Development of a Scale for Measuring Software Diffusion

Mr Grant J. Castner
BCom (Hons) UQ
Email: castner@commerce.uq.edu.au

Dr Colin B. Ferguson
BBus Vic.I.C., DipEd S.C.V. (Hawthorn), MEC N.E., GdipComp, PhD Deakin
MACS, ACA
Email: ferguson@commerce.uq.edu.au

Abstract
Existing research on diffusion of information technology in firms focuses on the factors that effect the diffusion with little study of diffusion itself. Most of the diffusion research also focuses on hardware diffusion and fails to consider the inter-related area of software diffusion. Consequently, robust measures of the software diffusion construct have not been developed. This paper addresses these omissions by developing a scale for the measurement of software diffusion. Software diffusion is the degree of dispersion of software throughout the firm.

The questionnaire incorporated features designed to increase the validity and reliability of the responses while also maximising the response rate.

The resultant 3-item scale satisfies the critical validity and reliability benchmarks normally associated with these types of scales. The scale incorporates both the adoption and implementation aspects of diffusion.

1. Introduction and Motivation

1.1 Introduction

Diffusion is ‘the process by which (1) an innovation (2) is communicated through certain channels (3) over time (4) among members of a social system’ ([24], p.10). In an information technology context, it is the extent to which information technology has been dispersed throughout an organisation [15]. Software diffusion is the degree of dispersion of the software throughout the organisation. It is not merely the availability of the software but the percentage of staff who actually use the software.

This study documents the development of a scale for the measurement of the software diffusion construct. Spreadsheet software is the example used in this study, but the scale can be adapted to any type of software. Development of the scale incorporates qualitative and quantitative testing of the software diffusion scale both before and after distribution of the final survey. Experts in the information systems and survey construction areas evaluated possible scale items and suggested scale items of their own. Next, classification procedures were performed with various groups to provide an indication of the unidimensionality of the scale. A number of information system managers then pilot tested the survey containing the diffusion scale to ensure that the items could be answered from a practical viewpoint. The pre-tests are necessarily qualitative because quantitative testing can only be conducted after collection of the survey data.

The survey itself has numerous design features, for example, additional contextual information and numerical scales, to ensure a higher response rate and more accurate responses. Post-testing of the scale included confirmatory factor analysis to evaluate discriminant and convergent validity. Reliability was assessed using a number of techniques, including Cronbach alpha. The scale development process of the diffusion construct is summarised in Figure 1.

If the software diffusion scale is valid, it indicates the items composing the scale measure software diffusion (convergent validity) and not other constructs (discriminant validity). Discriminant validity is only thoroughly tested by comparison of the diffusion items with other scale items thought to measure different factors. This study uses items that measure software infusion, a related construct, to assess the discriminant validity of the items measuring software diffusion. Software infusion is the degree of penetration of software into an organisation’s operations strategy, governance systems and social systems.
Figure 1: Scale development process

- **START**
  - Prepare initial survey items
  - Content validity procedure: expert evaluation
    - Modify instrument?
      - Yes → Redraft survey items
      - No → Sorting procedures
  - Modify instrument?
    - Yes → Redraft survey items
    - No → Conduct interview and pilot tests
      - Modify instrument?
        - Yes → Redraft survey items
        - No
  - **FINAL RESEARCH INSTRUMENT**
    - Tests of Validity (Confirmatory factor analysis)
      - Tests of Reliability
        - **FINISH**
1.2 Motivation

The primary motivation is to provide a more valid and reliable measure of diffusion, and in particular, software diffusion. Diffusion has become an important construct in information systems literature but there is little guidance on measurement of diffusion. Most studies have concentrated on measuring the factors influencing diffusion rather than diffusion itself [23]. Other studies have operationalised the construct using only one item scales [15] or only focused on the adoption aspects of diffusion [6]. Only one prior study provided a well-developed diffusion scale [25] but in that study there was no testing of the scale items prior to distribution of the survey and it focused on hardware diffusion. So the motivation behind this research is to develop a scale for software diffusion that is valid and reliable and that provides a basis for future diffusion research.

The second objective of the research is to suggest an overall method of scale development. Most of the testing procedures used in this study are not new but their use has been isolated. This paper illustrates how using the numerous validation procedures in a logical order facilitates better scale development.

