Accepted Manuscript

Re-Examining the Association between Quality and Safety Performance in Construction: From Heterogeneous to Homogeneous Datasets

Pauline Teo[1] and Peter E.D. Love[2]

Abstract: Recent research undertaken revealed a significant positive relationship exists between quality and safety performance. A major limitation of this research, however, was the nature of the sample; it was heterogeneous (i.e. a combination of US and international projects) and the sample was restricted to 18 projects. Building upon initial research, this paper re-examines the association between quality and safety using a homogeneous sample of 569 projects, which were derived from an Australian construction company with an annual turnover in excess of $1 billion Australian dollars (AU$). A total of 19,314 non-conformances and 17,783 injuries were used to determine the validity and reliability of previous research. A weak association between quality and safety performance was found ($p<0.01$). The p-values did not indicate any significant association between first aid and quality rates, except for the injury rate and rework frequency per million scope, which yielded an r-value of 0.307 and p-value 0.046 that is significant at 0.01 level. An association, however, between injuries and rework was identified ($r^2 = 0.70$). The discrepancy between this research's findings and that of previous work led to an examination of the issues of using ratios in correlation analysis. Thus, the statistical and arithmetic issues associated with the use of ratios are discussed and, it is recommended that estimating the relationships between quality

[1] Ph.D., Australian Research Council Research Fellow, Dept. of Civil Engineering, Curtin University, GPO Box U1987, Perth, WA 6845, Australia, Email: pauline.teo@curtin.edu.au, (Corresponding Author)
[2] Sc.D., Ph.D., John Curtin Distinguished Professor, Dept. of Civil Engineering, Curtin University, GPO Box U1987, WA 6845, Australia, Email: p.love@curtin.edu.au

and safety should be examined using regression techniques, or analysis of covariance. Linear regression was therefore performed with the injury data as the dependent variable, and rework frequency and personnel hours as the independent variables. The regression results demonstrated that there is a significant association between injuries and rework and man hours; it was revealed that both predictors accounted for 68.2% of the explained variability in injury frequency. The replication of the initial research has enabled a significant advancement in knowledge about relationship between quality and safety performance.

Keywords: Quantitative methods, injury, NCR, rework

Introduction

"We may say that a phenomenon is experimentally demonstrable when we know how to conduct an experiment, which will rarely fail to give us statistically significant results" Ronald Fischer (1971:p.14).

The replication of studies forms an integral part of science and is required for the advancement of knowledge. Essentially, the process of replication involves a study to be repeated using the same methods, different subjects, and experimenters. Replication, therefore, is important for a number of reasons, which includes (Heffner, 2016): (1) to provide assurance that results previously obtained are valid and reliable, (2) to determine their generalizability or the role of extraneous variables that have been examined, (3) to apply the results to real world situations (e.g., to practice work), and (4) to identify new research directions in consideration of previous findings from similar studies.

Wanberg *et al*. (2013) examined the relationship between several quality and safety metrics revealing that a significant association existed between them. A major limitation of this research, however, was the small sample of projects that were examined. Wanberg *et al*. (2013) acknowledged this limitation and subsequently recommended further research be undertaken using a larger sample, explicitly stating that their results should only be considered suggestive and thus stressed the need for further inquiry into this subject area. This paper builds upon and attempts to replicate the initial research undertaken by Wanberg *et al*. (2013), but instead uses a homogeneous and significantly larger dataset of 569 projects obtained from an Australian contractor. The paper commences by presenting the method and results of the Wanberg *et al*. (2013) study. The research from the analysis obtained from the projects is then compared with the findings reported in Wanberg *et al*. (2013). Through the replication of the Wanberg *et al*. (2013) research a number of shortcomings with the analysis were identified and are subsequently discussed. As a result, a more robust approach for analyzing quality and safety performance in construction is proposed.

Association between Quality and Safety Performance

An examination of the relationship between quality and safety outcomes in construction has been limited; this is specifically the case for rework and safety performance (Loushine *et al*., 2006; Hoonakker *et al*., 2010; Teo and Love, 2016). Bearing in mind the paucity of empirical research that has been undertaken to examine the relationship between these outcomes, Wanberg *et al*. (2013) used data from 32 projects (a combination of commercial, residential and civil projects) to examine the association between quality and safety performance; though only 18 projects were used in the analysis. The following safety performance data were obtained from each project: (1) recordable injury rate (i.e., number of recordable injuries per 200,000 worker-hours); and (2) first-aid injury rate (i.e., number of first-aid injuries per 200,000 worker-hours). Wanberg *et al.* (2013, p.4) defined recordable injuries as "any injury that results in death, days away from work, restricted work or transfer to another job, medical treatment beyond first aid, or loss of consciousness", and first-aid injuries as minor injuries that require one-time treatment.

