STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
*Stat., Optim. Inf. Comput.*, Vol. **7**, March 2019, pp 85–96.
Published online in International Academic Press (www.IAPress.org)

# Exploratory factor analysis for software development projects in Mexico

Sergio Alcaraz-Corona *, José Luis Cantú Mata, Fernando Torres-Castillo

*Department of Mechanical and Electrical Engineering, Autonomous University of Nuevo Leon, Mexico.*

**Abstract** Time and risk management have always been major concerns in a continously growing number of processes and methodologies used to develop all kinds of useful products and services. Software like any other industry not only needs the most appropriate development methodology but also an efficient project management strategy that is capable of estimating and managing development times and any potential risks as best as possible in order to deliver high quality software products and services on time. Hence, the success or failure of software development projects depends on the efficiency by which several key factors such as cost, time, and risks are managed and how other software development factors influenced them. This paper presents an exploratory factor analysis to study the effects produced primarily on time planning caused by common factors observed in outsourced software development projects such as communication, teamwork, personnel training, among others. In our work, we also build a structural model in which we analyze the relationships between sets of dependent and independent variables derived from the responses of a survey applied to 32 Mexican organizations that subcontracted their software development projects to outsourcing providers. The results are presented and the reliability of our model was validated using the SmartPLS software.

**Keywords** Exploratory Factor Analysis, Software Development Management, Time Planning, Risk Management

**DOI:** 10.19139/soic.v7i1.512

## 1. Introduction

Time has always been a critical factor in several situations. For instance, deficient time management may produce economic losses for a company that does not meet its commitments on time. Very often, companies rely on external service providers to fulfill their business needs given a lack of internal skills [1, 3, 5, 8, 18]. In particular, it is well known that software development is one of the most common business areas handed off to outsourcing companies. The outsourcing companies are expected to produce high quality software with shorter development cycles and reduced costs [7, 14]. In order to do that, software development companies or software suppliers often must use various tools and methodologies that not only can help reduce all development work, but also meet deadlines and budgets [2, 25]. However, developing software products and services is usually a complex task that requires both knowledge and experience in managing several major aspects related to the whole development process, e.g., costs, risks, personnel, and scheduling or time management issues among others [6, 15, 26]. In fact, the success or failure of a given project heavily depends on how well the previous aspects and several others are managed.

Regarding scheduling issues, it has been shown that software projects that are late usually imply higher expenses. Therefore, time planning, which in this work entails time estimation and management, have become one of the most important aspects of a growing number of software outsourcing projects where meeting deadlines and staying under budget are as important as satisfying all customer requirements [24]. In general, factors such as, development methodologies, level of technical skills, available personnel, teamwork, communication issues,

*Correspondence to: Sergio Alcaraz-Corona (Email: sergio.alcarazcrn@uanl.edu.mx). Facultad de Ingenierła Mecánica yEléctrica, Universidad Autónoma de Nuevo León. Pedro de Alba s/n, San Nicolás de los Garza, Nuevo León, México (66451).

ISSN 2310-5070 (online) ISSN 2311-004X (print)

among others, have been identified as relevant in project time planning and coordination efforts. Hence, the motivation of this work is to investigate the level of significance that the previous factors and several others exert on project time planning activities. For that, we performed an exploratory factor analysis (EFA) [17, 28] to address the following research question:

RQ1: How feasible is to determine the factors that have more relevance on project time planning using an EFA-based model in order to minimize risks.

To answer the previous question, we conducted a survey among 32 Mexican organizations of various types that had to subcontract some or all of their software development needs to outsourcing providers. From the responses, we obtained valuable information about the clients experience on such projects that allowed us to define a structural model along with two sets of independent and dependent variables to analyze their dependencies and level of significance through a series of hypotheses tests. By identifying the most relevant factors, not only time planning and management activities can be executed more effectively, but also any potential risks can be minimized namely, financial risks. Given the ubiquitous nature of software in a continuously growing number of applications, we think this kind of studies can provide valuable information on how to handle numerous technical and managerial aspects commonly found in software development projects within organizations. The rest of this paper is organized as follows. In the next section, we present some background information and motivation for this work by discussing various aspects that are usually present in virtually any process or methodology for developing business or commercial software products and services, and their impact on any time management activities. In Section 3, we describe our research methodology including our survey whose results were then used to build our model. Section 4 presents the statistical analysis and results. The paper concludes with a summary of the key findings and their implications for further research discussed in Section 5.