The paper proceeds as follows. The motivation for the study is detailed in Section 1.2. Discussion of the software diffusion construct is provided in Section 2. Section 3 describes the pre-testing procedures for the software diffusion scale and summarises the final items used in the diffusion scale. The next section details survey design and distribution. Post-testing of the construct is discussed in Section 5. Section 6 outlines the limitations and conclusions of the research and considers possible areas for future research.

2. Software diffusion

The level of diffusion of software is manifested by the degree of availability of the software and the number of departments and individuals who use the software. Availability relates to the number of occurrences of the software throughout the organisation. In other words, the number of computers or terminals that run the software. This has sometimes been described as information technology adoption [21]; [18]. Diffusion has been measured by the extent of implementation [16] and as the number of microcomputers per employee [6]. However, adoption does not necessarily mean diffusion within an organisation [15]. The same applies for the extent of implementation of information technology. For example, if the majority of the occurrences of the software are located within a small number of departments, diffusion has not occurred.

Consequently, the three initial items were constructed to measure the availability of the software, i.e., the number of instances of the software as well as the number of users of the software. These items are included in Appendix A.

3. Pre-testing of the diffusion scale

To validate the items measuring the underlying construct, a number of procedures are used. A thorough assessment of validity can only be conducted on the data collected from the survey. However, a number of qualitative tests are used to increase the validity of the scale before its distribution. This section details these validity procedures and summarises the resulting diffusion items that are used in the final survey.

3.1 Content validity procedure

Content validity is achieved when representative items are drawn from a universal pool. Because the pool of items is universal, an instrument valid in content is difficult to verify. One method of verification is to obtain the opinion of experts in the particular research area [10]. The approach used in this study is taken from [13].

 Twelve experts were asked to rate how strongly they believed the items estimated the construct of diffusion. A cover letter detailed the nature of the study and the purpose of the procedure. The respondents were then presented with a definition of diffusion, the items relating to diffusion and a single numerical scale from 1 to 10 with 1 representing a weak estimate and 10 a strong estimate of diffusion (see Appendix B). Respondents were asked to mark the scale depending upon how strongly they believed the items measured software diffusion. A section for any comments or improvements to the items was also provided.

The mean response was 8.00 (median: 8.00; standard deviation: 0.71). Being a qualitative procedure, there is no definitive guidelines as to what minimum score should be achieved. However, the high score does suggest that the items exhibit content validity.

Most of the raters’ suggested improvements relate to clarification of the terms and context of the items. For example, question three was clarified by the phrase “on a regular basis”. A number of additional items were suggested, for example, an item on the degree of training per employee. Most of the items were very similar to the existing items and to ensure that the length of the

---

1 While more than three items would have been preferred, any more questions would have increased the length of the survey, reducing the likely response rate.

2 All respondents were academics with detailed knowledge of information systems and/or survey question design.

3 The same procedure was used for software infusion.
questionnaire did not become unwieldy these additional items were not included.

3.2 Unidimensionality procedure

When discriminant validity and convergent validity are present, the individual items are said to be unidimensional [25]. Unidimensionality means the items load only on one factor. Convergent validity is demonstrated when the items anticipated to load on (explain) a particular factor do actually load on that factor. Discriminant validity is displayed when the items anticipated to load on a particular factor do not also load on other factors.

The testing procedure involved two groups, each of which consisted of four people. In each group, two people were postgraduate honours students, one a senior academic and the fourth an administrator. This range of persons allowed diverse opinions and knowledge backgrounds to be included within the group [23]. One of the academics did have an information system background, but this person was not familiar with the particular area of research. No other respondent had an information systems research background. This selection of group members was done so as to simulate the backgrounds of the actual respondents, who, while having detailed knowledge of the technical aspects of their respective information systems, are unlikely to have a research background. So while assessing unidimensionality, an assessment was also being made of the clarity of the items.

The test for each member of each group was done separately to ensure that no collaboration between participants could occur. Each member of the two groups was given a set of six randomly shuffled cards. Each card had one item typed on it. The respondents from the first group were asked to sort the cards into one or more categories. They were not given the categories of diffusion and infusion. The respondents from the second group were given the categories and definitions of software diffusion and infusion. They were asked to place the cards under the category that they thought appropriate. Respondents were given the option of not putting an item in either category.

