Khan, S. and Tzortzopoulos, P., 2016. An evaluation framework for an action research study on lean design management. In: *Proc. 24th Ann. Conf. of the Int'l. Group for Lean Construction*. Boston, USA, sect.1pp. 113-122, available at www.iglc.net

A FRAMEWORK FOR EVALUATING AN ACTION RESEARCH STUDY ON LEAN DESIGN MANAGEMENT

Sheriz Khan[1] and Patricia Tzortzopoulos[2]

ABSTRACT

There is no convention for evaluating action research (AR), and the one commonly used for evaluating research in general was deemed unsuitable for evaluating a postgraduate AR study on lean design management conducted by the researchers. The purpose of this paper is to present the framework developed by the researchers for evaluating the AR study, using criteria that are different from those traditionally used to evaluate research. It is hoped that this paper will contribute to lean construction research (LC) by highlighting the importance of using AR to measure the efficacy of LC production systems in their intended context of application and by letting future LC researchers know that they can develop their own criteria for evaluating their research rather than use generic criteria that may not be suitable.

KEYWORDS

Action research, LPS WWP, rigor, relevance

INTRODUCTION

Validity, *reliability* and *generalizability* are criteria commonly used to evaluate both quantitative (positivist) research and qualitative (interpretive) research. However, these criteria were developed by quantitative researchers to test the methodological rigor of quantitative research and have been used as the basis for criticizing qualitative research, especially case study research and AR, as lacking rigor and containing bias because of the subjective nature of data collection and analysis employed (Robson, 2002; Yin, 2009). AR has been criticized for its lack of rigor by applying quantitative criteria for rigor: rigor can be compromised if quantitative criteria are applied too rigidly to AR (Dick, 2014). This has raised concerns among qualitative researchers about the practicality of using quantitative criteria to evaluate qualitative research, so much so that there is a growing literature in qualitative research in general and AR in particular that encourages the use of

[1] PhD Candidate, School of Art, Design and Architecture, Department of Architecture and 3D Design, University of Huddersfield, Queen Street Studios, Queensgate, Huddersfield, HD1 3DH, England, United Kingdom. Telephone +1 (321) 674-0309, sherizkhan@yahoo.com

[2] Professor, School of Art, Design and Architecture, Department of Architecture and 3D Design, University of Huddersfield, Queen Street Studios, Queensgate, Huddersfield HD1 3DH, England , United Kingdom. Fax: (+44) 01484 472440, p.tzortzopoulos@hud.ac.uk

criteria other than those used in quantitative research (Dick, 2014). Some qualitative researchers (e.g., Lincoln, 1995; Remenyi et al., 1998; Golafshani, 2003) are content with simply redefining the quantitative criteria for qualitative research, while others (e.g., Miles and Huberman, 1994; Reason and Bradbury, 2008; Coghlan and Brannick, 2010) have proposed different criteria for evaluating AR. The researchers considered the practical relevance of their AR study more important than its methodological rigor so, in the absence of a convention for evaluating AR, they developed their own framework for evaluating the AR study, based on its practical relevance. Before presenting the evaluation framework, AR as adopted in the AR study will be defined, the AR study will be outlined, and a strategy used by the researchers to achieve rigor and minimize bias in the AR study will be discussed.

ACTION RESEARCH

AR is a strategy for implementing and evaluating an existing solution to a practical problem in its organizational context, with the knowledge acquired from the implementation and evaluation used to make recommendations for future application of the solution (Iivari and Venable 2009). It is an approach to research which is based on collaborative problem-solving relationship between researcher(s) and practitioners, which aims at both managing change and creating new knowledge (Coghlan and Davis, 2006). Researchers who adopt AR are likely to be practitioners who wish to improve understanding of their practice or more likely to be academics who have been invited into an organization by decision-makers aware of a problem requiring AR but lacking the requisite methodological knowledge to conduct it (Argyris and Schön, 1998; O'Brien, 2001). AR is a flexible cyclical process which allows action (change, improvement) and research (understanding, knowledge) to be achieved at the same time (Dick, 2002). Lewin (1946) is credited with pioneering AR which he portrayed as a spiral of learning steps consisting of planning action, taking action, evaluating action and amending the plan based on what was learned. The iteration within the AR spiral enables action and research to be built up cumulatively: through trial and error, both action and research can be pursued until achieved (Dick, 2014).