In terms of the quality performance, information on the total number and cost of defect/rework, and the total hours related to undertake these tasks were obtained from each project. Wanberg *et al*. (2013) referred to rework as involving activities associated with demolition, schedule pressure, and unstable work processes. The following quality rates as a proportion of project scope (i.e., contract value) and worker hours were developed: (1) number of defects per US$1 million project scope; (2) number of defects per 200,000 worker hours; (3) cost of rework per US$1 million project scope; (4) cost of rework per 200,000 worker hours; (5) number of worker hours spent on rework per US$1 million project scope; and (6) number of worker hours related to rework per 200,000 worker hours.

Correlation and regression analyses were carried out on 12 different combinations for the two safety and six quality rates. The Pearson-r was calculated to determine the relationship between the variables and the coefficient of determination (r^2), which provides the strength of the relationship, and the associated p-values. The following four of the 12 hypotheses propagated by Wanberg *et al*. (2013) were demonstrated to be statistically significant with p-values <0.05:

1. Recordable injury rate and number of worker-hours related to rework per US$1 million project scope (n= 9, r^2 is 0.937, and p-value is 0.032);
2. Recordable injury rate per 200,000 worker-hours and the number of worker-hours related to rework per 200,000 worker-hours (n=9, r^2 is 0.977, and p-value is 0.011);
3. First-aid rate per 200,000 worker-hours and number of defects per US$1 million project scope (n=15, r^2 is 0.548, and p-value is 0.009); and

4. First-aid rate per 200,000 worker-hours and number of defects per 200,000 worker-hours (n=16, r^2 is 0.722, and p-value is 0.0011).

Wanberg *et al.* (2013, p.9) concluded that the recordable injury and first-aid rates were positively correlated to rework and to number of defects, respectively; thus a project with a poor quality performance has a higher likelihood of injuries. Given the small sample size, with *n* ranging from 9 to 16, the reliability, validity and generalizability of the results are questionable, especially as the data was comprised of a heterogeneous mix of projects from the United States and other countries.

Research Approach

Exploratory research is undertaken to replicate and examine the preliminary research carried out by Wanberg *et al*. (2013) because their research had not been clearly defined and/or understood and relied upon a small sample and data that was heterogeneous (Shields and Rangarjan 2013). When the purpose of research is to gain familiarity with a phenomenon or acquire new insight to formulate a more precise problem or develop hypothesis, exploratory studies are a pertinent and justifiable approach to adopt (Babbie 2007; Shields and Rangarjan 2013). Thus, an exploratory approach was used to further examine the seminal work undertaken by Wanberg *et a*l. (2013) on the relationship between nonconformances (NCR) and safety incidents that arose during the construction of projects undertaken by an Australian contractor with an annual turnover in excess of $1 billion Australian dollars (AU$) per annum.

The contractor that afforded access to the data for analysis and interpretation provides engineering and contracting services to infrastructure, energy and resources, and transport sectors. Quality and safety form an integral part of the organization's mission and strategy. Testament to this dedicated focus is the number of national awards the organization has received for its safety performance and in its ability to deliver and construct facilities to the highest quality, on time and to budget.

The data made available covered the period from January 2007 until October 2015. Because of the commercial sensitivity of the data given, a detailed breakdown and examples of quality and safety data provided is unable to be provided. The incidents from the database with which the researchers were provided with included a wide variety of issues such as product and system NCRs that resulted in rework, injuries, investigations, environmental incidents, unsafe acts and behaviors.

Quality

In terms of quality indicators, the types and number of NCRs from the database were examined and summarized in Table 1. The contracting organization classified NCRs into three categories according to the type of correction that was required: (1) NCRs that required an action on a nonconforming product to make it conform to requirements were classified as 'rework'; (2) NCRs that precluded it from the original intended use were classified as 'scrap', and (3) NCRs in which concessions were granted by the client to be used, despite not conforming to the specified requirements, were classified as 'use-as-is'.

Of the 569 projects examined 210 (37%) projects reported that they experienced NCRs. A total of 19,314 cases of NCRs were recorded with 47% (n=9,098) being classified as 'rework', 48% (n=9,229) as 'used-as-is', 3% scrap (n=540), and 2% (n=448) were not classified. The mean number of NCRs per project was 92. In Table 1, a total cost of AU$97 million was incurred for all NCRs during the period sampled. This equates to AU$468,472 per project across all project types. The total direct cost of rework that was experienced was approximately AU$82 million and with an average of AU$419,473 per project.