The intention of the classification procedure for both groups was to assess convergent and discriminant validity. Convergent and discriminant validity are supported if an item is consistently placed under one category (e.g. diffusion) to the exclusion of the other category (e.g. infusion). Withholding the categories from the first group has the purpose of minimising interpretational confounding. Interpretational confounding occurs “as the assignment of empirical meaning to an unobserved variable [i.e. factor] other than the meaning assigned to it by an individual a priori to estimating unknown parameters” [23].

The results of the unidimensionality procedure for the first group are illustrated in Tables 1 and 2. Items one, two and three constitute the proposed diffusion scale.

All four respondents placed the items in the correct or equivalent categories (see Table 1). For diffusion, respondents used the terms “amount of use”, dispersion or “degree of use”. These terms are synonymous with or very closely related to the concept of diffusion. The third judge was more specific. The diffusion items were separated into staff use and hardware use. Diffusion requires that the spreadsheet software be available within the organisation as well as used by the staff.

Table 2 summarises the inter-rater reliability for each item [23]. The inter-rater reliability is a qualitative measure that indicates how often each item was correctly classified. A high rating demonstrates that a particular item has high discriminant and convergent validity. All diffusion items achieve complete inter-rater reliability.

In the second classification procedure, where the respondents were given the categories of infusion and diffusion, all three diffusion items were correctly classified by all respondents. The findings of the two classification procedures suggest that the items are measuring a distinct construct, namely software diffusion.

3.3 Pre-survey interview and pilot tests

The previous tests enable an assessment of the validity of the items to be included in the survey. However, they do not provide any guidelines on whether the survey could be answered from a practical point of view. Information systems managers may not have the requisite knowledge to properly answer all questions on the survey. An interview was undertaken with an information system manager from an organisation that fell within the definition of the survey sample to assess the ability of information systems managers to respond to the survey. The interviewee and three other respondents were asked to complete the survey and make any comments on the wording or design of the survey.

The pilot tests indicated that respondents could knowledgeably answer the questionnaire. We were present before and after the completion of the pilot tests to answer questions and receive comments on the questionnaire. One respondent suggested minor changes to the layout of the questionnaire that were incorporated into the final draft. The small number of respondents for the pilot test preclude any conclusions being drawn on the reliability or validity of the survey.

3.4 Summary of diffusion items

The testing procedures resulted in a 3-item diffusion scale. The first item comes from a prior study [15] and asks the information systems manager how widespread the software is throughout the organisation. The second item

4 All software infusion items were also correctly classified.
concerns adoption [18]. The third item was developed by the researchers and asks the managers to record the number of staff using the software. The diffusion items are included in Appendix A.

4. Survey design and distribution

This section details the process of survey development used to minimise both sampling and non-sampling errors while increasing the response rate. A number of procedures were employed to reduce the ambiguity of the survey measures and to increase their reliability and validity (see [19]). A mail survey has particular advantages over other research methodologies. It removes the bias associated with evaluation apprehension and experimenter expectancies because the researcher is not present during completion of the questionnaire [9]. It allows respondents to answer the questions in their natural setting, improving external validity [11].

<table>
<thead>
<tr>
<th>Name</th>
<th>Names given for diffusion</th>
<th>Relevant question(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Judge 1</td>
<td>Amount of use; Dispersion</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Judge 2</td>
<td>Amount of use</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Judge 3</td>
<td>Amount of use by staff</td>
<td>1, 3</td>
</tr>
<tr>
<td></td>
<td>Amount of hardware use</td>
<td>2</td>
</tr>
<tr>
<td>Judge 4</td>
<td>Degree of use</td>
<td>1, 2, 3</td>
</tr>
</tbody>
</table>

Table 1: How respondents categorised the three diffusion items

<table>
<thead>
<tr>
<th>Diffusion Name</th>
<th>Question 1</th>
<th>Question 2</th>
<th>Question 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Judge 1</td>
<td>Yes¹</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Judge 2</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Judge 3</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Judge 4</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Inter-rater reliability 100% 100% 100%

¹ A Yes response indicates that the respondent placed the item in the correct or equivalent category.
An Unsure response indicates that the item may fit in the category specified or in the other category or in a third category.
A No response indicates that the respondent believed that the item did not fit in either category.