THE ACTION RESEARCH STUDY

Using AR the researchers facilitated the implementation of LPS WWP during the design development phase of two building design projects and together with the design practitioners evaluated its effects on planning reliability and workflow variability. The AR study was carried out at two AE firms in Florida. The descriptions of the projects studied are summarized in Table 1. The hotel design team consisted of a project manager, an architect, two intern architects (IAs), a structural engineer, a mechanical engineer, an electrical engineer, a plumbing engineer, four engineers-in-training (EITs), a BIM manager, and six BIM technicians. The apartment design team consisted of a project manager, the architect, an IA, a structural engineer, an MEP (mechanical-electrical-plumbing) engineer, three EITs, a BIM manager, and five BIM technicians. The design

development phase of the hotel project began 6 May 2013 and ended on 23 August 2013; for the apartment project, it began on 8 July and ended on 25 October 2013.

Table 1: Summary of the projects in the AR study

Project	Location	Floors	Size	Construction Cost	Design Phase	Duration
Hotel	Melbourne Beach, FL	7	14,865m²	$23.94 million	Design development	16 weeks
Apartment	Sebastian, FL	6	8,919m²	$13.60 million	Design development	16 weeks

The AR study was divided into two studies: a four-week exploratory study to assess the current design planning practice and design planning reliability, followed by a twelve-week action study aimed at increasing design planning reliability and reducing workflow variability. The exploratory study revealed that top-down, push planning was being practiced in both projects: a design management team, met for an hour or two on Friday mornings, agreed on the design tasks on a master schedule that *should* be performed in the coming week and, without making sure that they *could* be done, *push* them down with instructions and/or sketches to the BIM technicians to create models and generate drawings from them and to the IAs and EITs to research and prepare technical specifications. The IAs, EITs and BIM technicians were left out of the weekly task planning (WTP) process. The average Percent Plan Complete (PPC) over the four-week exploratory study period was 73.1 for the hotel project and 72.3 for the apartment project. This low planning reliability was a cause of high workflow variability.

By itself, push planning is not an effective approach to task scheduling. However, it is necessary in building design, and failure to supplement it with pull planning essentially deprives building designers of a technique for producing desired results (Ballard, 1999). The researchers therefore recommended that the traditional WTP (push planning) be supplemented with LPS WWP (pull planning) during the final twelve weeks of the design development phase. As a result, the entire design team met in the firm's conference room each Friday afternoon to participate in the design planning process and make commitment to finish the tasks on the master schedule that were to be performed in the coming week by agreed dates. Tasks were decomposed into smaller, doable assignments. Assignments that should be performed but which were hampered by incomplete prerequisites or unresolved constraints were not scheduled. No assignment was scheduled unless an agreement was reached on who was responsible for timely prerequisite handover and who will perform the assignment and by when. If it was determined that more manpower or other resources would be needed to complete a task on the master schedule by a certain time, then more manpower or other resources would be allocated to that task. PPCs and FRAs (failure reason analyses) played an important role in the LPS WWP implementation and evaluation process.

The AR study took the form of a spiral of steps, signifying the cyclical, iterative and recursive nature of AR (see Figure 1), starting with the four-week exploratory study in which deficiencies in the current design planning practice were identified and remedial actions were planned, followed by the action study in which the researchers and the practitioners engaged in twelve weekly action research cycles of planning, implementing,

monitoring and evaluating LPS WWP, aiming to improve each cycle of implementation by applying the lessons learned and avoiding the mistakes made in the previous cycle. The cyclic process was repeated until a sufficient understanding of and actionable solution for the workflow variability problem was achieved. The general goal was to create a simple, practical, repeatable process of iterative learning, evaluation and improvement that would lead to increasingly better results for the practitioners. So, at the heart of each cycle was learning through critical evaluation, i.e., changing patterns of thinking and action that were well established in two groups of practitioners. The idea of the learning cycle is also common to Kolb's (1984) experiential learning cycle and Deming's (1986) quality cycle (Plan-Do-Study-Act or PDSA) which drew upon Shewhart and Deming's (1939) Plan-Do-Check-Act (PDCA) cycle.

