BUILDING AN ARGUMENT WITH HYPOTHESIS TESTING

SeongKyun Cho[1], Glenn Ballard[2]

ABSTRACT

This paper presents a research methodology for testing the hypothesis, “The implementation of Lean Construction improves project performance”, and for supporting a recommendation that South Korea embrace Lean Construction. To meet general quantitative research requirements, the methodology will include hypothesis testing; measurement of variables in the hypothesis, a large N sampling and a small N case selection strategy, and interpretation of the findings. The results of executing the methodology will be published in future papers.

KEY WORDS

Market competitiveness, Quantitative measurement, Statistical analysis, Lean Construction, South Korea

INTRODUCTION

South Korea invests a high percentage of its GDP in construction; 18.4% in 2004, 18.0% in 2005, 17.7% in 2006, 17.5% in 2007, 18.0% in 2008, and 18.4% in 2009 (Construction Association of Korea: CAK, 2009) . According to the press release made by the Construction & Economy Research Institute of Korea on March 19th 2010[3], this percentage will be likely to decrease to 11 to 11.5% in 2020 based on the trend of the construction sector in other well advanced nations. Construction is clearly vital to the national economy. If there are systemic factors causing inefficiency, the nation’s economy might have a big problem.

However, industry leaders (Lee et al, 2003) have criticized the government for not selecting the most competent contractors, which can reduce overall market competitiveness of the industry. International comparisons also raise concerns. According to ENR, in 2006, the last year for which data is currently available, South Korea’s share of the international construction market, excluding engineering, was 2.9% while its share of engineering is much less than that, 1.9%. Contrary to this, the United States, having been number one since 1997, had 17.1% in construction and 42.1% in engineering in 2006 (MLTM, 2007). We might reasonably conclude that South Korea’s construction industry depends primarily on labor and labor management rather than engineering expertise.

Lean Construction is a philosophy of organizational management characterized by pursuit of an ideal: to provide customers products and services exactly fit for purpose, within customer conditions of satisfaction, with no waste (Ballard et al,

[1] MS, Civil and Env. Engineering. Department, 407-A McLaughlin Hall, Univ. of California, Berkeley, CA 94720-1712, USA, Phone +1 510/725-7929, seongKyuncho@berKeley.edu

[2] Director, Project Production System Laboratory, http://p2sl.berKeley.edu, and Adjunct Associate Professor , Civil and Env. Engineering. Department, 215-A McLaughlin Hall, Univ. of California, Berkeley, CA 94720-1712, USA, Phone +1 415/710-5531, ballard@ce.berKeley.edu

[3] See http://www.cerik.re.kr/

2007). In theory, acting on the Lean Construction philosophy should improve project performance. Should South Korea embrace Lean Construction in order to improve the performance of its construction industry?

That policy recommendation would be well supported if it could be shown that Lean Construction projects do in fact perform better than non-Lean Construction projects, that Lean Construction practices are not widely[4] or well implemented in South Korea, and that there is room for improvement in the performance of South Korean construction projects. The challenge to be overcome is the lack of available data on Lean Construction projects. This paper proposes a research methodology that overcomes that challenge.

The research hypothesis is first presented, then the research methodology, followed by conclusions. The results of applying the methodology will be presented in future papers.

RESEARCH HYPOTHESIS

Our research questions are:

1. Has Lean Construction improved the performance of projects?
2. Does the construction industry of South Korea lack implementation of Lean Construction?
3. Is the performance of the South Korean construction industry worse than those which employ Lean Construction more?

These questions can be answered by testing the hypothesis: *The implementation of Lean Construction in a project will improve the performance of the project.* The independent variable of this hypothesis is the degree of implementation of Lean Construction in a project and the dependent variable is performance of that project. If we can support the hypothesis, that answers question (1). Measurement of the independent variable in the hypothesis, together with evaluation of management practice in South Korean projects, will be the answer to question (2). Comparison of performance (the dependent variable) between South Korean and U.S. projects that employ Lean Construction more, will give us the answer to question (3).

RESEARCH METHODOLOGY

Our research design is non-experimental with only post observation (Trochim, 2006), which has a basic weakness in maintaining internal validity[4]. Lack of internal validity can expose our research to various threats, such as selection bias caused by self selection and historical bias caused by external historical events (Hoyle et al., 2002).

