Systems engineering return on investment

by

Eric C. Honour

BSSE, MSEE

A thesis submitted for the degree of

Doctor of Philosophy

Defence and Systems Institute

School of Electrical and Information Engineering

University of South Australia

This file is an extract from the full thesis, containing only the executive summary and the Conclusions chapter. For a full copy, please contact the author at ehonour@hcode.com

January 2013

Summary

This Systems Engineering Return on Investment (SE-ROI) research project explored the quantifiable relationships between systems engineering (SE) activities and program success. The work discovered statistically significant relationships between SE activities and three success measures: cost compliance, schedule compliance, and stakeholder overall success. SE-ROI is discovered to be as high as 7:1 for programs with little SE effort and 3½:1 for median programs. Optimum SE effort for median programs is 14.4% of total program cost; the work provides an *a priori* estimation method to determine this optimum for specific programs based on 14 characterization parameters. These findings address a significant state-of-the-art gap in that SE effort levels have typically been based on subjective heuristics rather than quantified success parameters.

The research developed an interview methodology and interview instruments to obtain a rich set of data from completed programs. The research was supported by a Research Advisory Group of over 60 international members who evaluated the research plans, methods, and instruments during development, ensuring a robust research approach. Program interviews were performed on 51 completed programs in 16 organizations. Interview participants were the Program Manager and Lead Systems Engineer of each program. Programs were from a wide variety of both contracted and amortized development domains; had a wide range of cost, schedule and success; and evidenced SE effort level from near-nil to large.

Through the use of Principal Component Analysis and a hill-climbing search for best correlation, the program SE effort levels were adjusted for the specific program characteristics, increasing correlations from $R^2 = 14\%$ to as much as $R^2 = 80\%$. The high degree of resultant correlation indicates that the appropriately weighted program characterization parameters largely remove the confounding factors that usually obscure the relationship between SE effort and program success.

The resulting relationships show that all SE activities correlate significantly with cost compliance, nearly all SE activities correlate with schedule compliance, and most SE activities correlate with stakeholder overall success. There is some indication of causality in qualitative and theoretical factors. While not definitive, the implication is that the level of selected SE effort is causative of the program success. If true, then the

use of the SE effort estimation method herein would result in the best available program success.

Some additional findings are also presented. The data shows no significant correlation between SE effort levels and system technical quality. There is indication that this lack of correlation is due to program emphasis on requirement thresholds rather than on stakeholder-defined technical quality. Optimizing technical leadership/management levels is shown to provide a unique benefit in simultaneously associating with cost compliance, schedule compliance, and overall stakeholder success. The work also contributed to the discovery of a commonly held SE ontology that could be expressed in eight SE activities. The worth of this ontology was evident in its easy understanding by all interview participants.

Acknowledgements

This work would not have been possible without the participation, encouragement, and wisdom provided by many.

The University of South Australia (UniSA) has played a significant role in the completion of this work. While the original three-phase research plan was created in 1997, its progress was slow and difficult until UniSA graciously offered its sponsorship of Phase III as a doctoral candidacy. The 'official' nature of the research from that point served two primary purposes: (a) it provided deadlines and impetus to progress, to act in opposition to the pull of daily business that often held the research back, and (b) it opened doors to other advisors and for programs to interview that were not open to an independent researcher. In particular, I wish to acknowledge Prof. Stephen Cook, who twisted my arm to formalize this relationship and who reluctantly ended up as primary supervisor; A/Prof. Joseph Kasser, primary supervisor during the formative stages, who helped to keep the scope within control; and A/Prof. Timothy Ferris, whose detailed reviews were always insightful and helpful. Other researchers and academics on staff at UniSA Defence and Systems Institute (DASI) often helped with encouragement, review, contacts, and tidbits of essential knowledge. In addition, the entire DASI administrative staff was a delight to work with, with a special note for the selfless and always-present help of Dale Perin.