The total cost of scrap was AU$6.8 million, with a mean of AU$79,300 for each project. A total cost of AU$7.6 million was determined for 'used-as-is' NCRs, with a mean of AU$51,027 for each project. 'Undefined' NCRs had a total of AU$832,946 and a mean of AU$41,647. Fifty percent of the NCRs issued were attributed to 'rework', which accounted for 84% of their total cost. The remaining 16% of NCR costs were distributed as follows: 8%, 'used-as-is'; 7%, 'scrap'; and 1% were not defined.

Table 2 provides the average cost of NCRs for each project type. 'Building' incurred the highest mean cost per NCR at AU$10,689, followed by 'Infrastructure' at AU$4,605, and 'Rail' at AU$2,751. A similar pattern again was observed for 'rework' and 'use-as-is', although in the case of 'scrap', 'Infrastructure' experienced higher costs than 'Building'. Because of the insufficient data relating to the cost of rework and the number hours required to undertake this task, these metrics were not computed. The following quality metrics comparable with Wanberg *et al*. (2013) research were computed: (1) NCR frequency per AU$ million scope, (2) NCR frequency per million hours, (3) rework frequency per AU$ million scope, and (4) rework frequency per million hours. Corresponding to Wanberg *et al*. (2013) analysis, correlation was carried out between quality and safety rates.

Safety

Of the 569 projects that were examined, 456 reported injuries, and a total of 17,783 injuries were recorded. Other safety incidents were categorized as near misses, rail safety, and unsafe acts and conditions (Table 3). Injuries were further categorized into four main types: (1) lost-time injury (LTI), (2) first-aid injury (FAI), (3) alternate work injury (AWI), and (4) medical treatment injury

(MTI). Table 4 provides the definitions of the four types of injuries used to compute the injury and first-aid rates used in this study.

For each type of injury, the frequency rates were expressed as a ratio per million personnel hours at a project level. Table 5 summarizes the descriptive statistics for LTIs, FAIs, AWIs, MTIs and total recordable injuries (TRIs) frequency rates. Table 6 provides a breakdown of the mean value of the safety injury frequency for building, infrastructure, and rail projects. To replicate the Wanberg *et al.* (2013) approach, the number of LTIs and AWIs were extracted to compute the injury rates per million personnel hours for each project. First-aid rates were computed on the basis of the number of FAIs and MTIs per million personnel hours. In addition, TRIs were extracted and added to the comparative analysis.

Comparative Analysis

In line with Wanberg *et al*. (2013), the analysis of quality and safety rates used were computed for the sample of 456 Australian projects, which are presented in Table 7. The results indicated that the Pearson-r values (0.007-0.317) and Coefficient of determination r^2 (0-0.100) were low indicating a weak association between quality and safety rates, which were contrary to the results presented in Wanberg *et al*. (2013). The p-values did not indicate any significant association between first–aid and quality rates, except for the injury rate and rework frequency per million scope, which yielded an r-value of 0.307 and p-value 0.046 that were significant at 0.01 level. The discrepancy between this research's findings with that of Wanberg *et al*. (2013) led to a further examination of the issues of using ratios in correlation analysis.

Ratios in Correlation Analysis

Apart from the limited nature of the Wanberg *et al*. (2013) sample size, another issue pertains to the determination of the correlation between ratio variables with a common divisor (e.g. *x/z* and *y/z*); in the case of the Wanberg *et al*. (2013) research, the common divisor was personnel hours (e.g., Pearson, 1897; Dunlap *et al.*, 1997; Kim, 1999). Pearson (1897) identified that it was a fallacy to examine correlation coefficients of ratios with a common divisor – this simply results in a spurious correlation. Aldrich (1995, p. 367) defines a spurious correlation to be one which is produced by an arithmetic process and not by the organic relationship between the quantities being examined. For example, even though *x* and *y* are uncorrelated random variables, the two ratios against a common divisor random variable *z* can have a misleadingly large correlation coefficient value. This means that despite being uncorrelated, the *x* and *y* variables will become correlated when divided by a common divisor *z*.

To illustrate this phenomena, Jackson and Somers (1991) took a dataset comprised of a series of randomly simulated uncorrelated *x* and *y* variables, and the scatterplot between the ratio variables *y/x* and *x* shows a distinct negative correlation, whereas a positive correlation was demonstrated instead between *y/z* and *x/z*. Fundamentally, the strong correlation between the ratio variables may well mean that there is no relationship between the variables. It is suggested that this problem was present within the Wanberg *et al.* (2013) research.