Cause and effect thinking and hypothesis testing usually take on quantitative research based on Large N (number of cases) Analysis (LNA) while qualitative research seeks meanings constructed individually, socially, or historically with the intent of developing a theory or patterns (usually before hypothesis testing) based on Small N Analysis (SNA) (Crenswell, 2003). We propose to mix these two research methodologies, following Lieberman (2005).

[4] Strong internal validity requires random assignment of cases to different independent variables (Bernard, 2000).

Definition of Terms

- Hypothesis testing: Hypotheses propose a relationship between an independent variable(s) and dependent variables. Hypothesis testing evaluates the proposed relationship, typically using statistical techniques such as correlation analysis.
- LNA: Analysis of a large number of cases such as a survey of 1,000 persons.
- SNA: Analysis of a small number of cases selected based on some principles.
- Model building SNA: Analysis of a small number of cases deliberately selected to deviate far from the best fitting model (regression equation) suggested by the results of LNA hypothesis testing. This SNA is used to build a new model different from the previous best fitting model.
- Model testing SNA: Analysis of a small number of cases selected to best fit a model (regression equation) suggested by the results of LNA hypothesis testing. This is for supporting the best fitting model.

Lieberman's Research Design

Lieberman's mixed method (Lieberman, 2005) is as follows. An LNA is used to test a hypothesis. If the LNA gets satisfactory results, then an SNA is used for model testing. In the SNA, cases are selected that best support the hypothesis. If the SNA's results support the model suggested by the LNA, the hypothesis testing is successful. But if they do not, we need to determine why. If the previously selected cases were idiosyncratic, we can select other cases for another model testing SNA. If the cases are not idiosyncratic, we need to build a new model, which means the measurement of the hypothesis must be modified.

If the initial LNA hypothesis testing does not yield satisfactory results; i.e., the best fit regression line is not a good fit with actual data points, a model building SNA is used to create a modified regression line. In that case, the selection of cases would be done on three strategies: 1) selection among singularities, 2) selection among similar and dissimilar outcomes similar independent variables, and 3) selection among similar and dissimilar independent variables at similar outcomes. If this SNA supports the new model, we can decide if we do a new LNA for testing the new model. However, if the SNA fails to support the new hypothesis, then that is the end point of the research.

Our Research Design

The unique feature of our research is the use of two LNAs with the same hypothesis. The first LNA, on the projects outside of South Korea that employ Lean Construction better than those in South Korea, would test the hypothesis purely based on Lean Construction theory. This LNA will produce a model, which best explains the relation between independent variable and dependent variable of the hypothesis; i.e., the extent of Lean implementation and project outcomes. A best fit regression line created from the survey data points is then produced. If the degree of fit is acceptable, we will do a model testing SNA to strengthen the model. If the degree of fit is not acceptable, we will do a model building SNA to specification of variables, perhaps simply by deleting survey questions, hopefully yielding a better fit regression line; in other words, a model better explaining the relationship of the variables in the hypothesis.

Then we will do the second LNA on the Korean construction industry with the advanced measurement but with the same hypothesis. The cycle between LNA and SNA is repeated again here. As the final result, we would come to get the best model to support the hypothesis with the best measurement. The first LNA's roles are (1) to modify the original measurement and deliver the modified version to the second LNA; i.e., to determine what data is needed to best reveal the relationship between variables in the hypothesis, and (2) to produce a model explaining the relation between lean implementation and performance on the selectively chosen cases outside of Korea, and (3) to provide reference information to locate the situation of South Korea in case there is little information with which to diagnose the degree of lean implementation. The detail flow chart is shown on Figure 1.

Sampling

The next issue is sampling strategy. A randomized sampling is the only way to be confident about generalizing (Hoyle et al., 2000). The 2nd LNA, on South Korea's construction industry, would use randomized sampling. But the 1st LNA, on non-Korean projects, does not need to use it because its purpose is to structure the detailed expression of the hypothesis, and to provide reference information to the LNA on South Korean projects. It would use expert sources to gather cases that have more complete implementation of Lean construction, drawing on sources such as the University of California, Berkeley's Project Production Systems Laboratory (p2sl.berkeley.edu), the Lean Construction Institute (www.leanconstruction.org), and the International Group for Lean Construction (www.iglc.net). Even though there is little variation in lean implementation on Korean projects, the result of the LNA on this purposeful sampling would give us insight to compare Korean projects and Lean projects in terms of Lean implementation.