Untold numbers of friends in the International Council on Systems Engineering (INCOSE) provided special help, encouragement, and knowledge to this effort. I can single out Dr. George Friedman, Dr. Elliot Axelband, and Dr. Azad Madni for their constant and necessary push toward successful completion. Others, who preceded me through the immense work to complete doctorates, gave me a helpful pull forward, including specifically Dr. Ricardo Valerdi and Dr. Sarah Sheard. I also thank Dr. Bill Ewald for his encouragement and always-positive attitude toward my work. Particular thanks go to Dr. Brian Mar, who worked with me and encouraged me in the earlier phases but has since gone on to a better world.

The SE-ROI Research Advisory Group participants, over 60 strong, provided significant help by reviewing the formulation of the research plans and methods, with many timely comments and changes to improve the effort. They also often provided the necessary contacts to obtain the program interviews.

And finally, I cannot express sufficiently my gratitude to the unnamed organizations that participated in interviews, and the individuals from those organizations who were interviewed. The work could not have been done without the generous access you allowed me to your programs through proprietary boundaries. I hope that this work fulfills your expectations, and that the results advance the discipline of systems engineering to your benefit and the benefit of mankind.

This work is dedicated to

- *My wife, Beth, whose unstinting support over 15 years has encouraged, cautioned and prodded this success,*
- *The systems engineering discipline, making the world better through technology,*

and mostly to

- *God, who gave me the skills and reasons to perform it.*

7 Conclusions and recommendations

This chapter summarizes the thesis by highlighting the major findings and indicating future research areas that come out of this work.

7.1 Major findings

The work described in this thesis developed and executed a research program to explore the quantitative relationships between systems engineering and program success. The program created an interview instrument, based on a peer-reviewed ontology, which could gather salient data for the research. Using that instrument, the principal investigator obtained 1.5-hour interviews with the program management and systems engineering leaders of 51 system development projects. The interview data was subjected to rigorous mathematical and statistical processing to extract and test the relationships.

All of the following major findings are supported by the research work. The first six findings are of highest importance to the SE discipline as a whole, and to the system development programs that use SE.

1. There is a quantifiable relationship between systems engineering effort levels and program success.

The original research question A was stated as

- (RQ_A) **Is there a quantifiable correlation between the amount, types and quality of systems engineering efforts used during a program and the success of the program?**

The empirical work in this research has shown that there exists such a quantifiable relationship, with correlation coefficients well in excess of the test values for a significance level of $\alpha = 0.05$.

The relationships for total SE effort are shown in Figure 37, Figure 38, and Figure 39, all of which show program success measures (cost compliance, schedule compliance, and overall success, respectively) plotted against Equivalent SE Effort (*ESEE*) as a percentage of program cost. In all three cases, the relationship shows poor program success at low levels of *ESEE*, improving to desirable program success at moderate levels of *ESEE*, and then again deteriorating to poor program success at higher levels of *ESEE*. This relationship is consistent with the theoretical work of Honour (2002b).

Similar relationships were discovered, with minor exceptions, for each of the eight subordinate SE activities. Table 33 summarized the findings that are depicted in Figure 42 through Figure 73. All eight SE activities demonstrated a statistically significant correlation against at least two of the three program success measures. All SE activities were significantly correlated with cost compliance; seven activities were significant against schedule compliance, and five activities were significant against overall success. Given the demanding significance level of $\alpha = 0.05$, these results provide an overwhelming affirmation of this major finding.

<u>Importance</u>. This major finding acts as a warning to programs: the level of SE effort matters to the success of programs, as does the mix of that effort across the constituent activities of SE.

2. Systems engineering has a significant, quantifiable Return on Investment.

The quantified relationship between cost compliance and Total SE effort was subjected to standard financial calculations for Return on Investment, in which the Return was measured as program cost reduction and the Investment was measured as additional cost applied to Total SE effort. The results (Section 5.2.1.1) clearly showed a quantifiable ROI for the re-allocation of funding to additional SE effort. For programs operating at near-nil SE effort, that ROI is as high as 7:1, a program cost reduction seven times as great as the added SE cost. For programs operating near the median of the interviewed programs, the ROI is 3.5:1.

<u>Importance</u>. The result is compelling program management information: greater SE effort is associated with programs at significantly less cost overrun. This statement is true for programs operating at and below the median of SE effort.