For instance, Wanberg *et al*. (2013) concluded significant relationships with the following ratios on the basis of common denominators (worker hours): (1) the number of recordable injuries per 200,000 worker hours and the number of worker hours related to rework per 200,000 worker hours, with r^2-value of 0.977 and p-value <0.01; and (2) the number of first-aid injuries per 200,000 worker hours and the number of defects per 200,000 worker hours, with r^2 value of 0.722 and *p*-value <0.01. The large r and r^2 values may indicate spurious correlations and, therefore, should be interpreted with caution.

In practice, the use of ratios to remove the effect of size or scaling in analyses and benchmarking are ubiquitous; for example, loss time injury frequency rate (LTIFR), medical treatment injury frequency rate (MTIFR), first-aid Injury Frequency Rate (FAIFR) and total recordable injury frequency rate (TRIFR), are used as safety indicators in many industries. Safework Australia (2016), for example, also has published national standards for injury occurrence using ratios of number of LTIs per million personnel hours, which are used for benchmarking purposes. In the Wanberg *et al.* (2013) research, the recordable injuries, FAIs, and rework are expressed as a ratio per million personnel hours and per US$ million project scope. The purpose is to remove the effect of project size on the variables of interest (in this case, the number of injuries and defects); this is a common technique known as *data standardization* (Curran-Everett, 2013), which is, to adjust the numerator for variations in the denominator variable. However, the use of and the statistical properties of ratios, have been widely debated (e.g., Tanner, 1949; Atchley *et al.*, 1976; Packard and Boardman, 1988; Kronmal, 1993; Sollberger and Ehlert, 2016).

Tanner (1949) reported that the use of ratio standards is theoretically fallacious and misleading, and their use can distort the relationships of the variables of interest. To illustrate this, Figures 1a and 1b demonstrate the mathematical issues of using ratios as a standard for benchmarking purposes, such as using LTIFR and FAIFR, represented by the slope of the solid line. Given a national safety benchmark represented by the slope of the solid line, Figure 1 provides two scenarios in which the actual data intersects with the ratio standard at x personnel hours: (1) Figure

1a illustrates a lower injury frequency rate for the actual data than the benchmarking standard (as indicated by the lower gradient of the fitted data than the ratio standard), and vice versa; (2) Figure 1b identifies a higher injury frequency rate than that of the ratio standard. In Figure 1a, projects with greater than x personnel hours (for example, x_1 personnel hours) perform better than the national standard because the number of injuries that occurred (y_1) is less than the national average number of injuries forecasted by the benchmarking standard (y_2). With the same actual frequency rate, projects with less than x personnel hours (for example, x_2 personnel hours) have a poorer safety performance than the national standard because the actual number of injuries is more than the national average ($y_3 > y_4$). In this case, the use of ratio standards can be advantageous for projects with greater than x personnel hours and penalize projects with lesser than x personnel hours. In Figure 1b, the reverse occurs, in which the slope of the fitted actual data is steeper than the ratio standard. Projects with greater than x personnel hours perform worse off than the national standard, whereas projects with less than x personnel hours are better than the national standard. The interpretation of ratios standards only will be meaningful provided that the relationship between the numerator and denominator is linear and passes through the origin.

Kronmal (1993) also showed that the use of ratios in multiple regression analyses can lead to incorrect or misleading inferences, and warns against their use as either dependent or independent variables. The use of ratios to control for the differences in the denominator may misrepresent the relationship between the numerator and denominator (Curran-Everett, 2013). Furthermore, when the actual injury frequency function does not follow a linear relationship with personnel hours, as denoted in Figure 2; then the rate at which injuries occurs varies tangentially at different personnel hours (i.e., slopes represented by tangent a_1 and b_1). Specifically, at a personnel hour, the rate of injuries is represented by the slope of the tangent a_1, and at b personnel hour, the rate is represented by the slope of tangent b_1. However, in practice, injury rates are calculated using the number of injuries per million personnel hours, which are represented by the slopes of the lines a_2 and b_2. The lower rate of injury b_2 may be interpreted as improvements to the safety processes and systems when, in fact, it was because the increased personnel hours in relation to the injury frequency function. In addition, as the ratio cannot provide an insight to the absolute values of the variables, the analysis of ratios can lead to the loss of valuable information of the numerator and denominator variables (Sollberger and Ehlert, 2016). For example, a project that has experienced high levels injuries and personnel hours can have a similar rate of injury as one with low levels. The ratio also does not demonstrate the interactions or relationships between the variables of interest and how it interacts with a third variable, such as a NCR.