## 2. Background and motivation

Ever since it first appeared, software engineering has always sought to apply a systematic approach for developing, operating, and maintaining software-based systems [2, 25]. However, as the demand for faster and more efficient software products and services increases, software development methodologies have been constantly evolving to satisfy such demands. Because of that, the scope of software engineering has not been limited to technical activities of software development such as analysis, design, coding, and testing software. In fact, it also covers administrative tasks that can only be successfully completed through an efficient project planning strategy that can ultimately help meet all customer requirements [22]. Thus, an essential part of any project planning strategy should consider both, time estimation and time management as two of the most important factors in any software development project [24]. Inadequate handling of either or both time factors can originate a series of problems for both customers and development organizations. Clearly, some of the most common problems are multiple delays, additional costs, and in general, a loss of credibility and a negative image for the outsourcing organization [4]. Therefore, a correct time planning strategy is so important in any software development project given the various forms to review progress and status information [27].

Regarding time estimation, one approach indicates that time estimates can be derived from past projects that share similar characteristics, i.e., historical data from previous efforts can be used as a basis for proposing new estimates. The latter approach falls into a category known as analogy-based models [15]. However, for new software projects the complexity of time estimation increases and other models would have to be considered to come up with such important information [7, 25, 26]. In any case, a successful project management strategy depends on how well the various software development and project management activities are organized and coordinated [22]. Among the various activities that play a major role in time planning and management activities, the development methodology used to build the software has been identified as one of the most important factors. For instance, it is well know that any software development methodology divide all the necessary work into several phases or stages usually defined as requirements definition, design, coding and several forms of software testing and quality assurance activities embedded into them [2, 25]. Clearly, the previous work phases need to be properly managed and planned out in order to minimize any potential risks [21], i.e., it should also include accurate estimates regarding timeframe durations for each phase based on other key factors such as personnel size, their technical skills, available resources, among others. Of course, customers and developers should communicate effectively at all times to ensure deadlines are satisfied as well as all customers demands [16, 20]. Otherwise, communication deficiencies may cause deliverables to slip.

Besides the development methodology and communication issues, there are other factors that also impact time estimation and time management efforts. For instance, the members of the team responsible for developing the requested software product or service should possess a set of technical and non-technical skills that would allow them to deliver a quality product on time [9]. Hence, the level of technical competencies along with the number of available members assigned to a given project represent two additional important aspects for time scheduling tasks. As existing applications continue to evolve and the number of new requirements grows, software development organizations have taken steps to improve their production processes in order to be more competitive in the globalized world.

So far, we have identified and described the main sources that somehow affect any activities related to time planning required in business or commercial software development projects. Clearly, there could be other sources or factors which essentially make any software development time planning activity a complex task that requires both ample knowledge and experience in project management, namely project time planning and risk management. As indicated before, we first conducted a survey among 32 Mexican companies that subcontracted some or all of their software development needs to external or outsourcing providers in order to obtain valuable information from which we were able to define a model intended to analyse the level of significance between sets of independent and dependent variables through an EFA-based study. Moreover, we also used the Smart-PLS software to help us find the correlations between the two sets of aforementioned variables and test our proposed hypotheses [13]. In the following sections, we explain how our method is capable of measuring which development factors are more significant or have more influence on both time planning and risk management tasks.

## 3. Method

We consulted the database of the Software Council of the state of Nuevo Leon, Mexico (CSOFTMTY) which is an alliance between universities, government, and enterprises searching for economic growth, quality, and innovation mainly in the software industry. From the data set analysed, we developed a questionnaire of 29 multiple-choice questions (indicators), each one using a five-point Likert scale (*1. Totally disagree, 2. Disagree, 3. Neither disagree nor agree, 4. Agree, 5. Totally agree*), as our main data collection tool [19]. For only indicator $V_9$ the scale had to be changed accordingly (*1. Once every two years, 2. Once a year, 3. Twice a year, 4. Three times a year, 5. Four or more times a year*). Notice, however, that the questions are formulated from the perspective of a client company evaluating different aspects of the entire development work cycle provided by some software service provider.

The target population for this survey-based study was also selected from the CSOFTMTY database where we identified a total of 73 organizations of various types that were invited to participate during the second half of 2011 although at the end, only 32 of those responded. Since we wanted to collect project-related data from experienced people, the majority of respondents were either technical managers, high-level executives or project managers. Hence, each participant responded all multiple choice questions based on their own background knowledge and experience. A breakdown of the participants according to their role is the following: 71.88% are directors, 12.5% are sub directors, 9.37% are project leaders, and 6.25% are group managers. Meanwhile, the gender distribution is 90.63% for men and 9.37% for women. Even though there are other criteria such as lines of code (LOC) or effort estimates based on the number of available personnel [25, 26], the size of the projects consulted in this study was classified as small, medium, or large based on their approximate total costs, i.e., $1–500,000 USD for small, $500,001–1,500,000 USD for medium, and over $1,500,000 USD for large-sized projects. In this case, the project size distribution was roughly 40% for small, 35% for medium, and 25% for large efforts.