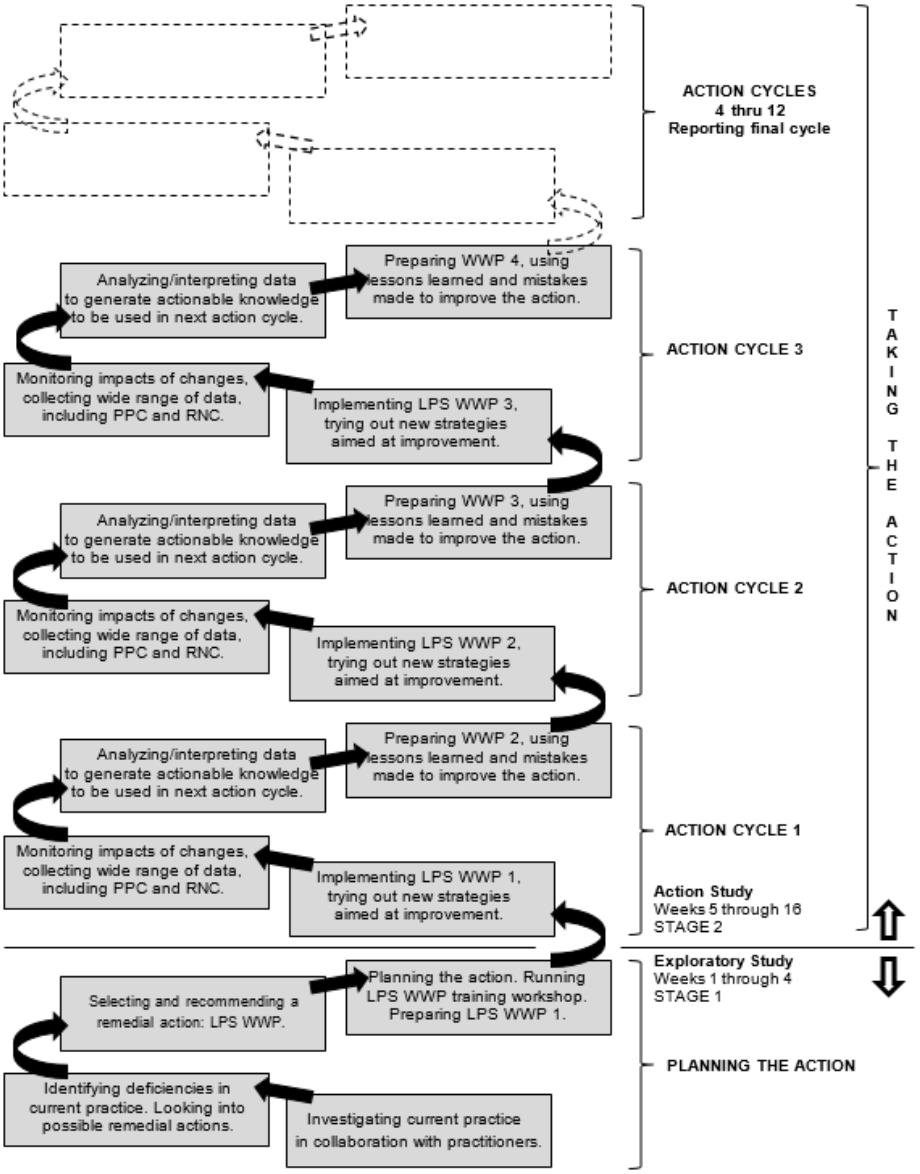

Figure 1: The AR study depicted as a cyclical learning process

PPC Measures Before And After LPS WWP Implementation

PPC measures served as a tangible incentive for the project teams to improve the predictability and reliability of the LPS weekly work plans and provided empirical evidence of the effectiveness of LPS WWP as a design planning and control tool. As shown in Figures 2 and 3, in both design projects, LPS WWP PPCs were higher than traditional WTP PPCs. There was 12% rise in average overall PPCs in the hotel design project and a 14% rise in average overall PPCs in the apartment design project, suggesting that there was an increase in planning reliability and thus reduction in workflow variability during the LPS WWP implementations. The hotel design development phase finished three days ahead of schedule, and the apartment design development phase finished two days ahead of schedule, which amounted to a 2.50% and a 3.75% increase in production cost efficiency, respectively, in this phase of the design projects.