Given the statistical and arithmetic issues associated with the use of ratios, it has been recommended that estimating the relationships, using regression techniques, or analysis of covariance is preferred, rather than analyzing with ratios (Tanner, 1949; Winter, 1992; Curran-Everett, 2013; Sollberger and Ehlert, 2016). In terms of standards, Tanner (1949) suggested that regression standards (instead of ratio standards) that describe the relationship between the variables should be employed instead.

Quality and Safety Incidents

Correlation analysis is carried out between the frequency of quality and safety incidents that occurred within 456 projects. The Pearson-r, r^2 and associated p-values between frequencies of injuries, with frequencies of NCRs, 'rework', 'scrap' and 'use-as-is', were computed and summarized in Table 8. The Pearson-r measures the strength of the linear association between two variables. The closer the r-value is to ±1, the stronger the relationship between the variables. The

r^2 is a statistical measure of how well the data fits a linear relationship. The r^2 measures the proportion of variance that is explained by the variation in the other variable in the linear relationship. The *p*-value determines the significance of the relationship, with *p*-values <0.01 as highly significant.

In contrast to the results on the basis of injury and quality incidents, the correlation results from Table 8 demonstrate a significant association between frequencies of injuries and quality incidents. The Pearson-*r* values ranged between 0.653 and 0.896 and the r^2 ranged between 0.426 to 0.803, which demonstrated a significant association. In particular, the association between injuries and rework was significantly strong ($r^2 = 0.701$ or 70%, $p = .000$).

Table 9 provides the correlation results of other types of safety incidents (sum of rail safety, unsafe acts and conditions) and near misses with quality frequencies. The results demonstrated that other safety incidents also present a strong correlation with quality frequencies (*r* ranged between 0.647 and 0.862; r^2 ranged between 0.419 and 0.743, and was significant at 0.01 level), as compared with near misses.

Regression

Rather than regression relying on the number of injuries and rework per million personnel hours as dependent and independent variables, regression was performed with injury and rework frequency and personnel hours as the variables. Linear regression was performed with the injury data as the dependent variable, and rework frequency and personnel hours as the independent variables. Table 10 summarizes the values of the regression coefficient of each independent variable and corresponding *p*-values. The output showed that the association between injury and rework frequency, and between injury frequency and personnel hours were statistically significant (p = .000). Both dependent variables are suitable predictors of injury frequency. The regression equation is expressed as

Predicted injury frequency=2.609 + 0.202 × Rework +5.557×10^{-5}Personnel hour [Eq.1].

The regression results of the above model, demonstrated that there was a significant association between injury, and rework and personnel hours [$F(2,186) = 202.500$; $p = .000$]. Both predictors accounted for 68.2% of the explained variability in injury frequency ($R^2 = 0.682$, and $R = 0.8285$). The regression coefficients can be used, instead of injury rates, for the purposes of benchmarking and as lead indicators for safety. The results provided empirical evidence that there is a positive association between the quality and safety incidents; that is, the occurrence of unplanned work that can materialize from NCRs, defects or rework, is strongly associated with safety incidents. Although this was a large sample size, further research needs to be undertaken to examine this relationship in detail.

Through the process of replicating the Wanberg *et al*. (2013) research, statistical and interpretational issues relating to the conventional method of safety assessment used in the construction industry (e.g., LTIFR and FAIFR), in which the number of injuries is expressed as a ratio to a million personnel hours, have surfaced. Moreover, these rates assume a linear relationship between injury frequency and personnel hours, which may not be a suitable assumption depending on the distribution and characteristics of the data. A limitation to linear regression is that injury frequency data are nonnegative integers; more suitable frequency statistical methods can be employed, such as a Poisson (Chua and Goh, 2005) or Negative Binomial models (Love and Teo, 2016).

Conclusion

Replication is fundamental to creating reliable outcomes in construction. Yet, the process of replication is seldom undertaken and reported within the construction and engineering literature. In this paper, the research findings reported in Wanberg's *et al*. were re-examined using quality and safety data from a sample of 569 projects. The research was unable to replicate the findings presented in Wanberg *et al*., but instead revealed a significant association between injuries and rework that was significantly strong ($r^2 = 0.701$). During the process of replication, it was observed that Wanberg et al. had relied upon the use of ratio to remove the effect of size or scaling because of the small size of their sample. As a result, this led to spurious results being reported.