Contrary to this, the 2nd LNA would use randomized sampling. The web site operated by the Construction Association of Korea (CAK) contains a list of names and relevant personnel of 5,880 completed projects over 10 million dollars. Before selecting cases randomly, we decided to stratify the population to reflect unique features. First, we assumed that whether the owner of a project is private sector should have a relatively big impact on performance (the dependent variable) because private owners will make every effort to increase the efficiency of their money, while public sector owners will try not to violate contractual or legal constraints while spending the government's money. And second, we assumed that whether a project's contract separates design from construction has a consequential result in performance of the project based on Lean Construction's use of collaboration to deliver value and eliminate waste. These assumptions, which will tested in the 2nd LNA, are the criteria of stratification so that we have four strata, public-separating (1,958 cases), public-combining (312), private-separating (3,409), and private-combining (the others). Then, we will select 100 projects from each stratum randomly[5].

[5] According to Sudman (1983), the minimum number of cases per strata is about 20

Measurement

The next issue is measurement of variables. According to Adcock et al. (2001), measurement is comprised of conceptualization, which is formatting a systemized concept[6] through reasoning about a background concept (in our case, it is independent /dependent variable)[7] in the light of goals of research; operationalization, which is developing, on the basis of the systemized concept, one or more indicators; and scoring cases, which is applying these indicators to score the projects being analyzed. Operational definition for scoring indicators of our research is on-line surveys (LNA). In fact, measurement of the independent variable, implementation of Lean Construction, has been undertaken by several researchers, including Liker (2004), Womack & Jones (1996), Hofacker et al. (2000), and Ballard et al. (2007). However, their principles are too abstract to perform large N analysis or have many interdependencies among their indicators that can cause multi-collinearity, in which case none of the independent variables predict significantly well the dependent variables even though the overall model fits well the data (Motulsky, 1995). Thus, we decided to make our own measurement of the independent variable, the implementation of Lean Construction, as Table 1, of which systemized concepts are: 1) incentives based on team performance; 2) innovation beyond current best practice; and 3) reduction of variation in executing plans. Each systemized concept has more specific indicators.

The purpose of these indicators is to measure projects whose participants may not know Lean theory well. Consequently, we had to deconstruct currently used lean tools, such as A3, PDCA, value stream analysis, real time estimation, target value design, Last Planner and so on (cf. Thomson et al, 2009) into understandable and independent indicators. Also, we don't believe specific technologies such as BIM or A3 reports are appropriate indicators. We would assign a low score to a strictly hierarchical organization using BIM at current best practice. PDCA, value stream analysis, and target costing are sufficiently measured by the indicators in Table 1.

For easier understanding of Table 1, it is important to understand how to score a project. One of the systemized concepts, "Incentives based on team performance", has eight indicators. There are several questions to measure each indicator in our survey. The respondents would give answers to the questions so that the score of each indicator is determined. The summation of the eight indicator's scores would be the systemized concept's score. The other systemized concepts' scores are determined in the same way. Finally, the summation of the score of each systemized concept would yield the total score of the project, the independent variable's numeric value. This is the original measurement, which could be enhanced or modified by the aforementioned research design.

[6] A specific formulation of a concept used by a given scholar or group of scholars
[7] The broad constellation of meanings associated with a given concept

