. Polit, D.F. and C.T. Beck, Nursing research: Generating and assessing evidence for nursing practice. 2008: Lippincott Williams & Wilkins.

★★★