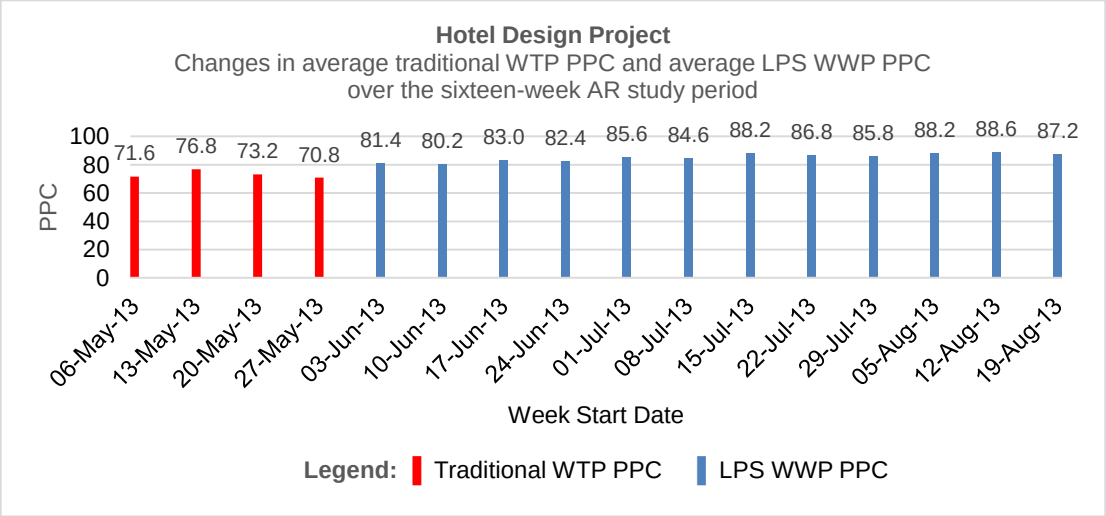

Figure 2: Hotel design project--PPCs before and after implementing LPS WWP

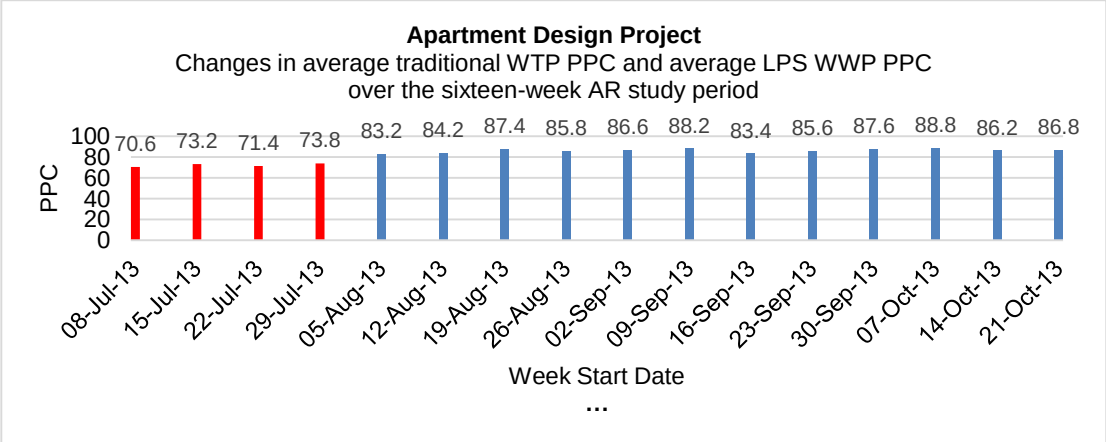

Figure 3: Apartment design project--PPCs before and after implementing LPS WWP

FAILURE REASON ANALYSES

A key feature of the continuous improvement process was the study of the reasons why assignments promised in the weekly work plans to be completed by a certain time were not completed by that time. FRAs were conducted to help improve each weekly cycle of LPS WWP implementation. This involved analyzing the causes of failure to complete daily assignments, thus facilitating learning from mistakes and helping to prevent those mistakes from happening again. The four main reasons for non-completion of assignments are shown in Table 2.