The legitimacy of the statistical and arithmetic problems associated with the use of ratios is recognized, and the recommendation to estimate the relationships between quality and safety should be examined using regression techniques or analysis of covariance. Linear regression was performed with the injury data as the dependent variable, and rework frequency and personnel hours as the independent variables. The regression results demonstrated that there was a significant association between injury, and rework and personnel hours; it was revealed that both predictors accounted for 68.2% of the explained variability in injury frequency. The replication of Wanberg *et al*.'s initial research enabled a significant advancement in knowledge about the relationship between quality and safety performance. It is recommended that future research focus on understanding the causal relations that exist between rework and safety incidents. On the basis of the new findings that are reported, it is suggested that if rework can be reduced and contained, then significant improvements in safety will ensue.

Acknowledgments

The authors would like to thank the anonymous reviewers for their constructive comments, which helped improve the quality of this manuscript. In addition, the authors would like to thank the contracting organization for agreeing to participate in this study. The authors would like to acknowledge the financial support provided by the Australian Research Council (DP130103018), which enabled this research to be undertaken.

References

Aldrich, J. (1995). Correlations genuine and spurious in Pearson and Yule. *Statistical Science,* **10**(4), pp.364-376.

Atchley, W. R., Gaskins, C. T., and Anderson, D. (1976). Statistical properties of ratios. I. empirical results. *Systematic Zoology,* **25**(2), pp.137-148.

Babbie, E. (2007). *The Practice of Social Research*. 11th Edition. Belmont CA: Thompson - Wadsworth. pp.87-89.

Chua, D., and Goh, Y. M. (2005). Poisson model of construction incident occurrence. *Journal of Construction Engineering and Management*, **131**(6), pp. 715-722.

Curran-Everett, D. (2013). Explorations in statistics: the analysis of ratios and normalized data. *Advances in Physiological Education,* 37(3), pp.213-219.

Dunlap, W. P., Dietz, J., and Cortina, J. M. (1997). The spurious correlation of ratios that have common variables: A Monte Carlo examination of Pearson's formula. *The Journal of General Psychology,* **124**(2), pp.182-193.

Fischer, R. (1971). *Design for Experiments.* 9th Edition, Hafner Publishing Company, NY Available at: http://www.phil.vt.edu/dmayo/PhilStatistics/b%20Fisher%20design%20of%20experiments.pdf, Accessed 28th July 2016)

Heffner, C.L ( ) *Introduction to Research.* Chapter 1.1. Research Methods, AllPysch, Virtual Classroom Available at: http://allpsych.com/researchmethods/introduction/, Accessed 16th September 2016.

Hoonakker, P., Carayon, P., and Loushine, T. (2010). Barriers and benefits of quality management in the construction industry: An empirical study. *Total Quality Management and Business Excellence,* **21**(9), pp.953-969.

Jackson, D. A., and Somers, K. M. (1991). The spectre of 'spurious' correlations. *Oecologia,* **86**, pp.147-151.

Kim, J.-H. (1999). Spurious correlation between ratios with a common divisor. *Statistics and Probability Letters,* **44**(4), pp.383-386.

Kronmal, R. A. (1993). Spurious correlation and the fallacy of the ratio standard revisited. *J. Journal of the Royal Statistical Society. Series A (Statistics in Society,* **156**(3), pp.379–392.

Loushine, T. W., Hoonakker, P. L. T., Carayon, P., and Smith, M. J. (2006). Quality and safety management in construction. *Total Quality Management and Business Excellence,* **17**(9), pp.1171-1212.

Love, P.E.D. and Teo. P. (2016). Statistical analysis of injury and non-conformance frequencies in construction: A Negative Binomial regression model. *ASCE Journal of Construction Engineering and Management* (In review)

Packard, G. C., and Boardman, T. J. (1988). The misuse of ratios, indices, and percentages in Ecophysiological Research. *Physiological Zoology,* **61**(1), pp.1-9.

Pearson, K. (1897). Mathematical Contributions to the Theory of Evolution.--On a Form of Spurious Correlation Which may Arise when Indices are used in the Measurement of Organs. *Proceedings of the Royal Society of London,* **60**(1896 - 1897), pp.489-498.

Shields, P. and Rangarjan, N. (2013). *A Playbook for Research Methods: Integrating Conceptual Frameworks and Project Management.* Stillwater, OK, New Forums Press

Sollberger, S., and Ehlert, U. (2016). How to use and interpret hormone ratios. *Psychoneuroendocrinology,* **63**(1), pp.385-397.

Tanner, J. M. (1949). Fallacy of per-weight and per-surface area standards, and their relation to spurious correlation. *Journal of Applied Physiology,* **2**(1), pp.1–15.

Wanberg, J., Harper, C., Hallowell, M., and Rajendran, S. (2013). Relationship between construction safety and quality performance. *Journal of Construction Engineering and Management,* **139**(10), 04013003.