As part of the models definition, the questions are then organized in constructs based on the specific aspects or topics of a software development project they are focusing on. Moreover, for convenience each question is assigned a unique identifier or code to facilitate the graphical representation of the various dependencies that may exist among indicators and constructs. Table 1 and 2 show the 29 items that form our questionnaire in conjunction with their corresponding constructs where each one represents one of the following common terms usually found in any software development project: $X_1$ = Employee training programs; $X_2$ = Outsourcing; $X_3$ = Communication; $X_4$ = Innovation; $X_5$ = Teamwork; $X_6$ = Project management; $X_7$ = Project completion; $Y_1$ = Lack of internal development skills; $Y_2$ = Development time planning; $Y_3$ = Risk analysis.

In this work, the $Y_j$ variables represent our dependent variables and our goal is to measure the approximate level of significance or influence that the $X_i$ latent variables exert on the $Y_j$ ones.

Below we present the set of hypothesis tests that were defined based on the proposed constructs, and the data that were collected and analyzed.

$H_1$: Employee training programs ($X_1$) are significant for a lack of internal development skills ($Y_1$).
$H_2$: Outsourcing ($X_2$) is significant for a lack of internal development skills ($Y_1$).
$H_3$: Communication ($X_3$) is significant for a lack of internal development skills ($Y_1$).
$H_4$: Communication ($X_3$) is significant for development time planning ($Y_2$).
$H_5$: Innovation ($X_4$) is significant for development time planning ($Y_2$).
$H_6$: Teamwork ($X_5$) is significant for development time planning ($Y_2$).
$H_7$: Project management ($X_6$) is significant for development time planning ($Y_2$).
$H_8$: Project completion ($X_7$) is significant for risk analysis ($Y_3$).
$H_9$: Lack of internal development skills ($Y_1$) is significant for risk analysis ($Y_3$).
$H_10$: Development time planning ($Y_2$) is significant for risk analysis ($Y_3$).

Figure 1 presents the complete graphical representation of our model including all the relationships between constructs and indicators.

## 4. Model fit

In order to validate the efficiency of our model, the following tests were carried out:

First, we made a normality test on the data focusing on the skewness and kurtosis. As shown in Table 3, the obtained values are within 2, and therefore, we can conclude that the proposed variables are normally distributed.

Next, a median test was conducted to determine whether the data comes from the same population and in which we consider the following hypotheses:

$H_0$: The median values of all k populations are the same

$H_a$: At least one of the populations has a distinct median value

As shown in Table 4, the null hypothesis of the medians'comparison is rejected whenever the significance value is less than 0.05 which also corresponds to a confidence level of 95%. Therefore, we can conclude that for the relationships between $X_6$ and $Y_2$ as well as $X_7$ and $Y_3$, at least one of the populations has a distinct median.

A multicollinearity analysis is performed in which Table 5 shows the values obtained for each of the latent variables associated with their dependent variables. The Variance Inflation Factor (VIF) should have values under 5 [12]. As shown, each VIF value is less than 5 and the tolerance index values are within acceptable levels, i.e., the values are neither close to 0 nor above 1. The latter shows that there is no collinearity between latent variables.

Convergent Validity (CV) evaluates whether a set of indicators measure a particular construct and not some other concept [11]. Moreover, the Average Variance Extracted (AVE) represents the average variation that a latent variable exerts over the observable variables [10]. It can be shown that values above 0.5 are acceptable [12], and