Table 2: Number and percentage of occurrences of non-completion of assignments

Reason	Project			
	Hotel		Apartment	
	Occurrences	Percentage	Occurrences	Percentage
Waiting for prerequisite work	22	36%	20	38%
Insufficient input information	19	31%	13	25%
Underestimation of time	17	28%	16	31%
Rework	3	5%	3	6%

ACHIEVING RIGOR/MINIMIZING BIAS IN THE AR STUDY

Validity, i.e., the accuracy of scientific findings, fit well with quantitative research, but the concept of *validity* is inappropriate for the qualitative part of an action research study, which relies on subjective interpretation of qualitative data collected in complex, uncertain and unpredictable systems (Dick, 2014). When quantitative methods are used in AR, as in the AR study, in which PPC measures were collected and a questionnaire survey was conducted, the quantitative concept of *reliability*—consistency and repeatability of the results over time—can apply to that part of the research, and conventional means of achieving reliability can be used (Dick, 2014). In some forms of AR, reliability and validity are achievable if they are applied less rigidly than traditionally applied (Dick, 2014). Furthermore, it is expected that all good scientific research should in some way or other be generalizable (Dick, 2014). In the strict sense of the term, *generalizability* can apply to the quantitative part of the AR study. However, to the extent that the practitioners in the two design projects and design practitioners elsewhere can use the new understanding of the research problem in similar situations or similar projects, *generalizability* is possible, and a contribution to knowledge can be claimed (Dick, 2014).

The AR study was designed to achieve rigor through data triangulation, i.e., convergence of evidence from different sources (see Figure 4): data was collected from the thirty-three practitioners, using multiple data collection techniques, including participatory and non-participatory observations, semi-structured and follow-up interviews, closed and open-ended questionnaires, individual and group discussions, and document reviews. Rigor was also achieved through methodological triangulation, i.e., convergence of evidence from different methods—quantitative and qualitative, aimed at corroborating the same facts. Methodological triangulation was be achieved through the use of both quantitative and qualitative data collection techniques. Using multiple sources

of data and combining methods, as well as multiple projects, strengthened the AR study (Patton, 1990). Active involvement in AR can increase researcher bias; however, triangulation may help to reduce it (Robson 2002). Practitioner participation in the evaluation of the AR study also helped to reduce researcher bias.

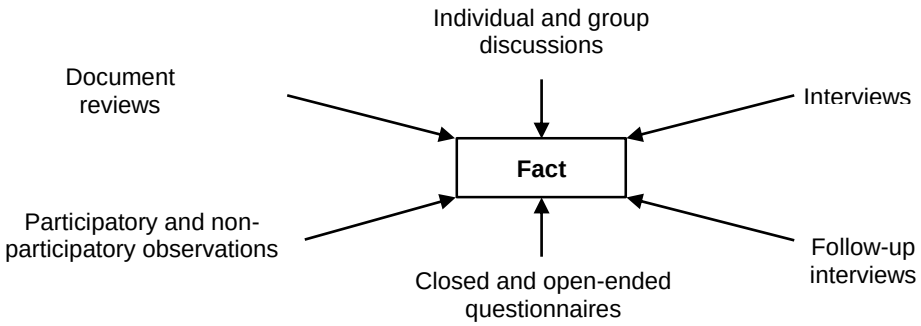

Figure 4: Data triangulation through convergence of multiple sources of evidence

Lastly, the cyclic process of planning, implementing, monitoring and critically evaluating LPS WWP helped the researchers and the practitioners to refine the action strategy as they learned more about their situation. The unfolding nature of the cyclical learning process contributed to the rigor of the research: the early cycles helped the researchers and the practitioners decide how to conduct the later cycles; and, in the later cycles, the interpretations developed in the early cycles were tested and challenged and refined.

THE EVALUATION FRAMEWORK

Continuous, internal evaluation is at the core of AR. As mentioned earlier, the two action studies consisted of a series of informal evaluations of LPS WWP in action. The goal was to continually refine the implementation of LPS WWP in light of the understanding developed in each earlier cycle in order to increase planning reliability and reduce workflow variability.

There is no formal framework for evaluating AR. However, criteria for evaluating AR were found in one or two places in the AR literature. Coghlan and Brannick (2010), for example, proposed following criteria for evaluating AR:

1. Correctness of the original diagnosis
2. Correctness of the action taken
3. Correctness of the way the action was taken

And Reason and Bradbury (2008) proposed what amounts to a checklist for quality in the form of a number of questions an action researcher can ask:

1. Did the research reflect cooperation between the researchers and the design practitioners?
2. Did the research enable actions guided by iterative evaluation as part of the process of change and improvement in the existing practice?
3. Did the research advance your practical and experiential knowledge?
4. Did the research engage in significant work?