Table 2: Correct categorisation of the three diffusion items by respondents without categories provided
4.1 Sample

The sample comprised 500 medium-sized\(^5\) Australian private sector firms. The 500 organisations, including both private and public companies and partnerships, were randomly selected from [2] and [22].

Medium-sized firms were selected for practical considerations. Large organisations may have numerous information systems managers that control software-related decisions for their individual departments. Diffusion is then not consistent across departments. The conclusions of this research apply to the individual departments but not necessarily to the firm as a whole because the results would be an average of the individual departments. While inconsistent software purchasing may also occur in medium-sized firms, the likelihood is substantially reduced. The alternative was to survey the department managers but the resources and time required were not available.

Information systems managers were asked to respond to the research instrument because they are more likely to have knowledge of the types of spreadsheet used, the number of computers on which the spreadsheet operates, the number of staff that use spreadsheets and the timing of spreadsheet replacement. Results from an interview conducted with an information systems manager suggest that most managers have sufficient knowledge to complete the instrument.

4.2 Survey instrument

The respondent was required to list all spreadsheet types and versions used, if any, and to answer the remaining questions based on the most widely-used spreadsheet in the organisation. The remaining questions were the scale items for software diffusion and infusion.

There were a number of general features included in the survey to increase its reliability. All frequency scales used numerical scales rather than verbal scales. For example, the item on how many computers run the spreadsheet software had a scale from 0 percent to 100 percent rather than a verbal scale such as “none”, “some”, “most”, “all”. Verbal scales decrease the reliability of the measurement because a respondent’s definition of “most” may differ due to the current situation or through time and experience [19].

Relevant contextual information was included in each question. For example, in the item on what percentage of staff use spreadsheet software, context was provided by inclusion of the word “currently” and the phrase “on a regular basis”. This achieves superior reliability because the respondent is not left uncertain as to the relevant point in time to which the item refers nor what is meant by “use”.

4.3 Administration of the questionnaire

The questionnaire was distributed on Tuesday, 28 May 1996. Tuesday was selected so that most respondents received the questionnaire on Thursday, giving respondents the chance to complete it before the weekend. Due to time and resource constraints no follow-up was employed. To increase the response rate, a number of general practices were implemented [11]. The cover letter was printed on official letterhead of the university and was co-signed by a senior member of the academic staff. Respondents were given an address, telephone number and an e-mail address if they had any problems with the questionnaire. Each cover letter was individually signed. The cover letter and return envelope included an identification number so that the respondent could be recognised but confidentiality maintained.

4.4 Responses

A total of 179 questionnaires were returned. Ten of these questionnaires were not completed because of the organisation’s policy not to respond to questionnaires. The remaining 169 questionnaires result in the relatively large response rate of 33.80%, mitigating the threat of response bias. The usual response rate for information system mail surveys is 20-25% (see [15]). Two questionnaires were deleted from the sample as they did not fit the sample criteria. One of the questionnaires was completed by the information systems manager of the parent company, being a company with revenue exceeding $200 000. The other organisation was in the process of liquidation resulting in unique factors driving the software replacement decision rather than software infusion and diffusion.

Suggested absolute sample requirements for structural equation modelling (confirmatory factor analysis uses structural equation modelling) are 100 [5] or 150 [1]. A suggested relative sample criteria is five times the number of free parameters to be estimated [4]. Because respondents did not answer all questions, the final sample size is 130, which meets both the absolute and relative sample criteria. One outlier was also removed because it contributed most to multivariate kurtosis (non-normality) of the data.

5 Post-testing of the diffusion scale

Post-testing of software diffusion concentrates on statistical methods of determining the validity and reliability of the diffusion construct. The following sections first outline the descriptive statistics of the diffusion items. Confirmatory factor analysis, the primary method of assessing validity, is then described and employed to test the diffusion items. Various methods of evaluating reliability are then described and used.

\(^5\) A medium-sized firm is defined as having annual revenues of between $50 million and $200 million for the financial year ending in 1994.
5.1 Descriptive Statistics

Descriptive statistics for the diffusion items are described in Table 3. Some scales are from zero to ten while others range from zero percent to 100 percent where a percentage was requested. All results are presented as if all questions are in the range from zero to 10.