Table 1. Survey questions and Hypothesis

| Construct | Hypothesis | Code | Indicators |
|---|---|---|---|
| Employee training programs $X_1$ | $H_1 : X_1 \rightarrow Y_1$ | $V_1$ | Contractors showed a high level of performance |
| | | $V_2$ | The software development methodology satisfied all customer requirements |
| | | $V_3$ | The total project cost was reasonable based on personnel capacities |
| Outsourcing $X_2$ | $H_2$: $X_2 \rightarrow Y_1$ | $V_4$ | Client sought outsourcing providers due to a lack of internal capacities |
| Communication $X_3$ | $H_3$: $X_3 \rightarrow Y_1$ | $V_5$ | Contractors adapted to all client necessities |
| | | $V_6$ | Contractors showed a right judgment and kept the client informed about any issues |
| | | $V_7$ | When needed, contractors displayed an adequate command of the English language |
| | $H_4$: $X_3 \rightarrow Y_2$ | $V_8$ | Contractors communicated effectively |
| Innovation $X_4$ | $H_5$: $X_4 \rightarrow Y_2$ | $V_9$ | How often new software products or services are released |
| Teamwork $X_5$ | $H_6$: $X_5 \rightarrow Y_2$ | $V_{10}$ | The determination and initiative of client personnel was important to meet objectives |
| | | $V_{11}$ | Clients long term vision to foresee any future requirements |
| | | $V_{12}$ | The set of tangible resources (financial resources, physical assets) were sufficient |
| | | $V_{13}$ | All project objectives were satisfied |
| | | $V_{14}$ | Adequate individual and collective interpersonal abilities |
| Project management $X_6$ | $H_7$: $X_6 \rightarrow Y_2$ | $V_{15}$ | The project management strategy help achieve the financial commitments for the project |
| | | $V_{16}$ | All project phases were effectively planned |
| | | $V_{17}$ | Control over the final product |
| Project completion $X_7$ | $H_8$: $X_7 \rightarrow Y_3$ | $V_{18}$ | Contractors appeared to be qualified and up to date in their technical skills |
| | | $V_{19}$ | Contractors showed honesty, integrity and ethical values |
| | | $V_{20}$ | The final product was completed within the established dates |

Source: Authors

as Table 6 shows, all AVE values are above 0.5 and their average value is 0.748 which altogether satisfy the CV criterion [10, 12].

Composite Reliability (CR) refers to the internal consistency of a latent variable without assuming that the indicators are reliable, but instead it assigns them priorities. Any values between 0.6 and 0.7 are considered appropriate as inferior limit [12]. As shown in Table 6, all CR values are above 0.8.

Table 2. Survey questions and Hypothesis, cont.

| Dependent variables | Hypothesis | Code | Indicators |
|---|---|---|---|
| Lack of internal development skills $Y_1$ | $H_9$: $Y_1 \rightarrow Y_3$ | $V_{21}$ | The provider added more contractors to the project due to a shortage of skills |
| | | $V_{22}$ | Contractors showed at any time that they needed to acquire more work related experience |
| | | $V_{23}$ | The provider added more contractors to the project to deliver the product sooner |
| Development time planning $Y_2$ | $H_{10}$: $Y_2 \rightarrow Y_3$ | $V_{24}$ | The provider managed time effectively in order to meet all project objectives |
| | | $V_{25}$ | The provider always kept enough personnel to attend each development phase |
| | | $V_{26}$ | The time allocated for each development phase was adequate |
| Risk analysis $Y_3$ | - | $V_{27}$ | During the entire project development, was there any risk of increasing the total cost |
| | | $V_{28}$ | During the entire project development, was there any risk of compromising the products quality |
| | | $V_{29}$ | During the entire project development, was there any risk of cancelling the whole project |

Source: Authors

Table 3. Normality test

| Construct | Skewness | Kurtosis |
|---|---|---|
| $X_1$ | -0.793 | -0.178 |
| $X_2$ | 0.384 | -0.315 |
| $X_3$ | -0.674 | -0.741 |
| $X_4$ | 0.658 | -0.399 |
| $X_5$ | 0.197 | -1.219 |
| $X_6$ | -0.283 | -1.022 |
| $X_7$ | -1.432 | 1.140 |
| $Y_1$ | 0.336 | -0.910 |
| $Y_2$ | -1.062 | 0.259 |
| $Y_3$ | 0.018 | -1.507 |

Source: Analysis of results using SPSS.

Discriminant Validity proves that a construct measures a concept distinct from other constructs. This type of validity was performed in two parts, the first part consists of the Fornell-Larcker method which compares the squared value of the highest correlation (0.454) against the AVE for each variable. It can be shown that the AVE is superior, and therefore, the second part can be executed. Basically, in the second part we obtained the average cross loading values for each latent variable which are then compared against the composite reliability values [11].