Winter, E.M. (1992). Scaling: partitioning out differences in size. *Pediatric Exercise Science*, **4**, pp.296-296.

Table 1. Types of NCRs

Frequency/ Value ($)	NCR Types	Projects (N)	NCR (N)	Min.	Max.	M.	Std. Deviation
Frequency	Rework	197	9098	1	1,436	47	127
	Scrap	87	540	1	79	6	13
	Use-as-is	166	9,229	1	2,896	56	239
	Undefined	42	448	1	114	11	23
	Total	210	19,314	1	4,525	92	336
Value ($)	Rework	195	81,797,250	.01	10,079,000	419,473	1,176,038
	Scrap	85	6,740,467	.01	1,939,6110	79,300	233,262
	Use-as-is	149	7,603,028	.01	1,783,402	51,027	165,993
	Undefined	20	832,946	600	296,116	41,647	71,944
	Total	207	96,973,691	0.01	12,561,056	468,472	1,337,578

Table 2. Mean cost of NCR for each project type

Project Type	Mean Cost of a NCR	Mean Cost of a NCR			
		Rework	Scrap	Use-as-is	Un-defined
Building	10,689	13,696	10,696	$1,531	5,921
Infrastructure	4,605	8,695	13,287	$875	3,748
Rail	2,751	3,356	7,076	$1,666	0.00
Total	5,021	8,992	12,482	$824	1,859

Table 3. Statistics for different types of safety incidents

Incidents Types	**Incidents (N)**	**Mean (M)**	**Std. Deviation**
Injury	17,783	39	122
Near misses	497	3	3
Rail safety	1,678	18	101
Unsafe act	229	4	12
Unsafe condition	206	2	2
Total	**20,393**	**44**	**138**

Table 4. Types of injuries

Types of Injuries	Definition
Lost-Time Injury (LTI)	An LTI is an injury sustained by an employee that will ultimately lead to the loss of productive work time in the form of worker delays or absenteeism. An injury is considered a lost time injury only when the worker is unable to perform the regular duties of the job, takes time off for recovery, or is assigned modified work duties for the recovery period.
Alternate Work Injury (AWI)	An AWI refers to an injury sustained by an employee, who became unable to perform the scope of his duties and is subsequently assigned to perform other tasks while awaiting recovery.
First Aid Injury (FAI)	Treatment normally performed by a First Aider and not resulting in a LTI, AWI, RWI or MTI.
Medical Treatment Injury (MTI)	A medical treatment injury is a work related injury or illness (physical or psychological), which has not been classified as a LTI or AWI, and which required treatment beyond first aid.
Total Recordable Injury (TRI)	Refers to all fatalities, lost time injuries, cases restricted for work, cases of substitute work due to injury, and medical treatment cases by medical professionals (e.g., doctors and nurses).

Table 5. Safety injury frequency rates

Safety Injury Frequency Rates	N	Min	Max	Mean	Std. Deviation
LTI	185	0.3	809	17	66
AWI	260	0.3	1,295	26	90
MTI	322	0.3	580	30	60
FAI	400	1.4	1,078	75	109
TRI	389	0.7	1,295	50	113

Table 6. Mean safety injury frequency rates for each project type

Injury Frequency Rates	Building	Infrastructure	Rail
LTI	5	11	43
AWI	12	24	54
MTI	15	26	102
FAI	79	67	11
TRI	88	77	50

Table 7. Comparison of quality and safety rates with Wanberg *et al.*'s (2013) research

Safety Rates	Quality rates	Pearson *r*	Coefficient of determination r^2	p-value	Wanberg's Pearson *r*	Wanberg's Coefficient of determination r^2	Wanberg's p-value
TRIFR	NCR count per AU$million scope	0.089	0.008	0.514			
	NCR count per million hours	-0.007	0.000	0.925			
	Rework count per AU$million scope	0.159	0.025	0.259			
	Rework count per million hours	-0.021	0.000	0.788			
First aid rate (MTIFR and FAIFR)	NCR count per AU$million scope	0.147	0.022	0.347	0.740	0.548	0.009
	NCR count per million hours	0.024	0.001	0.770	0.850	0.722	0.001
	Rework count per AU$million scope	0.221	0.049	0.171			
	Rework count per million hours	0.104	0.011	0.222			
Injury rate (TRIFR less MTIFR and FAIFR)	NCR count per AU$ million scope	0.092	0.008	0.556	−0.551	0.304	0.449
	NCR count per million hours	-0.054	0.003	0.519	−0.511	0.261	0.489
	Rework count per AU$million scope	0.317	0.100	**0.046			
	Rework count per million hours	0.054	0.003	0.527			