Normality of the data is a requirement when parametric statistics are used, as is the case for the results presented below. The descriptive statistics and skewness and kurtosis coefficients indicate that all the items are normally distributed.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 1</td>
<td>7.38</td>
<td>8</td>
<td>2.30</td>
<td>0</td>
<td>10</td>
<td>-0.96</td>
<td>0.46</td>
</tr>
<tr>
<td>Question 2</td>
<td>7.03</td>
<td>8</td>
<td>2.76</td>
<td>1</td>
<td>10</td>
<td>-0.88</td>
<td>-0.40</td>
</tr>
<tr>
<td>Question 3</td>
<td>4.59</td>
<td>5</td>
<td>2.61</td>
<td>0</td>
<td>10</td>
<td>0.12</td>
<td>-1.06</td>
</tr>
</tbody>
</table>

5.2 Confirmatory factor analysis

Confirmatory factor analysis is the primary method of ensuring the unidimensionality (convergent validity and discriminant validity) of the factor, software diffusion. The factors and their relevant indicators are illustrated in Figure 2. The analysis necessarily requires both software diffusion and software infusion to be tested concurrently, so results for both are presented, but the focus is on software diffusion.

Confirmatory factor analysis is the appropriate procedure as opposed to exploratory factor analysis when, as is the case here, there is a priori theory supporting the inclusion of certain items with certain factors. Exploratory factor analysis does not provide an explicit test of unidimensionality, because each factor is only defined as the weighted sum of all observed variables rather than indicators corresponding to factors developed by theory.

Confirmatory factor analysis first involves analysing the fit of the model. The fit of the model depends on the level of convergent and discriminant validity. The chi-square statistic is one measure of the goodness of fit, where the null hypothesis is the data fit the model well. A significant chi-square statistic means that the null hypothesis is rejected and implies poor model fit. However, because of the sensitivity of this statistic to even minor discrepancies in model fit especially with large sample sizes, other goodness of fit indices are considered.

One of these alternative measures is an adjusted chi-square statistic, where the chi-square statistic is divided by its degrees of freedom. If the adjusted chi-square statistic is less than five or alternatively less than two the model is considered to be a good fit. Further indices of fit include the normed fit index (NFI), the non-normed fit index (NNFI) and the comparative fit index (CFI). The recommended lower threshold is 0.90 for the three indices.

Constraining the correlation between the factors of diffusion and infusion to unity provides a test of discriminant validity. A significant degradation in model fit with the unity constraint imposed supports the conclusion that there are separate factors rather than simply one factor. The relevant statistic is the difference between the chi-square statistics of the constrained and unconstrained models.

Two further tests of convergent and discriminant validity are the Wald test and the Lagrange Multiplier test. The Wald test gives a list of parameters that when fixed, improve model fit.

---

6 Unidimensionality can be expressed mathematically. First, the equation for a single indicator (factor) is:

\[ x_i = \lambda_i \xi + \delta_i \]

where \( x_i \) is the ith indicator from a set of unidimensional indicators, \( \lambda_i \) is the corresponding factor loading and \( \delta_i \) is the corresponding error term.

Assessment of unidimensionality is based on two product rules that follow from this mathematical definition.

The rule of internal consistency (convergent validity):

\[ \rho_{ij} = \rho_{i\xi} + \rho_{j\xi} \]

This rule states that the correlation of two items \( i \) and \( j \) which are alternative measures of a single underlying construct \( \xi \) is equal to the product of their respective factor loadings.

The rule of external consistency (discriminant validity):

\[ \rho_{ij} = \rho_{i\xi} \rho_{j\xi} \]

The correlation of two items \( i \) and \( p \) which are measures of two different constructs \( \xi \) and \( \xi^* \) is equal to the product of item \( i \)'s loading on factor \( \xi \) indicator \( p \)'s loading on factor \( \xi^* \) and the correlation between factor \( \xi \) and \( \xi^* \).

7 EQS for Windows Version 5.1, a structural equation modelling program, is used to perform the analysis.
The Lagrange Multiplier test has an opposite function. It provides a list of parameters that when free to vary improve model fit.

The final test to assess unidimensionality is to analyse the path co-efficients. The path coefficients are the standardised regression coefficients derived from the structural equations implicit in the model illustrated in Figure 2. A z-statistic provides an assessment of the significance of the path coefficients.