3. No correlation was found between systems engineering and system technical quality.

The research work included four success measures, not only three. The fourth success measure was intended to quantify the technical quality of the resulting system, using the Key Performance Parameters (KPP) believed to matter most to the stakeholders. The statistical work unfortunately showed no significant correlation between any SE activity and the system technical quality. (See Table 33.)

A review of the qualitative interview information provided an explanation, in that the program leaders (PM and SE) rarely used KPPs as a driving factor in the program. Instead, the driving technical factor was usually the set of requirements defined for the development. Programs typically excelled at KPPs only when either (a) improving the KPPs carried no additional effort or cost, or (b) technical excellence was the program driver.

Importance. This major finding is a caution for the SE discipline, to take care lest SE become merely an adjunct of program management. The role of SE in a project is to monitor and guide the technical success. Today, it appears that technical requirements are the *de facto* measure of technical success, rather than the technical qualities that matter to the stakeholders. This definition is attractive to program managers and contracts, but does not produce the best systems.

4. There is an optimum amount of systems engineering for best program success.

The quantifiable correlations between SE effort and program success all evidence a 'bathtub' behavior in which there is a clear optimum value of SE effort in each relationship. That optimum has been calculated by determining the point at which ROI goes to zero; up to this point, additional application of SE effort has a positive effect on the program, while beyond this point the ROI becomes negative.

For Total SE, the optimum amount of effort for a median program is 14.4% of the total program cost. For non-median programs, this value can vary roughly between 8% and 19% of the total program based on the varying program characteristics.

Similar values exist for each of the eight subordinate SE activities, with similar variation for non-median programs. Optimal effort values for the median programs are shown in Table 34 and range from 1.3% to 3.9% of total program cost.

Importance. This major finding provides a useful metric for program development and for program reviews because it shows the typical levels of SE effort that are associated with the most successful programs.

5. Programs typically use less systems engineering effort than is optimum for best success.

For the median of the interviewed programs, the calculated ROI is 3.5:1. This indicates that additional total SE effort would result in a program cost reduction 3.5 times as great as the cost of the additional effort. It is apparent from this data that the median programs in the interview set operated with considerably less SE effort than is optimum.

How much less can be seen in Table 34, which shows the median level of SE effort is 8.5% of total program cost against an optimum level of 14.4%. These numbers are important: for a median $14M program operating at 8.5% SE effort, the observed cost overrun was on the order of $1.5M; for a similar program using $200K greater SE effort, the cost overrun was only $1.0M. (Both numbers from the averaging line of Figure 37.) By allocating $200K greater SE effort, such a typical program could reduce its cost by $500K.

Importance. This major finding is a caution to program managers: ensure that their program is operating at a sufficient level of SE effort, a level higher than they are accustomed to using.

6. A method is provided to estimate the optimal systems engineering effort for a given set of program characterization parameters.

The second major research question was stated as

- (RQ_B) **For any given program, can an optimum amount, type and quality of systems engineering effort be predicted from the quantified correlations?**

This research has developed a systems engineering effort estimation method that is based on the quantified correlations. That method is described in 6.2.1 in five steps that use the mathematics of this research. The method is suitable for use in very early stages of a system development to provide meaningful effort estimation even for the initial SE activities. It has been shown that the method necessarily results in a level of SE effort that is optimum for program success for the given program characteristics.

There is no guarantee that following the estimations will result in the best success for a program. Other confounding factors still exist, and the causality has yet to be proven. However, the estimation provides SE effort levels that are proven to be associated with the most successful programs in this interview data set. This associative relationship is more useful than any prior method.

Importance. This major finding provides an effective program/technical management tool that can be used in system development programs to appropriately size the SE effort.

7. For systems engineering effort estimation, some program characterization parameters are of much greater importance than others.

The mathematics of the SE effort estimation method provides great insight into the factors that are of most importance. If these factors are mis-estimated, the SE effort will be poorly estimated. The importance of the factors occurs through two effects:

- The strength of their contribution to adjustment of SE effort from median to the specific program. More important parameters create a larger adjustment. This effect can be seen in Figure 35, Figure 36, and Table 37 that show the effort adjustment due to each factor.
- The effect of the factors on the correlation. More important parameters align the programs more closely to the central tendency, thereby removing the effects of many confounding issues and improving the estimation assurance. This effect can be seen in Figure 76 showing the change in correlation due to each factor.