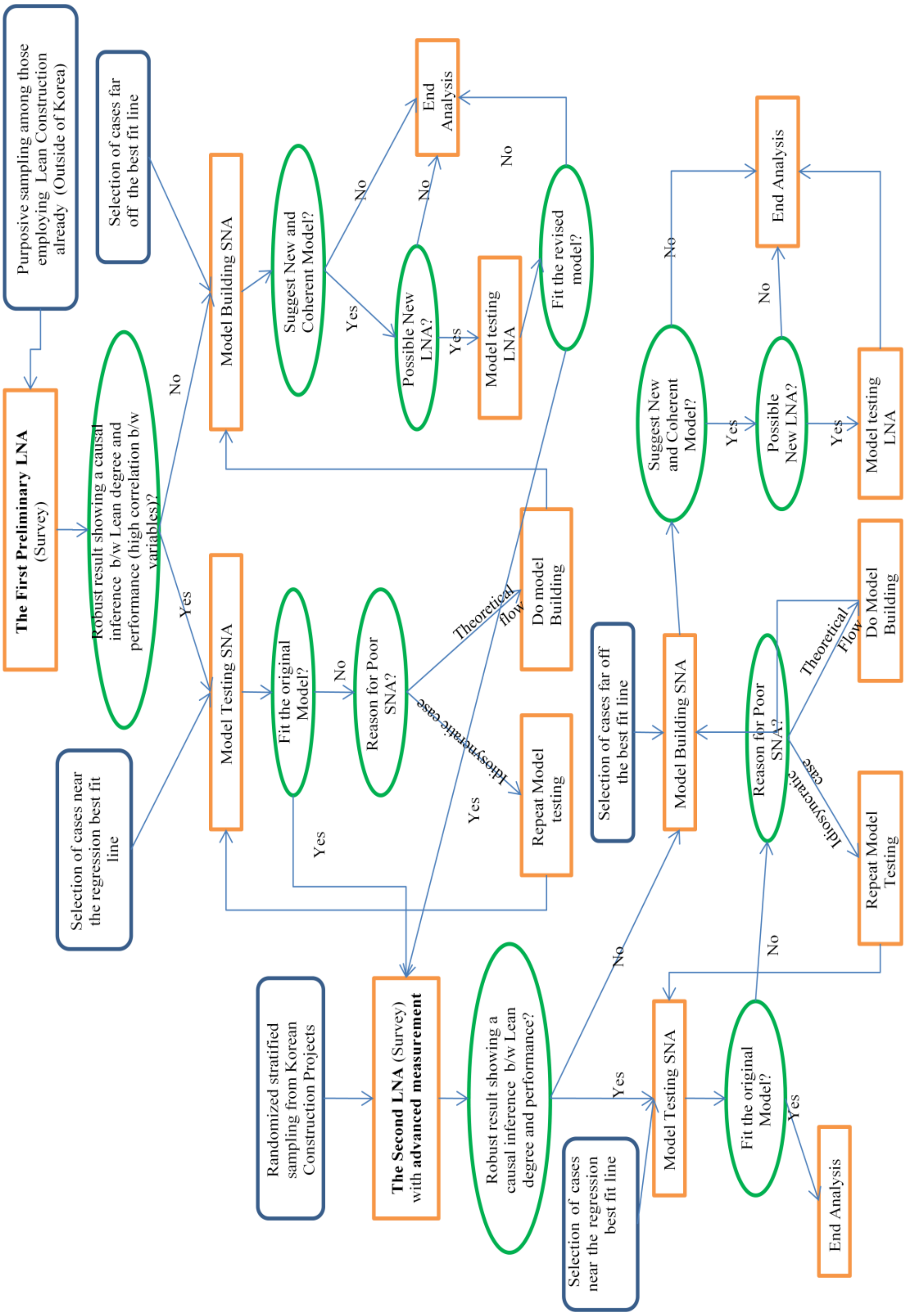

Figure 1: Research Design, a modified version of Figure 1 in Lieberman (2005)

Compared to the measurement of the independent variable, measuring the dependent variable is easier because many articles have defined key performance indicators. CII (2003) defines cost performance, schedule performance, incidence rate, and so on, as key performance indicators. UK's *Rethinking Construction* set reduction of construction cost, time, and accidents; and increase of profitability, productivity, and predictability, as key performance indicators (CTF, 1998). The Danish benchmarking system defined duration performance, change of cost and unit price, accident frequency, and defects during construction and after handing over, as key performance indicators (Cheung et al., 2004). Similar indicators have been developed by Cheung et al (2004) and Chan et al (2004). Based on this previous research, we set cost change, unit cost, duration change, remediation of defects during one year after handing over, safety accidents during construction, and overall subjective satisfaction, as our performance indicators.

Of course, these indicators are of the type 'conforming to plan' rather than of type 'exceeding current benchmarks'. Every project must conform to its plan that is made based on agreements by parties at some points. Thus, our indicators are minimal conditions, which a project should have in view of market competitiveness.

We made online five different surveys from the indicators in Table 1 as well as performance more specifically according to respondent (owner, contractor, or architect) and project type (contractually separated, or combined between design and construction). Even though indicators in Table 1 look simple, making questions from them must take generally acceptable survey form. For example, to measure participation of contractors in setting expected cost, we included two questions in the contractor type's survey. We first asked "Did the main contractor participate in the owner's estimation of expected cost?". If the answer is Yes, the contractors are then asked "Of all speciality contractors, what percentage provided inputs relevant to setting the expected cost?" This answer could be 0, 0-25, 25-50, 50-75, 75-100, 100%[8]. In addition, we add more questions about needs for institutional changes and about details on types of commercial contracts to make analysis easier. The survey questions were tested by several persons named in acknowledgements in facial validation tests, investigating whether the measurement measures what its name suggests (Hoyle et al, 2000). By inserting opposite concepts in the survey, we made discriminant validation (Trochim, 2006) test possible during data analysis.