Unidimensionality involves an assessment of whether each survey item loads only on the factor predicted by a theoretical model. In this study, questions 1, 2 and 3 should load only on the software diffusion factor. The results of the confirmatory factor analysis are illustrated in Figure 3. The results given are the standardised path coefficients and they all are significant at the 1% level. The $\chi^2$-statistic of 16.28 (df=8, $p < 0.05$) suggests that the model is not a good fit. However, the chi-squared statistic is sensitive to large sample sizes and alternative measures may suggest a better fit. Dividing the $\chi^2$-statistic by its degrees of freedom yields a test statistic of 2.04, slightly above the maximum recommended level of 2 but well below 5, the other suggested threshold [8]; [26]. The normed fit index (NFI), the non-normed fit index (NNFI) and the comparative fit index (CFI) are all above the recommended 0.90 level. These results imply a reasonable model fit.

The diffusion items have acceptable factor loadings with only question 5 being slightly below the 0.70 threshold. However, to ensure unidimensionality of the diffusion items, the items from the software diffusion and infusion constructs must be compared. The results indicate that the infusion scale items do not exhibit unidimensionality. First, questions 5 and 6 are well below the recommended 0.70 minimum for factor loadings, suggesting that these questions have low convergent validity. This is reinforced by the Wald test statistic that suggests that the error term associated with question four be fixed at unity, meaning that infusion is really only a reflection of question four rather than a combination of the three items.

The Lagrange Multiplier test for adding parameters suggests that paths should be drawn from questions 1 and 2 to the infusion factor, indicating that these items may have low discriminant validity. This result is reinforced when the covariance between the diffusion and infusion factors is

$\chi^2 = \sum (\hat{\beta} - \beta)^2 / \sigma^2$

[8] The structural equations required for the confirmatory factor analysis in this study are:
\[
\begin{align*}
\xi_1 &= \lambda_1 x_1 + \delta_1 \quad (a) \\
\xi_2 &= \lambda_2 x_2 + \delta_2 \quad (b) \\
\xi_3 &= \lambda_3 x_3 + \delta_3 \quad (c) \\
\xi_4 &= \lambda_4 x_4 + \delta_4 \quad (d) \\
\xi_5 &= \lambda_5 x_5 + \delta_5 \quad (e) \\
\xi_6 &= \lambda_6 x_6 + \delta_6 \quad (f)
\end{align*}
\]

[9] The NFI, NNFI and CFI were 0.94, 0.94 and 0.97, respectively.

[10] Other sub-samples taken from the survey data have a factor loading for question 5 greater than 0.70, reinforcing its retention in the diffusion scale. Question 5, if omitted, results in the diffusion construct being unidentified.
constrained to unity. The constrained model has a $\chi^2$-statistic of 18.05 (df=9, $p<0.05$), only slightly worse than the unconstrained model. The difference between the two $\chi^2$-statistics is 1.78 (df=1, ns) meaning that the null hypothesis that there is only one factor (rather than the two hypothesised) cannot be rejected.

To summarise, the confirmatory factor analysis suggests that there is only one factor, with questions 1, 2, 3, and 4 loading onto that one factor. However, this is probably an artifact of the low convergent validity exhibited by questions 5 and 6 and question 4 does not relate to the dispersion of the software throughout the organisation. The non-inclusion of question 4 means that the diffusion scale comprises questions 1, 2 and 3.

![Figure 3: Confirmatory factor analysis results](image)

5.3 Reliability of the diffusion scale

In addition to assessing the validity of the results, the reliability of the data is also relevant. Cronbach or coefficient alpha is one method used to statistically assess the reliability of the survey data. It is not a test of unidimensionality but rather assumes unidimensionality. Other methods of assessing reliability used in this study are a measure of composite factor reliability\(^\text{11}\) and a measure known as Average Variance Extracted (AVE) \([14]\).\(^\text{12}\)

Composite factor reliability measures whether there is a sufficient relationship between the scale items and their respective constructs. 0.70 is the suggested lower threshold before reliability can be assumed \([20]\). Average variance extracted (AVE) indicates the amount of variance that is captured by the factor as opposed to the variance due to error. 0.50 is the suggested cut-off for AVE \([25]\).