The comparison of the data in these figures shows the factors in Table 41 to be most important for proper estimation of SE effort levels.

Table 41. Most important factors for SE effort estimation

Factor	**SE Effort Adjustment** ($Weight_j$, unitless)	**Correlation (Assurance) Improvement** (ΔR^2, unitless)
Level of Definition at Start	+53	0.25
Development Autonomy	+30	0.37
Level of Integration (system vs. subsystem)	-30	0.38
System Size	-25	0.20
Proof Difficulty	+22	0.10
Technology Risk	-19	0.12

Importance. This finding shows that it is necessary to take the important program factors into account when estimating program SE effort. Adjustment factors have been calculated to facilitate this.

8. Of the defined SE activities, technical leadership/management is unique in providing optimum program success simultaneously in cost, schedule, and stakeholder acceptance.

Section 6.1.9 observed the unique nature of the technical leadership/management (TM) effort as displayed in Figure 70, Figure 71 and Figure 72, that all programs in the optimal region of TM effort were largely on-cost, on-schedule, and had the highest levels of stakeholder overall success. No other SE activity claims this same level of success; in all other activities, some programs fell short of these desired goals. Therefore, it appears that using the optimal levels of TM effort is associated with this high assurance of meeting program goals.

There is a caution with this finding, however, in that the same graphs show severe degradation in both schedule compliance and overall success when the TM effort becomes greater than optimum. Providing optimal TM effort is important, but it is also important not to provide too much.

Importance. This major finding provides an indication to program managers in what area to most carefully emphasize proper SE effort.

9. There is a commonly held ontology of systems engineering that is sufficient to be meaningful.

The work of Honour & Valerdi (2006) demonstrated the confused nature of SE terminology by the difference of terms used in various SE standards for similar concepts. This confusion has existed for many years, due to the use of SE in various ‘stovepipe’ domains that do not trade information with others. One early concern for the current work was whether it was even possible to obtain comparable data from different organizations, due to the vast difference in terminologies and program structures.

The development of the eight SE activities described in Section 4.1.3 showed that a common vocabulary existed, indicating an underlying ontology (‘shared understanding’) was in fact sufficient to be meaningful. The use of those activities during interviews was invariably meaningful to the interview participants, and every interview was able to translate data from the source organization's definitions into the common ontology of the research.

Importance. This major finding is useful in the SE discipline to bring commonality and closure to many issues of conflict about terminology and concepts. It is noted that such closure is developing even during the period of this current work, with the development of a widely supported Systems Engineering Body of Knowledge (SEBOK).

10. It is possible to effectively quantify systems engineering effort using empirical data.

One dissenting view about SE quantification was published during Phase I of the prior 'Value of SE' work described in Section 2.3. This dissenting view was 'The Shangri-La of ROI' (Sheard/Miller 2000), in which the authors wrote

> This paper shows that 1) There are no 'hard numbers.' 2) There will be no hard numbers in the foreseeable future. 3) If there were hard numbers, there wouldn't be a way to apply them to your situation, and 4) If you did use such numbers, no one would believe you anyway.

The paper was accepted with acclaim in a time when systems engineers struggled with program managers for recognition of their worth. It was widely referenced for many years, partly because the paper echoed the despair among systems engineers, and partly because the paper also provided useful indications of subjective and political methods to build a case for systems engineering.

This current work has now proven that it is in fact possible to effectively quantify SE effort using empirical data. There is nothing magical or different about SE than other disciplines in this regard. The author has applied the same empirical methods used in many other disciplines to create the results herein.

Importance. This major finding removes a common excuse for not doing the necessary quantification. Programs can quantify their SE effort and can relate that to the success of the program.

11. It is possible to obtain meaningful data about systems engineering and success through program proprietary boundaries.

Sheard & Miller (2000) also expressed another objection to quantifying SE:

> Data such as productivity, cost to produce systems, process improvement cost, and the like is usually considered highly confidential. Everyone would be more than happy to receive reports on industry trends and data on all the other companies, but no one is willing to provide such data. Even if the data was clearly defined, additional confidential data would be required to verify that the data had the same meaning from company to company. It is hard to imagine ever prying such numbers out of a large number of companies in a manner that has any sort of fidelity.