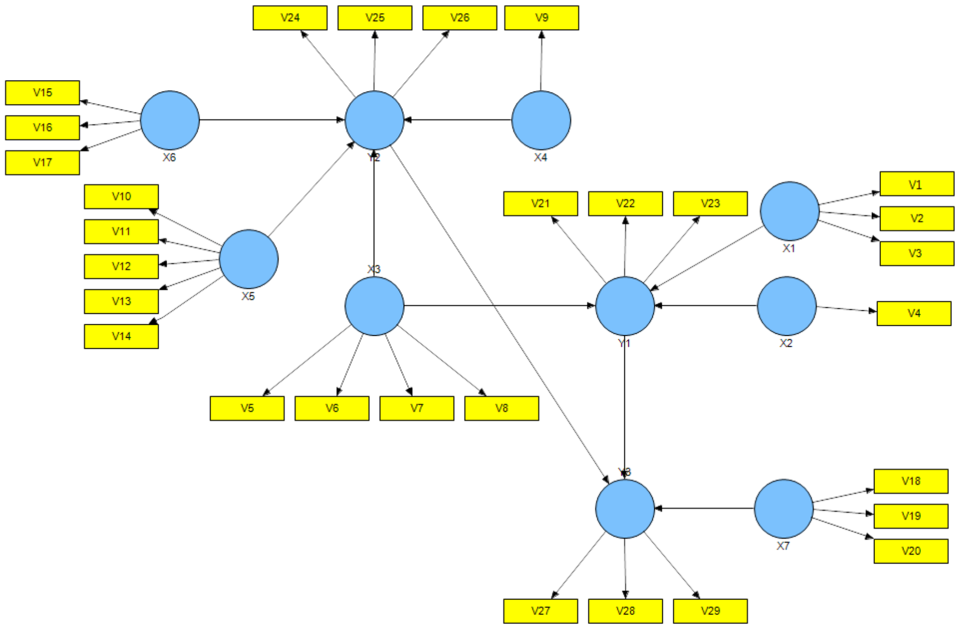


Figure 1. Graphical view of the model

Table 4. Kruskal-Wallis test

| | $Y_1$ | | | $Y_2$ | | | | $Y_3$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| K-W test | $X_1$ | $X_2$ | $X_3$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ | $Y_1$ | $Y_2$ |
| $\chi^2$ | 6.634 | 3.049 | 5.616 | 6.348 | 7.368 | 7.824 | 14.941 | 9.317 | 7.734 | 2.814 |
| Sig. | 0.157 | 0.550 | 0.230 | 0.175 | 0.118 | 0.098 | 0.005 | 0.025 | 0.052 | 0.421 |

Source: Analysis of results using SPSS.

Table 5. Multicolinearity test

| Dependent variable | Construct | Tolerance | VIF |
|---|---|---|---|
| $Y_1$ | $X_1$ | 0.718 | 1.392 |
| | $X_2$ | 0.853 | 1.172 |
| | $X_3$ | 0.632 | 1.581 |
| $Y_2$ | $X_3$ | 0.525 | 1.905 |
| | $X_4$ | 0.977 | 1.024 |
| | $X_5$ | 0.588 | 1.700 |
| | $X_6$ | 0.545 | 1.835 |
| $Y_3$ | $X_7$ | 0.968 | 1.033 |
| | $Y_1$ | 0.954 | 1.048 |
| | $Y_2$ | 0.934 | 1.070 |

Source: Analysis of results using SPSS.

Table 6. Quality criteria

| Construct | AVE | Composite reliability | Average cross loadings | $R^2$ | Cronbach's Alpha |
|---|---|---|---|---|---|
| $X_1$ | 0.720 | 0.885 | 0.416 | | 0.812 |
| $X_2$ | 1.000 | 1.000 | 0.169 | | 1.000 |
| $X_3$ | 0.740 | 0.919 | 0.359 | | 0.889 |
| $X_4$ | 1.000 | 1.000 | 0.049 | | 1.000 |
| $X_5$ | 0.666 | 0.888 | 0.448 | | 0.833 |
| $X_6$ | 0.840 | 0.940 | 0.405 | | 0.905 |
| $X_7$ | 0.594 | 0.815 | 0.290 | | 0.691 |
| $Y_1$ | 0.641 | 0.841 | 0.142 | 0.713 | 0.723 |
| $Y_2$ | 0.610 | 0.824 | 0.233 | 0.766 | 0.679 |
| $Y_3$ | 0.634 | 0.836 | 0.348 | 0.446 | 0.696 |

Source: Analysis of results using SMART-PLS.

Notice that for each latent variable, the composite reliability values are higher than the average cross loading values as shown in Table 6.

The $R^2$ results for dependent variables $Y_1$ (0.713), $Y_2$ (0.766), and $Y_3$ (0.446) are shown in the fifth column of Table 6. Any values above 0.750 are considered substantial whereas values above 0.500 are considered moderate, and values above 0.250 are weak [12].

Cronbach's alpha represents the internal correlation or reliability of a set of indicators that measure either a non-observable or directly measured variable. In this case, each of the proposed variables have been measured with their corresponding indicators, and the results are acceptable based on the fact that they all satisfy the lower limit of 0.6 [12]. On the other hand, $X_2$ and $X_4$ are measured with just one indicator and that explains the high value obtained (1.000) as shown in Table 6.