The composite factor reliability of the diffusion construct is 0.82, well above the suggested 0.70, indicating that there is a sufficient relationship between the scale items and the diffusion construct. The average variance (AVE) of the diffusion items (0.60) exceeds the recommended 0.50 meaning that the items explain more of the variation in the factor than error. An alternative measure of reliability is Cronbach alpha, which for the diffusion items is 0.82. The results suggest that the scale items for the diffusion items are highly reliable.

6. Limitations and Conclusions

The software diffusion scale exhibits both unidimensionality and reliability, indicating that the scale items do measure software diffusion. This section details possible limitations of the development of the software diffusion scale and then offers some concluding comments.

6.1 Limitations

One limitation of the use of the survey as a sampling technique is that it may result in reduced internal validity. This is because the survey solicits responses that are...
subjective in nature. It is a form of reactive measurement [17] because it focuses attention on the items being measured and also because respondents may attempt to give the perceived ‘correct’ answers. Further, the response bias is accentuated by the varied experience and knowledge of the respondent, the setting or timing of the survey, the instructions given, and by the items themselves [19]. These limitations have been attenuated by careful selection of respondents, the contextual cues provided within the questionnaire and the pre-testing procedures to ensure the comprehensibility of the questions. However, these procedures do not overcome the possibility that the correct person did not answer the questionnaire.

The trade-off between a higher response rate and reduced unidimensionality constitutes the principal limitation of the scale development. By not including additional items to measure infusion and diffusion, there was no scope to drop particular items from either construct without encountering the problem of model underidentification. Additional variables could also have been included if the response rate was not such a concern. However, a small sample lessens the reliability of some statistical tests, justifying the reduced length of the questionnaire.

6.2 Conclusions

This study describes the development of a scale to measure the construct, software diffusion. It described a range of both qualitative and quantitative tests conducted both before and after distribution of the survey. Pre-survey testing is important to increase the probability that the scale items are unidimensional and reliable. An ancillary purpose of the research is to outline procedures that increase and analyse the validity and reliability of survey scales. Further research avenues may include development of scales on hardware diffusion or use other types of software, for example, software developed within the organisation. The results suggest that software diffusion is a function of not only the number of instances of the software, but also its level of use by staff within the organisation. We believe this scale provides researchers with a robust means of measuring software diffusion in firms – an important area of future research.

7. References


Appendix A: The software diffusion and software infusion scales

**Software Diffusion**

1. In your opinion, how widespread is the use of spreadsheet software throughout your organisation?  
   [Scale: 0 (Not widespread) to 10 (Very widespread)]

2. Approximately what percentage of terminals or computers in your organisation currently run the spreadsheet software?  
   [Scale: 0% to 100% or unknown]

3. Approximately what percentage of the staff in the organisation currently use the spreadsheet software on a regular basis?  
   [Scale: 0% to 100% or unknown]

**Software Infusion**

4. To what extent has the spreadsheet software penetrated your organisation in terms of importance, impact or significance?  
   [Scale: 0 equals no penetration, up to 10 which corresponds to penetration to a large extent]

5. Approximately what percentage of use of the spreadsheet software would involve pre-saved information, functions, templates or macros?  In other words, use of additions to the software, such as macros, that were not included with the original software, but designed specifically by or for the staff in your organisation.  
   [Scale: 0% to 100% or unknown]

6. Approximately what percentage of work tasks be completed if the spreadsheet software and related files became unavailable? For example, if the users could not complete 80% of the things they usually do at work, the answer would be 20%. The unavailability could be due to a power failure, a system crash or maintenance being performed.  
   [Scale: 0% to 100% or unknown]

Appendix B: Content validity procedure for diffusion

**Diffusion of Spreadsheet Software**

Definition:  
Diffusion is the process by which an innovation is communicated through certain channels over time among members of a social system [24]. In an information technology context, it relates to the degree of dispersion of the information technology (in this case - spreadsheet software) throughout the organisation.

By placing a mark along the line shown below, please indicate the extent to which you believe the group of statements below, if answered by a respondent to the questionnaire, estimate the extent of spreadsheet software diffusion in an organisation. Please provide any explanations or other comments on the following page.  

(Note: Before placing a mark along the line, carefully read through each question. Do not answer the questions themselves)

<table>
<thead>
<tr>
<th>Weak estimate of spreadsheet software diffusion</th>
<th>Strong estimate of spreadsheet software diffusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
</tbody>
</table>

[Place three diffusion items here]