This has been a commonly held thought by many, that the proprietary boundaries and the sensitivity of the necessary data would completely preclude obtaining meaningful data. Three works: Valerdi (2005), Elm (2008), and this thesis, show that this fear is not true. Proprietary boundaries can be bridged with care, and sensitive data can be controlled successfully.

Importance. This major finding helps to pave the way for future studies that also may need access to proprietary data.

7.2 Future research indications

As with all research, each step forward shines the light on the path beyond. The current work has provided indications of possible research in several different directions. Any of the following areas may bear great fruit for the future of the SE discipline.

Recent SE paradigms. As noted in the limitations of Section 6.4.2, the source data for the current work included only projects that were executed using traditional SE paradigms. This constraint had not been intentional; it was a simple result of the difficulty of obtaining willing programs and the nature of the programs that became available. No programs were explicitly using Model-Based Systems Engineering, Lean, or Agile methods.

A good further research project would be to apply similar empirical methods to programs using these more recent SE paradigms. Proponents of the newer paradigms make strong claims for their worth on programs (Kennedy & Umphress 2011) and the research method demonstrated herein can provide the information to validate those claims.

Other SE domains or cultures. Another limitation to external validity, also noted in Section 6.4.2, is that all programs interviewed were within Western cultures. All but four programs were within English-speaking countries (USA, Australia). Most of the interviewed programs were for military systems. Most of the interviewed programs were within a contracted environment.

A good additional research project would be to use the same empirical methods for system development programs in other system domains and other cultures, obtaining sufficient data to compare the differences with the work herein.

Validating current findings on subordinate SE activities. The current work has presented mathematically solid relationships between total SE and program success, and between subordinate SE activities and some forms of program success. In explicating the relationships, the data was transformed first through PCA and then through weights on the program characterization parameters. These transformations were often performed with minimal numbers of data elements, as noted in Section 6.4.1. This minimal approach was necessary due to the segmentation of data into the different areas, as well as the removal of outliers and constrained programs that was described in Section 5.1.9.3. The result is that many of the $Weight_j$ values are derived based on 15-dimension or 21-dimension searches with only a few more data points than the degrees of freedom.

Given the indications of the current work, a good advance may result from obtaining greater sets of data specifically on the subordinate SE activities. With larger data sets, the results herein may be further validated as well as extended into new knowledge. The values of $Weight_j$ can be corroborated or corrected as needed.

True measures of technical quality. This work has shown a surprising lack of correlation between SE activities and a well-recognized measure of technical quality. It became apparent during interviews and in the post-analysis that most of the interviewed programs treated the set of requirements as the only measure of technical quality. The Technical Performance Measurement method of KPPs was used by almost none of the interviewed programs, even though the method has been documented and known for decades.

A fruitful area of research would be to evaluate the accuracy of different measures of technical quality. True quality is always defined in the perceptions of the stakeholders. Are requirements generally an accurate representation of that perception? How often does a system meet requirements and yet is deficient in the stakeholder perception? What other technical quality measures exist, and how well do they reflect true quality?

Qualitative analysis of the interview data. During the current work, the SE-ROI interviews gathered a significant set of qualitative information. This information was not necessary for this thesis, but was gathered simply because obtaining program information is so difficult. A review of the interview instrument in Appendix Appendix B shows that the information includes:

- Selection of key performance parameters
- Other success measures
- Methods used for each SE activity, with indications of their success
- Tools used for each SE activity, with indications of their success
- Metrics used to evaluate each SE activity, with indications of their value
- Lessons learned on the programs
- Other descriptive information captured during the interview

A useful research project would be to collate this information against the program success measures to evaluate the qualitative factors that contribute to success. Through this type of research, causal factors may be revealed that are largely hidden through statistical work alone.

Best practices and leading indicators. 'Best practices' are those activities, methods, or tools that contribute to success under defined conditions. 'Leading indicators' are the measurable or perceivable characteristics of a program that provide advance indications of future success or failure. The current work has identified correlative relationships, but has not provided any indication of the best practices or leading indicators that accompany the relationships. Although it has shown the level of SE effort that is associated with success, this thesis has not obtained data to show the details of what SE methods are included within that level of effort, nor what effect those methods might have on success. It has also not developed leading indicators other than the appropriate levels of early SE activity.