Analysis of $Q^2$ value. Table 7 shows the predictive relevance of the model. For instance, any $Q^2$ values above 0.35 have a high predictive relevance whereas any values between 0.15 and 0.35 have a medium predictive relevance. Lastly, any values between 0.02 and 0.15 have a low predictive relevance [13].

Table 7. Results for $Q^2$

| Construct | $Y_1$ | $Y_2$ | $Y_3$ |
|---|---|---|---|
| $X_1$ | 0.498 | | |
| $X_2$ | 0.222 | | |
| $X_3$ | 0.470 | 0.409 | |
| $X_4$ | | 0.029 | |
| $X_5$ | | 0.177 | |
| $X_6$ | | 0.188 | |
| $X_7$ | | | 0.221 |
| $Y_1$ | | | -0.012 |
| $Y_2$ | | | 0.181 |

Source: Analysis of results using SMART-PLS.

Table 8. $\chi^2$ Test

| Constructs | Causal relationship | $\chi^2$ | Sig. | Hypothesis |
|---|---|---|---|---|
| $X_1$-$Y_1$ | 0.837 | 161.62 | 0.004 | $H_1$: Accept |
| $X_2$-$Y_1$ | 0.560 | 75.56 | 0.000 | $H_2$: Accept |
| $X_3$-$Y_1$ | -0.841 | 208.38 | 0.000 | $H_3$: Accept |
| $X_3$-$Y_2$ | -0.931 | 178.94 | 0.000 | $H_4$: Accept |
| $X_4$-$Y_2$ | 0.236 | 41.06 | 0.258 | $H_5$: Reject |
| $X_5$-$Y_2$ | 0.630 | 191.00 | 0.000 | $H_6$: Accept |
| $X_6$-$Y_2$ | 0.682 | 172.11 | 0.000 | $H_7$: Accept |
| $X_7$-$Y_3$ | 0.548 | 66.46 | 0.005 | $H_8$: Accept |
| $Y_1$-$Y_3$ | -0.078 | 145.52 | 0.005 | $H_9$: Accept |
| $Y_2$-$Y_3$ | 0.494 | 121.64 | 0.000 | $H_{10}$: Accept |

Source: Analysis of results using SMART-PLS and SPSS.

## 5. Independence Test

In order to prove the hypotheses ($X_i \rightarrow Y_j$) we used a $\chi^2$ test to probe the relationship between two constructs by the following hypothesis:

$H_0$: The study construct is not related to the dependent variable.
$H_a$: The study construct is related to the dependent variable.

The test shown in Table 8 where all practical cases are higher than their theoretical counterparts except for the relationship between Innovation ($X_4$) and Development time planning ($Y_2$). Therefore, each of the proposed hypotheses is accepted except $H_5$.

## 6. Analysis of results

Employee training programs ($X_1$). Each employee assigned to work in a project is expected to apply various skills and perform several activities that would in turn help the project satisfy its requirements. This variable also represents all qualified personnel, development programs and any costs related to it. Moreover, it is important to always maintain a good working relationship with the service provider, share information and communicate effectively. Once the project is over, certain employees may be assigned to train users on how to operate the recently created software. In general, this variable is relevant whenever there is a lack of internal skills.

Outsourcing ($X_2$). This variable is relevant for companies whose main business has nothing to do with software development. In this case, the idea is to hand off all the company's software development needs to an external service provider in order to save time and money that otherwise the company would have had to invest to hire and train the right personnel to do it.

Communication ($X_3$). This variable is very important throughout the entire development work cycle but especially during the initial and final stages of a project. All personnel from both the client and outsourcing companies should keep in constant communication during all development phases [16, 20]. It is possible that communication deficiencies may cause a negative impact on other important variables, e.g., the negative values obtained for 'lack of internal development skills'($Y_1$) and 'development time planning'($Y_2$).

Innovation ($X_4$). This construct is not relevant because so far, one of the main business objectives of software development companies is to satisfy all software-related customer demands on time and within cost estimates, i.e., the former are not necessarily innovating at least at the time when this work was performed. Actually, customer or client organizations are the ones who are innovating with the aid of the software products and/or services they are continuously requesting.

Teamwork ($X_5$). Represents the group of people responsible for performing all the necessary activities to develop the end product. Clearly, there has to be an effective teamwork collaboration to minimize any potential risks, in particular anything that could impact the proposed schedule and its deliverables. Hence, this variable is also relevant.