CONCLUSIONS

The result of this research will reveal whether Lean Construction implementation is likely to have desirable impacts on performance of South Korean construction projects, and thus provide support for policy recommendations.

ACKNOWLEDGMENTS

We would like to thank Professor Laura Stoker of UC Berkeley for her tremendous help with the research design, operationalization, and validation test of measurement. Professor Coye Cheshire, PhD student Farook Hamzeh, PhD student Hung Nguyen, and PhD student Hyunwoo Lee gave expert reviews on our survey questions.

[8] You can see the complete form of the survey at http://www.surveymonKey.com/s/9QYQDV6 until June, 2011.

Table 1: Indicators of each systemized concepts of the independent variable

Systemized concepts	Indicators (reference)
Incentives based on team performance	Performance assessment after completion and unrestricted publication (Cheung et al., 2004); Performance-based selection of contractor and architect Main contractor develop cost and duration estimates based on current best practice (Ballard, 2006); Main and specialty contractors participate in setting the expected cost/duration (Ballard, 2006); Main and specialty contractors' participate in product design (Saunders, et al., 2005); Targets for cost and duration set as stretch goals to spur innovation (Sakal, 2005); Main and specialty contractors' participate in setting the target cost/duration (Sakal, 2005); All project team members share profit and loss based on a pre-agreed distribution, which is decided unanimously (Sakal, 2005);
Innovation; going beyond best practice	Allocation of target scopes and cost/duration to cross functional teams by facility system to create the design organization (Ballard, 2000-a); Design specialists' consulting other relevant design specialists (Ballard, 2000-a); Design specialists' consulting all relevant contractors (Ballard, 2000-a); Design specialists review all possible alternatives including other specialists' suggestions (Ballard, 2000-a); Design specialists eliminate design alternatives based on constraints, including cost/duration targets (Ballard, 2000-a); Concurrent product and process design; All relevant contractors' participate in process design (Ballard 2006) Preassembly in process design (Tsao et al., 2001); Minimizing batch sizes of materials or facilities in process designing (Arbulu et al., 2002); Inventory management in process designing (Walsh et al., 2004); Standardization in process designing (Tommelein, 2006);

Systemized concepts	Indicators (reference)
Reduction of variability in executing plans (Last planner)	Each worker investigates the readiness of the next workers before execution; pull (Tommelein, 1998); Existence of communication channels between adjacent processes (Tommelein, 1998); Each front line supervisor does constraint analysis (prerequisite work, contractual approvals, sequence, resource, duration, funds, weather conditions, labor & equip, pseudo works like an experiment, and so on) to make scheduled tasks ready to be performed when scheduled, 6 weeks or so before executing its works (Ballard, 2000-b); Removing all causes of the past weekly plans' failure before the next similar plans' execution (Ballard, 2000-b); Each front line supervisor investigates, before executing their tasks, if the causes of the past plans' failures have been removed. (Ballard, 2000-b) Definition of all handoffs in a phase between work groups, in terms of sequence, duration and other constraints, in phase planning (Ballard et al., 2003); Everyone responsible for finding and correcting defects in higher level work plans (master schedules, phase schedules, lookahead plans, etc.). (Ballard et al., 2003);

REFERENCES

Arbulu, R., Tommelein, I., Walsh, K., and Hershauer, J. (2002), "Contributors to Lead time in Construction Supply Chains: Case of Pipe Supports used in power Plants." *Proceedings of the 2002 Winter Simulation Conference.* Sandiego, CA, USA, Dec 8-11

Ballard, G. (2000-a). "Positive vs. Negative Iteration in design". *8th IGLC proceedings*, Available at www.leanconstruction.org [Retrieved Mar 26th 2010]

Ballard, G. (2000-b). "The last planner system of production control" PhD. Diss. School of Civil Eng., Univ. of Birmingham , UK, Available at www.leanconstruction.org [Retrieved Mar 26th 2010]