** Correlation significant at 0.01 level.

Table 8. Association between quality and injuries

Safety	Quality	df	Pearson-*r*	Coefficient of determination r^2	p-value
Injuries	NCR	198	0.896**	0.803	.000
	Rework	187	0.837**	**0.701**	.000
	Scrap	83	0.653**	0.426	.000
	Use-as-is	159	0.864**	0.746	.000

**. Correlation is significant at the 0.01 level (2-tailed)

Table 9. Association between quality and other types of safety incidents

Safety	Quality	df	Pearson-*r*	Coefficient of determination r^2	p-value
Other Safety Incidents	NCR	200	0.825**	0.681	.000
	Rework	188	0.771**	0.594	.000
	Scrap	84	0.647**	0.419	.000
	Use-as-is	160	0.862**	0.743	.000
Near Misses	NCR	97	0.304**	0.092	.000
	Rework	93	0.327**	0.107	.001
	Scrap	43	0.345*	0.119	.020
	Use-as-is	78	0.413**	0.171	.000

**. Correlation is significant at the 0.01 level (2-tailed), *. Correlation is significant at the 0.05 level (2-tailed).

Table 10. Summary of regression coefficients of the regression model

Independent variables	**Unstandardized Coefficients**		**Standardized Coefficients**	***t***	**Sig.**
	B	Std. Error	Beta		
Intercept	2.609	4.470		.584	.560
Rework frequency	.202	.052	.178	3.877	.000
Personnel hours	5.557E-5	.000	.733	15.947	.000

List of Figures

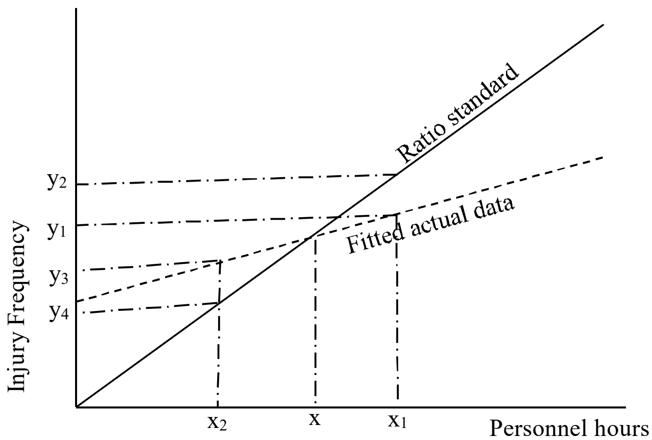

Figure 1a. Ratio standard greater than actual injury frequency rate

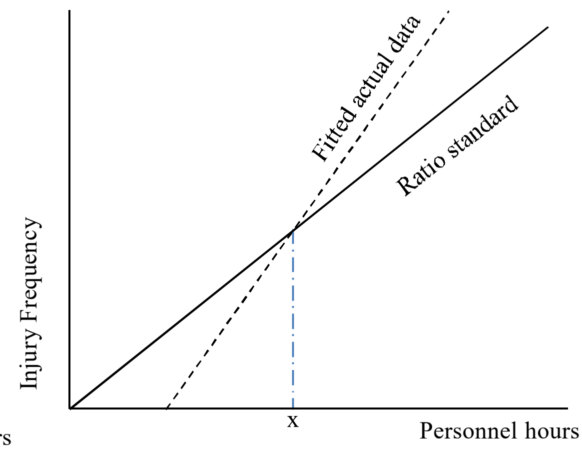

Figure 1b. Ratio standard lesser than actual injury frequency rate

Figure 1. Ratio standard and fitted actual data: (a) Ratio standard greater than actual injury frequency rate; (b) Ratio standard lesser than actual injury frequency rate

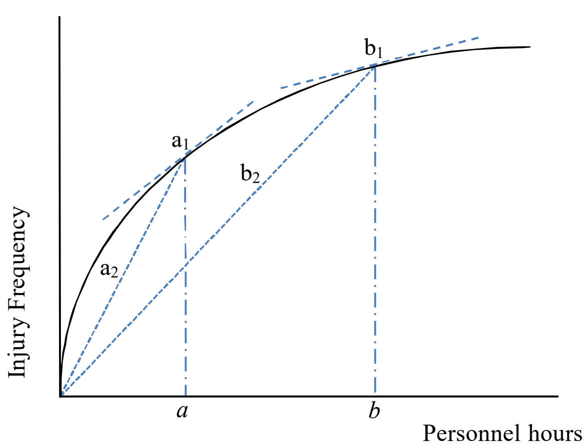

Figure 2. Injury frequency and personnel hours