5. Did the research result in new and enduring changes?

The criteria proposed by Coghlan and Brannick (2010), the checklist by Reason and Bradbury (2008) and concepts by March and Smith (1995), Kagiolglou et al. (1998), Smith and Morrow (1999), Bresnen and Marshall (2001), Cooper (2001), Tzortzopoulos (2004), and Brady et al. (2013) were used to develop a framework to evaluate the AR study that focused on evaluating its practical relevance. The framework was organized in three hierarchical levels: criteria, attributes and attribute definitions (see Table 3). The attribute definitions and the checklist proposed by Reason and Bradbury (2008) were used to design a questionnaire aimed at obtaining a comprehensive evaluation of the AR study by the researchers and the practitioners.

Table 3: Framework for evaluating usefulness and effectiveness of the actions

Criteria	Attributes	Attribute definitions
Correctness	*Diagnosis*	Accuracy of the original findings
	Treatment	Suitability of the action taken
	Execution	Conformity to the technique used for taking the action
Usefulness	*Applicability*	Appropriateness of the action for the situation
	Practicality	Ease of use of the action in terms of simplicity and clarity
	Flexibility	Adaptability of the action to the current practice
Effectiveness	*Efficacy*	Ability of the action to achieve the intended results
	Measurability	Ability of the action to be quantified
	Acceptability	Ability of the action to inspire trust in its value to practice

THE QUESTIONNAIRE

The questionnaire contained twenty closed questions, which were used to evaluate the correctness, usefulness and effectiveness of both the action and the research, and two open-ended questions, one for the practitioners to compare the design planning practice before and after LPS WWP was implemented and the other to find out what they considered to be the key drivers of and barriers to the adoption of LPS WWP in their firm. The closed questions were set up to be answered using a 1-to-5 Likert response scale, with *Strongly Agree* on one end and *Strongly Disagree* on the other end, and *Neither Agree nor Disagree* in the middle. The closed questions each began with *To what extent do you agree or disagree with the following statement?* followed by a statement (see Table 4).

Table 4: Statements in the questionnaire

1. The researcher's diagnosis of the problems associated with the traditional weekly task planning was accurate.
2. Supplementing traditional weekly task planning with LPS weekly work planning was the right action to take.
3. Traditional weekly task planning was properly supplemented with LPS weekly work planning using action research.
4. LPS weekly work planning was appropriate for the change needed in the design planning practice.
5. In terms of simplicity and clarity, LPS weekly work planning was easy to implement.

6. LPS weekly work planning was adaptable to traditional task weekly planning.
7. Collaborative production planning, task decomposition, make-ready planning, assignment completion commitment, PPC measurements and FRAs analyses introduced by LPS weekly work planning resulted in an increase in planning reliability and thus a reduction in workflow variability.
8. The increase in planning reliability and reduction in workflow variability resulting from supplementing traditional weekly task planning with LPS weekly work planning could be measured accurately.
9. The increase in planning reliability and reduction in workflow variability after traditional weekly task planning was supplemented with LPS weekly work planning has convinced you of the value to design planning practice of supplementing traditional weekly task planning with LPS weekly work planning.
10. LPS weekly work planning provided the workflow control mechanism that traditional weekly task planning lacked and that rendered it inadequate for the complex design project.
11. The action research strategy was correct, useful and effective with regard to allowing the researchers and practitioners to collaborate in systematic investigations, seeking practical solutions to workflow problems.
12. The action research strategy was correct, useful and effective with regard to enabling the researcher to influence practice directly instead of simply being an observer passively collecting data.
13. The action research strategy was correct, useful and effective with regard to encouraging meaningful discussions and hence a better understanding among all concerned of the practices of your firm and the problems in the project that was studied.
14. The action research strategy was correct, useful and effective with regard to permitting the practitioners to contribute effectively, ensuring that all information from target groups and individuals was obtained.
15. The action research strategy was correct, useful and effective with regard to allowing the researcher to diagnose and help solve design planning problems and design workflow problems during the design development phase of the building design project that was studied.
16. The action research strategy was correct, useful and effective with regard to fostering cooperation between the researcher and the practitioners.
17. The action research strategy was correct, useful and effective with regard to encouraging actions to be taken that were guided by iterative evaluation as part of the process of change in the existing design planning practice and improvement in design workflow.
18. The action research strategy was correct, useful and effective with regard to enabling actions that advanced your practical and experiential knowledge.
19. The action research strategy was correct, useful and effective with regard to facilitating the conduct of significant research.
20. The action research strategy was correct, useful and effective with regard to producing new and enduring changes to the design planning practices.