Project management ($X_6$). Evidently, this construct should not be overlooked as it is fundamental for other major variables, in this case, for 'development time planning '($Y_2$). Moreover, we used a factor that represents the 'vision of the software development process 'as an indicator ($V_{11}$). Such indicator tries to reflect the idea that projects, in general, should be built with a broader long term vision and not just to satisfy any immediate needs

Project completion ($X_7$). Unexpected events of various kinds may cause delays at any stage of the development process among several other risks [24]. Therefore, it is very important that any software provider has an effective project management strategy capable of handling the aforementioned events in order to reduce or eliminate risks that can affect the estimated completion of a given project.

Lack of internal development skills ($Y_1$). Even though the software provider is responsible for the whole development process, there may be instances where personnel from the client company may be asked to provide feedback or guidance with certain project related details. Thus, the less client employees know about the way business software is usually developed, the higher the risks that they might provide incorrect or misleading information.

Development time planning ($Y_2$). The development time becomes larger, the higher the chances of witnessing more risks that may increase costs, cause delays, among others.

As with any empirical study, our model is not freed from a number of threats of validity, both internally and externally as well as the construct itself. First, internal validity reflects the extent to which the proposed model supports its outcome. The criteria analyzed in section 4 provide support for a general assessment of the internal and construct validity of our method. On the other hand, external validity is concerned with the generalization of the obtained results to other environments or practical cases. In this case, our study only involved 32 respondents from the same geographical region which represents a sample that somehow limits any attempts to generalize the results to any software development projects. However, our model is capable of producing effective results for samples of size 30 or more to comply with normality tests.

Moreover, our model is designed to work with quantitative data only, and requires qualitative data to be transformed into some form of numerical representation prior to its application. In general, several factors such as cross-cultural differences, work practices, and work values in the case of offshore software development can also have a strong impact on the results of studies such as this one. Likewise, technical characteristics of each project, e.g., whether it is a small, medium or large-sized project could play a major role in the outcome. Clearly, business software development is a complex task that involves several factors that need to be properly addressed and managed, in particular, the ones that can become the source of major risks and jeopardize the development of the end product.

Given the large diversity of practical scenarios, we realize the list of dependent and independent variables presented in this case study are not comprehensive because each respondent may come up with several more factors based on their own experiences. Nevertheless, we are confident the set of indicators and constructs that were defined provide a fair description of most current outsourced software development projects, and therefore, we consider this kind of experimental studies can be valuable as they can show other ways to approach and analyze certain aspects related to the development of software products or services.

## 7. Conclusions

The proposed model is designed to work with quantitative data only, and requires qualitative data to be transformed into some form of numerical representation prior to its application. In general, several factors such as cross-cultural differences, work practices, and work values in the case of offshore software development can also have a strong impact on the results of studies such as this one. Likewise, technical characteristics of each project, e.g., whether it is a small, medium or large-sized project could play a major role in the outcome.

In this work, we have presented a novel way to define what are the most relevant factors behind time planning and management activities required in any major business or commercial software development project. Our methodology is mainly based on executing an exploratory factor analysis on a model that was built from the responses of a survey intended to extract several factors in order to test their relevance in time scheduling tasks. The decision to use an EFA-based study proved to be an effective and reliable way to identify such relevant factors, and therefore, we were able to answer the research question stated in the introductory section. Finally, we consider this kind of case studies can be very valuable as the software industry will continue to evolve and expand as well as their development methodologies.