It would be of tremendous benefit to the SE discipline to do empirical research to correlate both best practices and leading indicators with program success, so that programs can be guided in advance toward success.

Benchmarking the SE effort estimation method. The SE effort estimation method presented herein could possibly be an important tool for the SE discipline. The current widely used COSYSMO tool provides good management information for estimating

SE effort, but is based solely on the level of SE effort actually used in programs. This thesis has shown that the level actually used is only about 70% of the optimal level. This discrepancy is considerable, indicating that programs using COSYSMO continue to perpetuate the gap.

This thesis provides data based on 51 program interviews, but many of those were also extracted as outliers or constrained programs as noted in Section 5.1.9.3. The final results therefore used only 26-36 program interviews (70-80 total data points, if including the Value of SE data). While statistically significant, this is a small number of data points for a method that could drive program decisions affecting billions of dollars.

A good research avenue would be to use the SE effort estimation method for many trial programs, with follow-on tracking to determine its true accuracy and utility. This research might also consider merging this method with the COSYSMO method for even better benchmarking.

7.3 Summary

This research was concerned with the relationships between systems engineering and program success. The research specifically showed that relationships are evident in the data that significantly exceed the statistical bounds. The results allow the calculation of the cost-based Return on Investment for additional SE effort, showing that SE-ROI to be significant for programs in the median of those interviewed.

The findings from the research are that the relationship between SE effort level and program success demonstrates a distinctive optimum that can be calculated as 14.4% of total program cost for median programs. The median of interviewed programs operated at only 8.5%, significantly less than the optimum. For programs operating near the median of the interviewed programs, the ROI is 3.5:1. These findings can be considered to be typical of similar programs in similar domains and cultures. The implication of these findings for SE methodologies are to demonstrate that most programs operate with less SE effort than is optimum for program success. For most programs, adding SE effort can be expected to significantly reduce the total development cost.

In another finding, no significant correlation appears in the data between SE activities and system technical quality (as measured by stakeholder-based Key Performance Parameters). The implication of this finding is that current SE process definitions may be over-emphasizing the use of requirements as opposed to more basic technical measures of interest to the stakeholders. The conflict between KPP-based success and requirements-based success is evident.

The research also specifically supports a new SE effort estimation method based on optimizing program success based on an evaluation of program characterization parameters. The SE efforts that result from the estimate represent the SE effort level associated in this data with programs that had the best level of program success. All of these findings are applied both to total SE effort and to eight subordinate SE activities, with appropriate parameters and relationships for each activity. The effort estimation method provides values not only for the total SE effort but also for the eight activities. The implication of these findings is that this estimation method offers a more accurate level of SE effort estimation (for program success) than is currently available anywhere else.

As part of the SE effort estimation method, the research specifically shows that a small set of the program characterization parameters has the greatest impact on the effort estimation, indicating the program characteristics that most drive the level of SE effort. The implication of this finding for SE methodology is to indicate which program characteristics should be considered first in tailoring the SE methodology for a program.

In examining the eight subordinate SE activities, the research specifically shows that technical leadership/management provides a unique possible benefit in its optimal association with simultaneous success in cost, schedule, and stakeholder overall success. The implication of this finding for SE methodology is to indicate that technical leadership/management has a pre-eminent place among the SE methodologies.

The development of the research project and the interviews followed straightforward interview methodology, but still revealed insight that can be applied to the SE discipline. A commonly held SE ontology was discovered that could be expressed in eight subordinate SE activities and that was understood by all interview participants.

Through this ontology, useful demographics of the interviewed programs are shown that can be considered to be representative of similar SE domains and cultures. The implications of this ontology are that the SE discipline has a common thought process that seems to span across work domains and cultures. Acceptance of and encouragement of this common thought process can likely help the SE discipline to coalesce into greater formality.

Finally, the entire work demonstrates that it is possible to obtain meaningful and quantifiable data about systems engineering and success through empirical methods. The implication of this demonstration is that further empirical research is indeed possible, that can observe the SE discipline to the point of formulating effective underlying theory.