CONCLUSION

Until there is a convention for evaluating AR to follow, lean construction action researchers may have to develop their own set of criteria to evaluate their research.

REFERENCES

Argyris, C. and Schön, D. (1998). *Organizational Learning II. Theory, Method, and Practice*. New York: Addison-Wesley.

Ballard, G. (1999). "Improving Workflow Reliability." In: *Proc. 7th. Ann. Conf. of the Int'l Group for Lean Construction*, Berkeley, California, USA. July 26-28.

Brady, D., Tzortzopoulos, P. and Rooke, J. (2013). "The Development of an Evaluation Framework based on the Design Science Approach." In: *Proc. 21st. Ann. Conf. of the Int'l Group for Lean Construction*, Fortaleza, Brazil. July 29 –August 2.

Bresnen, M. and Marshall, N. (2001). "Understanding the Diffusion and Application of New

Management Ideas in Construction." *Engineering, Construction and Architectural Management*, 8(5/6), 335-345.
Coghlan, D. and Davis, K. (2006). "Facilitating Practitioner Research." *Revue Sciences de Gestion*, 50, 19-32.
Coghlan, D. & Brannick, T. (2010). *Doing Action Research in Your Own Organization*. (3rd. ed.). London: Sage.
Cooper, R. (2001). *Winning at New Products: Accelerating the Process from Idea to Launch.* Cambridge, MA: Perseus Publishing.
Deming, W. (1986). *Out of the Crisis: Quality, Productivity and Competitive Position*. Cambridge: Cambridge University Press.
Dick, B. (2002). "Action Research: Action and Research." Available on line at <http://www.aral.com.au/resources/aandr.html> (15 Feb 2016).
Dick, B. (2014). "Validity, Reliability, Generalizability." In D. Coghlan and M. Brydon-Miller (Eds.) *Sage encyclopedia of action research*. London: Sage.
Golafshani, N. (2003). "Understanding Reliability and Validity in Qualitative Research." *The Qualitative Report*, 8(4), 597-607.
Iivari, J. and Venable, J. (2009). "Action Research and Design Science Research – Seemingly Similar but Decisively Dissimilar." *Special Issue of MIS Quarterly*.
Kagioglou, M., Cooper, R., Aouad, G., Hinks, J., Sexton, M. and Sheath, D. (1998). *Final Report: Generic Design and Construction Process Protocol*. University of Salford, UK.
Lewin, K. (1946). "Action Research and Minority Problems." *Journal of Social Issues*, 2(4), 34-36.
Lincoln, Y. (1995). "Emerging Criteria for Quality in Qualitative and Interpretive Research." *Qualitative Inquiry*, 1(3), 275-289.
March S. and G. Smith. (1995). Design and Natural Science Research on Information Technology. *Decision Support Systems, 15* (4), 251-266.
Miles, M. and Huberman, A. (1994). *Qualitative Data Analysis: An Expanded Sourcebook* (2nd. ed.). Thousand Oaks, CA: Sage.
O'Brien, R., 2001. "An Overview of the Methodological Approach of Action Research." In: Roberto Richardson, (Ed.) *Theory and Practice of Action Research*, Universidade Federal da Paraíba, João Pessoa, Brazil.
Patton, M. (1990), *Qualitative Evaluation and Research Methods* (2nd ed.), Thousand Oaks, CA: Sage.
Reason, P. and Bradbury, H. (Eds). (2008). *Handbook of Action Research: Participative Inquiry & Practice.* (2nd. ed.). London, UK: Sage Publications.
Remenyi, D., Swartz, E., Money, A. and Williams, B. (1998). *Doing research in business and management: An Introduction to Process and Method*. London, UK: Sage Publications.
Robson, C. (2002). *Real World Research: A Resource for Social Scientists and Practitioner Researchers*. (2nd.). Oxford: Blackwell.
Shewhart, W. and Deming, W. (1939). *Statistical Method from the Viewpoint of Quality Control*. Washington, D. C.: The Graduate School, the Department of Agriculture.
Smith, R. & Morrow, J. (1999). "Product development process modelling." *Design Studies*, 20, 237-261.
Tzortzopoulos, P. (2004). *The Design and Implementation of Product Development Models in Construction Companies.* PhD Thesis, School of the Built Environment, University of Salford, Salford, UK.
Yin, R. (2009). *Case study research: Design and Methods*. (4th. ed.). Thousand Oaks, CA: Sage