## REFERENCES


1. K. G. Abraham, and S. K. Taylor, *Firms' use of outside contractors: theory and practice*, Journal of Labor Economics, vol. 14, no. 3, pp. 394–424, 1993.
2. P. Abrahamsson, O. Salo, J. Ronkainen, and J. Warsta, *Agile software development methods: review and analysis*, VTT Publications, vol. 478, 2002.
3. S. Ainin, S. Bahri, T. Faziharudean, and N. Mohd, *Impact of business process outsourcing practices on financial performance*, Asian Journal of Information Technology, vol. 11, no. 2, pp. 56–64, 2012.
4. M. A. Akbar, J. Sang, A. A. Khan, M. Shafiq, S. Hussain, H. Hu, and H. Xiang, *Improving the quality of software development process by introducing a new methodology–AZ-model*, IEEE Access, vol. 6, pp. 4811–4823, 2018.
5. A. Arora, V. S. Arunachalam, J. Asundi, and R. Fernandes, *The Indian software services industry*, Research Policy, vol. 30, no. 8, pp. 1267–1287, 2001.
6. M. Bano, D. Zowghi, N. Ikram, and M. Niazi, *What makes service oriented engineering challenging? A qualitative study*, IET Software, vol. 8, no. 4, pp. 154–160, 2013.
7. B. W. Boehm, C. Abts, A. Winsor Brown, S. Chulani, B. K. Clark, E. Horowitz, R. Madachy, D. J. Reifer, and B. Steece *Software cost estimation with COCOMO II*, Prentice-Hall, Englewood Cliffs, New Jersey, 2000.
8. E. Carmel, *The new software exporting nations: Impacts on national well being resulting from their software exporting industries*, Electronic Journal on Information Systems in Developing Countries, vol. 13, no. 3, pp. 1–6, 2003.
9. R. Elizalde, and S. Bayona, *Interpersonal relationships, leadership and other soft skills in software development projects: a systematic review*, World Conference on Information Systems and Technologies, Springer, Cham., pp. 3–15, 2018.
10. A. M. Farrell, *Insufficient discriminant validity: A comment on Bove, Pervan, Beatty, and Shiu*, Journal of Business Research, vol. 63, no. 3, pp. 324–327, 2010.
11. C. Fornell, and D. F. Larcker, *Structural equation models with unobservable variables and measurement error: Algebra and statistics*, Journal of Marketing Research, vol. 18, no. 3, pp. 382–388, 1981.
12. J. F. Hair, C. M. Ringle, and M. Sarstedt, *PLS-SEM: Indeed a silver bullet*, Journal of Marketing Theory and Practice, vol. 19, no. 2, pp. 139–152, 2011.
13. J. F. Hair, G. T. M. Hult, C. M. Ringle, and M. Sarstedt, *A primer on partial least squares structural equation modeling (PLS-SEM)*, SAGE Publications, 2014.
14. D. R. Hansen, and M. M. Mowen, *Cost Management: Accounting and Control*, Thomson South-Western, Fifth edition, 2006.
15. M. Jorgensen, *What we do and don't know about software development effort estimation*, IEEE Software, vol. 1, no. 2, pp. 37–40, 2014.
16. Y. Y. Jusoh, R. N. H. Nor, B. A. Mahmood, M. T. Wafeeq, M. A. Ali, and M. N. B. Jusoh, *Communication management in global software development projects*, IEEE Fourth International Conference on Information Retrieval and Knowledge Management (CAMP), pp. 1–7, 2018.
17. H. F. Kaiser, *An index of factorial simplicity*, Psychometrika, vol. 39, no. 1, pp. 31–36, 1974.
18. D. R. Kim, M. J. Cheon, C. D. Beugré, and T. K. Coverdale, *Information systems outsourcing: Exploration on the impact of outsourcing service providers'service quality*, Issues in Information Systems (IIS), vol. 4, pp. 528–534, 2003.
19. T. C. Lethbridge, S. Elliot Sim, and J. Singer, *Studying software engineers: data collection techniques for software field studies*, Empirical Software Engineering, vol. 10, no. 3, pp. 311–341, 2005.
20. S. A. Licorish, and S. G. MacDonell, *Communication and personality profiles of global software developers*, Information and Software Technology, vol. 64, pp. 113–131, 2015.
21. K. Lyytinen, L. Mathiassen, and J. Ropponen, *Attention shaping and software risk – a categorical analysis of four classical risk management approaches*, Information Systems Research, vol. 9, no. 3, pp. 233–255, 1998.
22. C. Manteli, I. Van de Weerd, and S. Brinkkemper, *Bridging the gap between software product management and software project management*, Proceedings of the 11th International Conference on Product Focused Software, ACM, pp. 32–34, 2010.
23. T. Mezher, W. Nasrallah, and A. Alemedinne, *Management of technological innovation in Lebanese industry*, PICMET, vol. 3, pp. 1064–1073, 2006.
24. P. Savolainen, J. J. Ahonen, and I. Richardson, *When did your project start?–The software suppler's perspective*, The Journal of Systems and Software, vol. 104, pp. 32–40, 2015.
25. I. Sommerville, *Software engineering*, Addison Wesley Longman Publishing Co. Inc., Ninth edition, 2010.
26. C. J. Torrecilla-Salinas, J. Sede?o, M. J. Escalona, and M. Mejías, *Estimating, planning and managing agile web development projects under a value-based perspective*, Information and Software Technology, vol. 61, no. 2, pp. 124–144, 2015.
27. Y. K. Wong, *Modern software review: Techniques and technologies*, IRM Press, First edition, 2006.

28. A. G. Yong, and S. Pearce, *A beginner's guide to factor analysis: Focusing on exploratory factor analysis*, Tutorials in Quantitative Methods for Psychology, vol. 9, no. 2, pp. 79–94, 2013.