Ballard, G. (2006). "Rethinking project definition in terms of target costing". *Lean construction journal*, 2008, pp. 1-19

Ballard, G., and Howell, G. (2003). "An update on last planner." *Proceedings of the IGLC 11th* , BlacKsburg, VA, USA, July 2003, pp. 329-341

Ballard G., Yongwoo Kim, Jang J., and Liu L. (2007, Oct). "Roadmap for Lean implementation at the project level." 1-109, Construction industry institute research report 234-11. p. 140

Bernard, R. (2000), *The social research methods-qualitative and quantitative approach*, London: Sage publication, Ch 2, 4, 5, and 8

CII Web site (2003), CII Best Practices, Available at www.construction-institute.org [Retrieved Feb 17th 2010]

Cheung, S., Suen, H., and Cheung,.K (2004). "PPMS-a Web based construction Project." *J.of Automation in construction* 13 (2004) pp. 361-376, May issue

Construction Association of K (CAK, 2009), "The statistics of the construction industry." Available at: www.caK.or.Kr [Retrieved July 11, 2009]

Crenswell, W. J. (2003), *Research Design: qualitative, quantitative, and mixed approaches,* Second edition, Thousand Oaks, CA, USA., SAGE publication, pp. 4-23

Hoyle, R., Harris, M., and Judo, C. (2002), *Research methods in social relations*, US: Thomson Learning.

Lee, G., Kim, S., Kim, S., and Cho, D. (2008), "Banking statistics of Korea in 2003", Available at http://news.sdi.re.Kr [Retrieved July 11th 2009]

Lee, S., and Hanmi Parsons (2003), *the 'K's construction industry*, Bosunggak, Seoul, pp.17-18.

Lieberman, E. (2005), "Nested analysis as a mixed method strategy for comparative research", *J. of American political science review*, 99 (3) 435-452, August issue

Liker, J. (2004), "The Toyota Way: Fourteen Management Principles from the World's Greatest Manufacturer". NY: McGraw Hill. Pp. 37- 41.

Ministry of Land, Transport, and Marine affairs of Korea (MLTM, 2007) "The fourth master plan for developing construction technology", Available at www.mltm.go.Kr [Retrieved Mar 13th 2010]

Motulsky, Harvey (1995). "Multicollinearity in multiple regression." [On line] Available at www.graphpad..com [Retrieved Feb 11th 2010]

National Building and Construction Committee, Science and resources (NBCC, 1999), "Building for growth: building and construction industry agenda", Available at http://www.revaluing-construction.com [Retrieved Mar 16th 2010]

Sakal, W. Matthew (2005). "Project alliancing: A relational contracting mechanism for dynamic project". *Lean construction journal, 2 (1),* pp. 67-79

Saunders, K. and Mosey, D. (2005). "PPC 2000: Association of consultant Architects standard form of Project Partnering Contract." *Lean Construction Journal, 2 (1)*. 62-66

Sudman, Seymour (1983), "Applied sampling". In Handbook of Survey Research, eds Peter Rossi, James Wright, and Andy Anderson. New York: Academic Press, CA: Sage. Ch. 2.

Thomsen, Chuck, Joel Darrington, Dennis Dunne, and Will Lichtig (2009), Managing Integrated Project Delivery, available at http://cmaanet.org [Retrieved May 30, 2010]

Tommelein I. D. (2006). "Process Benefits from use of standard product - simulation experiments using the pipe spool model." *Proceedings of IGLC 14th*, 25-27 July, Santiago, Chile, pp. 315-319

Tommelein, I. D. (1998). "Pull Driven Scheduling For Pipe-Spool Installation: Simulation of a Lean Construction Technique." *ASCE, J. of Const. Eng. & Management., 124 (4)*, 279-288

Trochim, W. (2006). "Research methods knowledge base." Available at http://socialresearchmethods.net [Retrieved Mar 22nd 2010]

Tsao, C. and Tommelein, I. D. (2001). "Integrated Product-Process development by Light Fixture Manufacturer." *IGLC 9,* 6-8 August, Singapore

Walsh, K., Hershauer, J., Tommelein, I, and Walsh, T. (2004). Strategic Positioning of Inventory to match Demand in a Capital Projects Supply Chain. *Journal of Construction Engineering and Management*, *130 (6)*, 818-826

Womack and Jones. (1996). "Principles of Lean cited from Lean thinking." Available at www.lean.org [Retrieved July 11th 2009]